Atmospheric Environment 115 (2015) 553–568

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Influence of the choice of gas-phase mechanism on predictions of key gaseous pollutants during the AQMEII phase-2 intercomparison

Christoph Knote [a, *], Paolo Tuccella [b], Gabriele Curci [b], Louisa Emmons [a], John J. Orlando [a], Sasha Madronich [a], Rocio Baró [c], Pedro Jiménez-Guerrero [c], Deborah Luecken [d], Christian Hogrefe [d], Renate Forkel [e], Johannes Werhahn [e], Marcus Hirtl [f], Juan L. Pérez [g], Roberto San José [g], Lea Giordano [h], Dominik Brunner [h], Khairunnisa Yahya [i], Yang Zhang [i]

[a] Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, CO, USA
[b] Department of Physical and Chemical Sciences, Center of Excellence for the Forecast of Severe Weather (CETEMPS), University of L'Aquila, L'Aquila, Italy
[c] University of Murcia, Department of Physics, Physics of the Earth, Campus de Espinardo, Ed. CIOyN, 30100, Murcia, Spain
[d] Atmospheric Modelling and Analysis Division, Environmental Protection Agency, Research Triangle Park, USA
[e] Karlsruher Institut für Technologie (KIT), Institut für Meteorologie und Klimaforschung, Atmosphärische Umweltforschung (IMK-IFU), Kreuzeckbahnstr. 19, 82467, Garmisch-Partenkirchen, Germany
[f] Section Environmental Meteorology, Division Customer Service, ZAMG – Zentralanstalt für Meteorologie und Geodynamik, 1190, Wien, Austria
[g] Environmental Software and Modelling Group, Computer Science School – Technical University of Madrid, Campus de Montegancedo, Boadilla del Monte, 28660, Madrid, Spain
[h] Laboratory for Air Pollution and Environmental Technology, Empa, Dübendorf, Switzerland
[i] Department of Marine, Earth, and Atmospheric Sciences, 2800 Faucette Drive, #1125 Jordan Hall, Campus Box, 8208, North Carolina State University, USA

HIGHLIGHTS

- We intercompared tropospheric gas-phase mechanisms used in AQMEII phase 2.
- Box model results show O_3 differs by 4 ppbv (5%), NO_x 25%, isoprene >100% and HCHO 20%.
- Key radicals OH and HO_2 differ 40 / 25% between mechanisms, NO_3 by more than 100%.
- Uncertainty due to gas-phase mechanism choice has to be considered in model simulations.

ARTICLE INFO

Article history:
Received 9 June 2014
Received in revised form
27 November 2014
Accepted 29 November 2014
Available online 2 December 2014

Keywords:
Air pollution
Box modeling
Gas-phase mechanisms
Tropospheric chemistry
Model intercomparison
AQMEII

ABSTRACT

The formulations of tropospheric gas-phase chemistry ("mechanisms") used in the regional-scale chemistry-transport models participating in the Air Quality Modelling Evaluation International Initiative (AQMEII) Phase 2 are intercompared by the means of box model studies. Simulations were conducted under idealized meteorological conditions, and the results are representative of mean boundary layer concentrations. Three sets of meteorological conditions – winter, spring/autumn and summer – were used to capture the annual variability, similar to the 3-D model simulations in AQMEII Phase 2. We also employed the same emissions input data used in the 3-D model intercomparison, and sample from these datasets employing different strategies to evaluate mechanism performance under a realistic range of pollution conditions.

Box model simulations using the different mechanisms are conducted with tight constraints on all relevant processes and boundary conditions (photolysis, temperature, entrainment, etc.) to ensure that differences in predicted concentrations of pollutants can be attributed to differences in the formulation of gas-phase chemistry. The results are then compared with each other (but not to measurements), leading to an understanding of mechanism-specific biases compared to the multi-model mean. Our results allow us to quantify the uncertainty in predictions of a given compound in the 3-D simulations introduced by the choice of gas-phase mechanisms, to determine mechanism-specific biases under certain pollution conditions, and to identify (or rule out) the gas-phase mechanism as the cause of an observed discrepancy in 3-D model predictions.

* Corresponding author. Atmospheric Chemistry Division, NCAR, 3450, Mitchell Lane, Boulder, CO, USA.
E-mail address: knote@ucar.edu (C. Knote).

http://dx.doi.org/10.1016/j.atmosenv.2014.11.066

We find that the predictions of the median diurnal cycle of O_3 over a set of emission conditions representing a network of station observations is within 4 ppbv (5%) across the different mechanisms. This variability is found to be very similar on both continents. There are considerably larger differences in predicted concentrations of NO_x (up to ± 25%), key radicals like OH (40%), HO_2 (25%) and especially NO_3 (>100%). Secondary substances like H_2O_2 (25%) or HNO_3 (10%), as well as key volatile organic compounds like isoprene (>100%) or CH_2O (20%) differ substantially as well. Calculation of an indicator of the chemical regime leads to up to 20% of simulations being classified differently by different mechanism, which would lead to different predictions of the most efficient emission reduction strategies.

All these differences are despite identical meteorological boundary conditions, photolysis rates, as well as identical biogenic and inorganic anthropogenic emissions. Anthropogenic VOC emissions only vary in the way they are translated in mechanism-specific compounds, but are identical in the total emitted carbon mass and its spatial distribution.

Our findings highlight that the choice of gas-phase mechanism is crucial in simulations for regulatory purposes, emission scenarios, as well as process studies that investigate other components like secondary formed aerosol components. We find that biogenic VOCs create considerable variability in mechanism predictions and suggest that these, together with nighttime chemistry should be areas of further mechanism improvement.

1. Introduction

The large number of unconstrained degrees of freedom in state-of-the-art regional chemistry-transport-models (CTMs) severely hinders knowledge gain through the evaluation of the result from 3-D simulations. Differences in the processes included and the way they are parameterized, unknown initial and boundary conditions for trace gas concentrations and aerosol properties, as well as unconstrained feedbacks among meteorology, aerosols and trace gases offer a range of explanations for an observed discrepancy between model results and measurements, and picking the most plausible one is often at the discretion of the observer. One way to overcome this problem is to disassemble the modeling system and compare core components (such as the set of gas-phase chemical reactions) in a tightly constrained setting, so as to characterize one model's component via comparison with that of its peers – eliminating all other sources of differences between models. This "diagnostic evaluation" (e.g., Zhang et al., 2006; Dennis et al., 2010) aims at understanding the reasons for differences between model performance.

In this work we evaluate the contribution of the choice of a gas-phase mechanism to predictions of key pollutants during the Air Quality Modelling Evaluation International Initiative Phase 2 (AQMEII, Alapaty et al., 2012). Most of the formulations of photochemistry used in current regional CTMs have been evaluated against chamber data upon inception (e.g., Stockwell et al., 1990; Yarwood et al., 2005; Carter, 2010) and a large body of literature exists comparing these gas-phase mechanisms (e.g., Kuhn et al., 1998; Gross and Stockwell, 2003; Luecken et al., 2008; Kim et al., 2011; Zhang et al., 2012). The regional CTMs participating in AQMEII Phase 2 and the respective gas-phase mechanism implementations were used and evaluated in previous projects (e.g., Grell et al., 2005; Wyat Appel et al., 2007; Knote et al., 2011). This does not provide an understanding, though, of the exact performance of a certain gas-phase mechanism compared to its peers within the AQMEII phase 2 intercomparison. Mechanisms were developed further since their last peer-reviewed intercomparison, the very set of mechanisms used in AQMEII Phase 2 has never been compared directly, and most previous mechanism intercomparisons were conducted under idealized emission conditions, or were set up to mimic chamber experiments.

In the work presented here we extracted the gas-phase chemical mechanisms used in AQMEII Phase 2 and intercompared their performance under tight constraints for environmental parameters, photolysis rates, removal processes and emissions. Box model simulations were made that represent mean boundary layer concentrations, and different sets of conditions (photolysis rates, temperature, diurnal evolution of the boundary layer height) were used to describe different seasons. We show differences in predictions of key pollutants like O_3 or NO_x and relate them to the pollution situation (NO_x, VOC levels) as well as the spatial distribution over the model domain. By conducting box model simulations using emissions at the same locations of observations used also in the analysis of the 3-D simulations we can attribute findings from the 3-D evaluation to the gas-phase mechanism or determine possible masking of mechanism performance by other processes. We identify reasons for discrepancies among the mechanisms due to the mechanism formulation and provide suggestions for mechanism improvement.

2. Box modeling

A total of 7 mechanisms, variants or updated versions of the Carbon Bond Mechanism (CBM, Whitten et al., 1980), the chemical mechanism of the Regional Acid Deposition Model (RADM, Stockwell, 1986) and the MOZART mechanism (Brasseur et al., 1998), were used in the intercomparison (Table 1). All have been developed for the description of tropospheric gas-phase chemistry with a focus on the description of key pollutants like O_3. All contain an explicit description of inorganic chemistry of NO_x (NO and NO_2) and a more or less condensed representation of the chemistry of reactive volatile organic compounds (VOCs) required for a realistic representation of radical cycling (especially OH, HO_2). Halogen chemistry is not considered in our simulations even in the mechanisms that contain such reactions. Any heterogenous reactions (e.g. on aerosol surfaces) are switched off as well.

A number of errors were found in the implementation of some of the mechanisms during our analysis (see Section S4 of the Supplementary material) and were reported back to the respective groups.

All mechanisms were provided by the participants in the form of Kinetic PreProcessor (KPP) models (Sandu and Sander, 2006), text descriptions of reactions and rate constants which are then converted by KPP into a set of Fortran 90 routines that allow for numerical integration of the system over time. The KPP descriptions of all mechanisms investigated can be found in the Supplementary

Table 1
List of gas-phase chemical mechanisms, groups, and modeling systems participating in the AQMEII intercomparison. "Group(s)" here refers to the participant(s) in the AQMEII phase 2 intercomparison using that mechanism. In case of multiple participants, in bold the group that provided us with the mechanism and the emissions they used.

Mechanism	Group(s)	3-D model
RADM2 (Stockwell et al., 1990)	**U. Murcia**; Karlsruhe Institute of Technology (KIT/IMK-IFU); Zentralanstalt für Meteorologie und Geodynamik (ZAMG); University of Ljubljana, SPACE-SI	WRF-Chem
RADMKA (Vogel et al., 2009)[a]	Eidgenössische Materialprüfungsanstalt (EMPA)	COSMO-ART
RACM-ESRL (Stockwell et al., 1997)[b]	U. L'Aquila	WRF-Chem
CB05Clx (Yarwood et al., 2005; Sarwar and Yarwood, 2006; Karamchandani et al., 2012)	North Carolina State University (NCSU)	WRF-Chem
CB05-TUCL (Yarwood et al., 2005; Whitten et al., 2010)[c]	U.S. Environmental Protection Agency (EPA)	WRF-CMAQ
CBMZ (Zaveri and Peters, 1999)	U. Politécnica de Madrid	WRF-Chem
MOZART-4 (Emmons et al., 2010; Knote et al., 2014)	National Center for Atmospheric Research (NCAR)	WRF-Chem

[a] Extended version of RADM2 with updated reaction rates, more detailed heterogenous reactions of N_2O_5 and HONO, a more detailed isoprene scheme (Geiger et al., 2003), and additional hydrocarbons.
[b] Extended version of RACM with a more detailed isoprene scheme (Geiger et al., 2003) and further updates by NOAA Earth System Research Laboratory.
[c] http://www.airqualitymodeling.org/cmaqwiki/index.php?title=CMAQv5.0_Chemistry_Notes.

material for reference. The Rosenbrock method with adaptive time stepping available in KPP (Sandu and Sander, 2006) is used for the numerical integration over time for all mechanisms. The mechanisms that have been implemented in WRF-Chem using KPP employ the same integration method. Participants that used WRF-Chem with the RADM2 mechanism in the 3-D model intercomparison, however, employed a different solver (quasi steady state approximation). For the sake of consistency we here also use the Rosenbrock solver for RADM2. All box model simulations were carried out on the Yellowstone computing system at the National Center for Atmospheric Research (Computational and Information Systems Laboratory, Boulder, CO: National Center for Atmospheric Research, 2012), and the R-project software (http://www.r-project.org, last accessed 09/13/2014) was used for postprocessing.

The "box" is assumed to be stationary horizontally (not changing location over time), have a constant horizontal area and extend vertically from the surface up to the top of the (time-varying) boundary layer. It thereby forms a time-varying mixing volume in which realistic emission amounts as used in the 3-D simulations can be injected. We further represent an idealized diurnal cycle through entrainment during boundary layer rise in the morning and decoupling of a residual layer at nightfall. The box is considered to be always and instantaneously well-mixed. Resulting concentrations hence represent the mean boundary layer value. Each simulation starts at midnight (local time) and is run for 7 days (168 h).

2.1. Initial conditions

Initial conditions for inorganic species are shown in Table 2. To reflect clean and polluted conditions the values for NO_x scale with the emission situation (the higher the NO_x emission, the higher the initial conditions of NO_x). The fraction NO/NO_x is set at 10% for the initial conditions. All species not explicitly mentioned in Table 2 are set to 0. This is reasoned by the fact that the remaining species are VOCs, whose representation differs between the mechanisms. While we acknowledge that this will create a bias for longer-lived VOCs, providing (arbitrary) boundary conditions for these species would likewise create a bias.

Table 2
Values for initial conditions used in box model simulations. Values for NO and NO_2 vary depending on emission situation. These values are also used for top entrainment (see text).

Species	Concentration (ppbv)
H_2O	1E + 07
CH_4	1750
CO	120
O_3	30
NO	0–1
NO_2	0–10
SO_2	1
N_2O	320
All others	0

2.2. Temperature and planetary boundary layer height

To represent the different meterological conditions that occur over the course of a full year (which is the time-frame of the 3-D intercomparison) we conducted box model simulations with three different sets of temperature, boundary layer height and photolysis rates (see next Section), depicted in Fig. 1:

- 'Winter': a shallow boundary layer, temperatures around freezing, and cloudy skies.
- 'Spring/autumn': warmer temperatures, a stronger diurnal cycle in PBLH, and stronger insolation.
- 'Summer': conditions of a clear-sky summer day, large diurnal cycle of the PBLH and warm temperatures.

Temperature ranges between 288.15 and 303.15 K ('summer')/278.15 and 293.15 K ('spring/autumn')/268.15 and 278.15 K ('winter'), following the solar zenith angle calculated by the photolysis module (see next paragraph) with a time lag of 1 h. The only effect of changing temperatures considered is the change in temperature-dependent reaction rates. The diurnal cycle of the height of the planetary boundary layer is modeled after boundary layer textbook knowledge (Stull, 1988). The PBLH ranges from 250 to 1500 m ('summer')/100 to 750 m ('spring/autumn')/50 to 500 m ('winter'), and follows the solar zenith angle to describe the morning rise (without time lag). During this rise, the additional volume engulfed by the box is considered to be entrained air from above, which contains the same concentrations as used for initial conditions (Table 2) – hence long-lived (inorganic) species are entrained with realistic concentrations, while short-lived VOCs are diluted against zero air. In the evening the PBLH stays at the maximum height until sunset, and then drops to the minimum height within an hour. This boundary layer collapse does not change concentrations, but affects the mixing volume available for new emissions and increases the effect of dry deposition (see 2.4). It hence represents a decoupling of the surface from the air above (the residual layer). Diurnal cycles are repeated for the duration of the simulation.

2.3. Photolysis rates

Photolysis rates are calculated by a stand-alone simulation of radiative transfer using the Tropospheric Ultraviolet and Visible

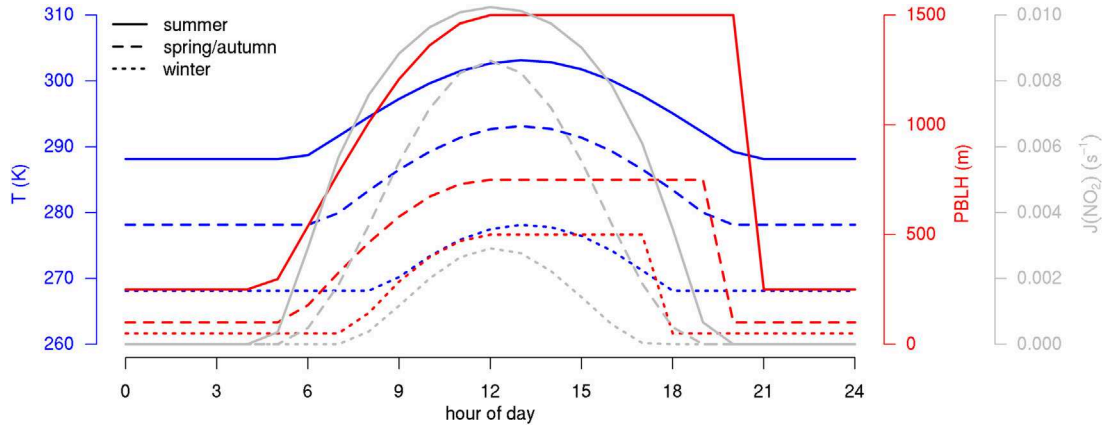

Fig. 1. Diurnal variation of temperature (blue), height of the boundary layer (red) and photolysis rates of NO_2 for the three conditions investigated. 'Summer' (solid), 'spring/autumn' (dashed), and 'winter' (dotted lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(TUV) radiation model (Madronich and Flocke, 1997) in version 5.1. Hourly photolysis rates are calculated for sea-level, no-aerosol conditions at 40° N on 14th July/14th April/14th January 2010 ('summer', 'spring/autumn', 'winter'), a cloud optical depth of 0/5/10, and an overhead O_3 column of 325 Dobson units using the U.S. Standard Atmosphere 1976. For lumped/surrogate VOCs the photolysis rates are usually approximated in the mechanisms as (a fraction of) the photolysis rates of (similar) explicit compounds, hence we provide the same information from the TUV standalone simulation. For rates of explicit compounds not included in TUV we obtained cross-sections and quantum yields from the groups and added those rate calculations to TUV.

2.4. Removals

Removal is considered indirectly through the diurnal cycle of the boundary layer (entrainment, cutoff of the residual layer), and additionally as time-invariant dry deposition first order loss process for the species listed in Table 3 in the form of

$$\frac{\partial C_k}{\partial t} = -v_{dep,k} \cdot C_k / \mathrm{PBLH} \quad (1)$$

with C_k the concentration of a given species k in molecules m^{-3}, t the time in seconds, $v_{dep,k}$ the deposition velocity in m s^{-1}, and PBLH the height of the planetary boundary layer in m.

2.5. Emissions

We test the mechanisms under a realistic range of emission conditions. Therefore, we employ the exact emission input (2-D time-varying fields of emission fluxes) used in the 3-D model intercomparison, and then sample from these datasets using different strategies. It is our intention to drive all mechanisms with identical emissions to ensure differences found in simulation results can be attributed to the gas-phase mechanism only. This consistency is easily achieved for all types of emissions but anthropogenic VOC emissions, which are mechanism-specific, as different approaches are taken to lump VOCs into groups/surrogates. The following paragraph outlines our approach.

Anthropogenic emissions for AQMEII Phase 2 were provided by U.S. Environmental Protection Agency (EPA) (for North America) and the Netherlands Organization for Applied Scientific Research (TNO) (for Europe) and are described in detail in Pouliot et al. (2015). Emission totals, their hourly evolution over time and their spatial distribution (horizontal resolution of 12 × 12 km (U.S.)/0.125 × 0.0625° (Europe)) were hence identical for all groups. Participating groups were allowed to choose their horizontal coordinate system and grid resolution, and were responsible for regridding the emissions provided onto their grid. Each group also had to convert emissions of non-methane volatile organic compounds (NMVOCs) to the (surrogate or lumped) species available in their respective mechanism.

Table 3
Dry deposition velocities considered (Hauglustaine et al., 1994; Zhang et al., 2003).

Species	Deposition velocity (cm s^{-1})
CO	0.03
SO_2	1.8
NO	0.016
NO_2	0.1
NO_3	0.1
N_2O_5	4
HONO	2.2
HNO_3	4
HNO_4	4
O_3	0.4
H_2O_2	0.5
CH_3OOH	0.1

In the 3-D model intercomparison, biogenic emissions were calculated independently by each group (and on their own grid and resolution) according to schemes that depend on land-use characterization and meteorology like the Model of Emissions of Gases and Aerosols from Nature (MEGAN, Guenther et al., 2006, http://lar.wsu.edu/megan/, last accessed 18/08/2014) or the Biogenic Emission Inventory System (BEIS, http://www.epa.gov/ttnchie1/emch/biogenic/, last accessed 18/08/2014).

We asked participating groups to provide us with files describing hourly emissions in their respective native model resolution and lumping of anthropogenic VOCs for the 14th of July 2010. The date was randomly picked and only ensures that we receive comparable emissions from the different groups. Using files in their native format ensured that we receive the exact input to their modeling system, including all changes made through interpolation and lumping.

To achieve consistency for explicit (inorganic) compounds, a base model is chosen for each continent to provide emissions of inorganic compounds (CO, SO_2, NO, NO_2, HONO, NH_3), methane, and biogenic VOCs (isoprene, α-/β-pinene, or total monoterpenes defined as the sum of α- and β-pinene). The base model for North

America was chosen to be the group from NCAR (MOZART-4; WRF-Chem), whereas U. L'Aquila (RACM; WRF-Chem) was chosen for Europe.

Anthropogenic VOC emissions were used from the emission files provided by each group. Each mechanism requires a different grouping of VOCs, and hence we could not use one set of emission files for all mechanisms. Different approaches were taken by the groups to get anthropogenic VOC emissions speciated for their mechanism, ranging from involved emissions modeling including source- and region-specific activity and composition data (e.g., using the Sparse Matrix Operation Kernel Emissions, SMOKE, University of North Carolina (2003)) and subsequent lumping onto a certain mechanism, to the use (and re-speciation) of emissions data already gridded onto a certain model grid and lumped onto a specific mechanism. Nonetheless, they are all based on the same total amount of NMVOC emissions provided by U.S. EPA/TNO. Figs. S20 and S21 in the Supplementary material show that there are no major differences in the amount of anthropogenic VOCs emitted by the different mechanisms.

Vertical distribution of emissions was ignored, all area emissions and anthropogenic point sources are summed up vertically. Emissions from wildfires were not considered. For each simulation specific locations were picked from the native resolution files – interpolating or averaging depending on the type of analysis: for the comparison using the emission situation at station locations (see next Section), emission amounts are the result of distance-weighted interpolation of the closest 4 model grid points. Differences in spatial resolution of the interpolated emissions between groups are not removed. For comparisons where we intended to remove the effect of grid resolution (i.e. the "raster" approach), averaged emissions from all grid points falling into a $1 \times 1^{\circ}$ box are used.

Anthropogenic emissions are identical for the simulations under different environment conditions ('winter', 'spring/autumn', 'summer'), whereas biogenic emissions are scaled from the original MEGAN input (scaling factors. 'winter': 0.01, 'spring/autumn': 0.1, 'summer': 1.0).

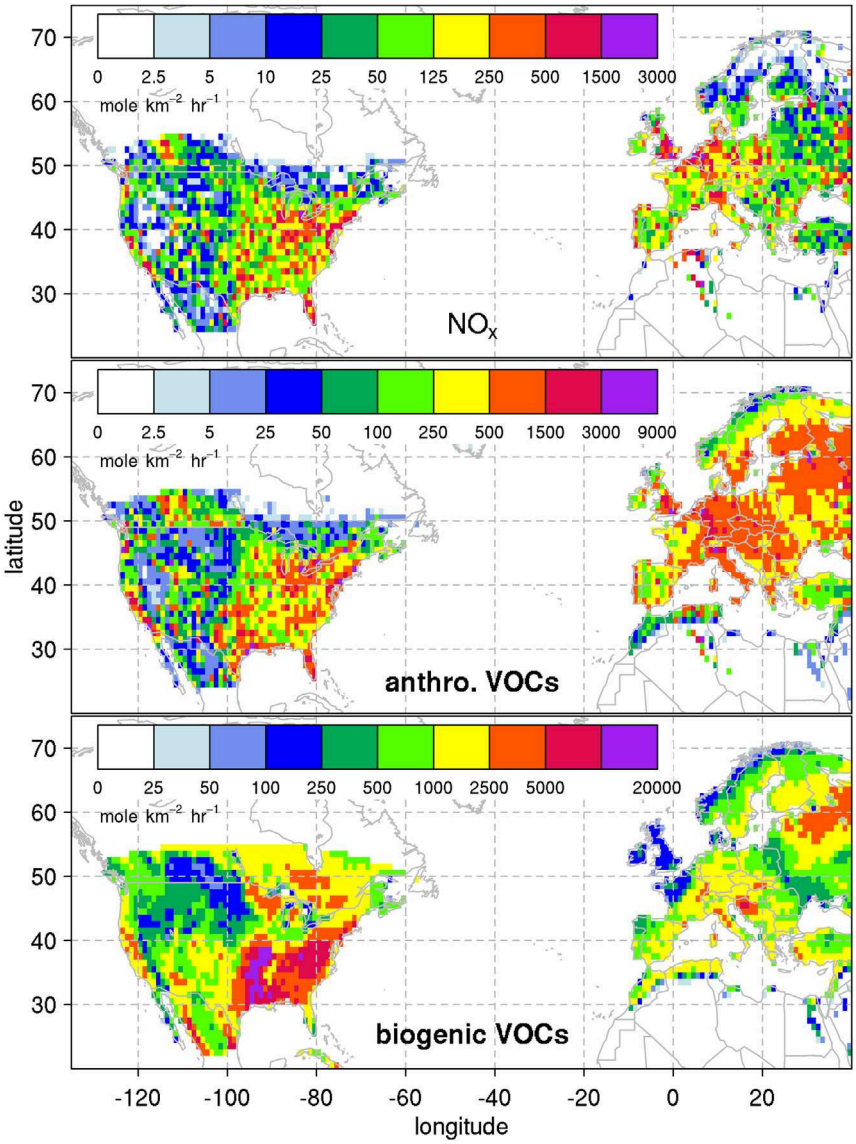

Fig. 2. $1 \times 1^{\circ}$ averaged hourly emissions on 14 July of NO_x ($NO + NO_2$, top plot), anthropogenic VOCs (sum of emissions in MOZART-4 for North America, RACM in Europe, in mole C, middle plot) and biogenic VOCs (sum of isoprene, α-, β-pinene from MEGAN simulation in mole C, bottom plot).

As guidance we provide time-averaged emission maps for NO_x, total anthropogenic and biogenic VOCs in Fig. 2. We present diurnal cycles of emissions for all mechanisms for select locations in Section S5 of the Supplementary material.

3. Intercomparison

The following section presents an intercomparison of the different mechanisms within the context of AQMEII Phase 2. In each subsection we analyze metrics (regulatory figures, averaged diurnal cycles, indices) of select pollutants simulated by the different mechanisms under a range of emission situations by sampling emissions in the AQMEII domains with different strategies. We conducted simulations under different environmental conditions to evaluate the range of conditions in all seasons.

For each sampling point we made 7-day long (168 h) box model simulations with diurnally varying temperature, PBL height and emissions as explained above. In our analysis, differences between the mechanisms are analyzed with the multi-mechanism mean as reference. The first 24 h of each run are ignored as they represent a

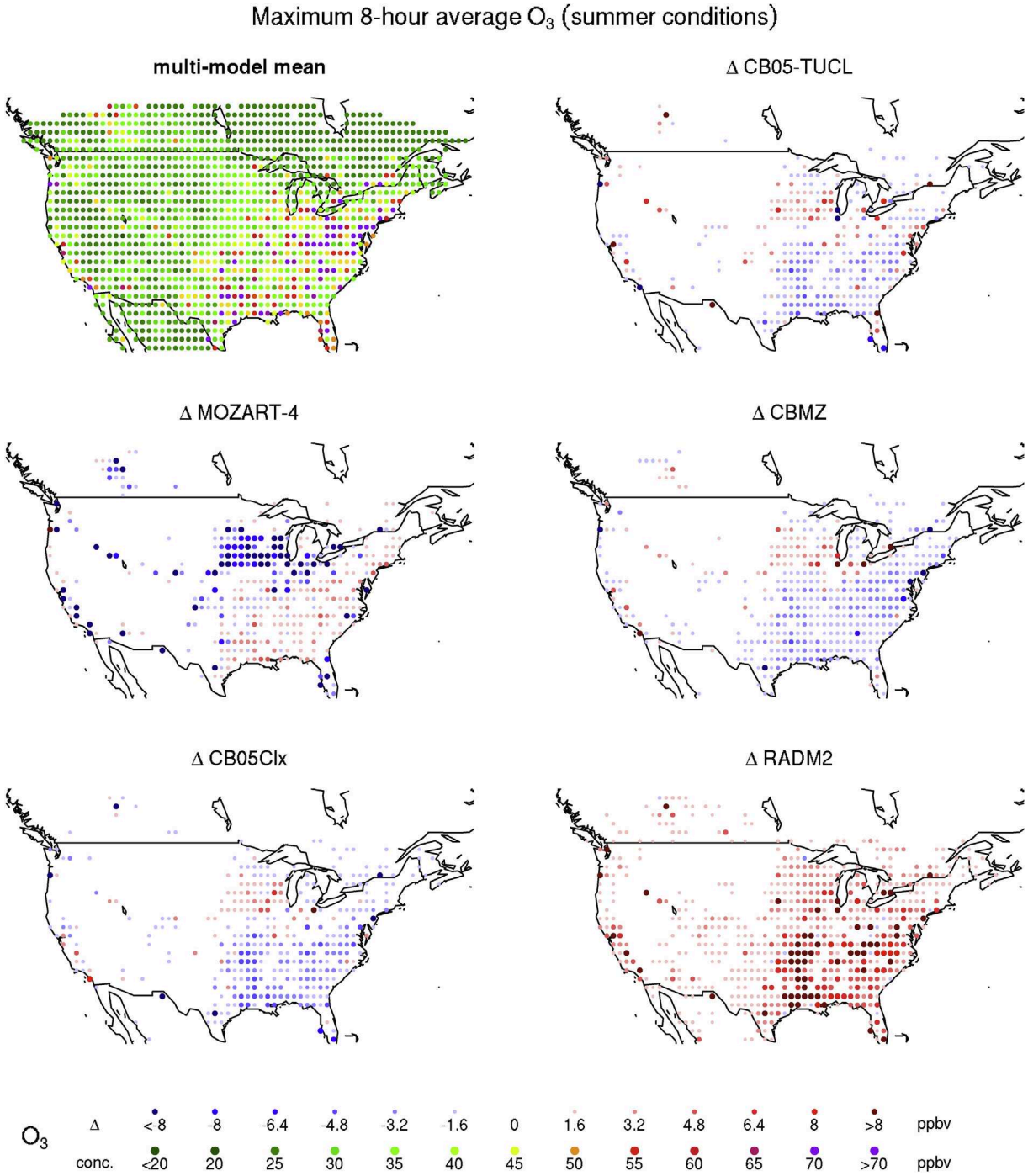

Fig. 3. Top left: sampling locations for the raster approach, color-coded by O_3 concentration but not scaled. Remaining panels: color coding shows biases of the different mechanisms relative to the multi-model mean O_3 concentrations, symbol size scales with O_3 concentration. All values represent averages over the hours 0 to 72. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

spin-up period, and all statistics reported are based on days 2–7 (hours 25–168).

We first evaluate differences in the spatial pattern of O_3 concentrations between the mechanisms over a comparable range of emission conditions, removing differences in horizontal resolution by averaging the emissions input to a common grid. Secondly, we highlight differences in the diurnal evolution of pollutants by calculating statistics over box model results using the emissions conditions at the locations of a station network that has been used in the 3-D model evaluation of Im et al. (2015). Thirdly we look at predicted O_3 as a function of NO_x and VOC concentrations. Then we investigate indicators that predict NO_x/VOC limited regimes to understand how mechanisms would react to changes in emissions. Finally, we briefly review important (inorganic) rate constants determining O_3 and OH formation and loss.

The reader is reminded that no comparison with observations is made in our analysis, and any under-/overestimations and biases presented here are always versus the multi-model mean. It should also be noted that the correction of errors found in the mechanism implementations of MOZART-4 and CB05Clx did have (considerable) influence on the results. For the sake of consistency with the 3-D model intercomparison, simulations were done with the uncorrected mechanisms. Figures showing results using the corrected mechanisms can be found in Section S4 of the Supplementary material.

3.1. Spatial differences in maximum 8-h average O_3

We first sample all land points within the common model domain (intersection of all models) in 1° steps, using emissions averaged over an area of 0.5 × 0.5° to remove effects of different resolution/projections. Running the box model with each mechanism for each grid point (see e.g. Fig. 3, top left panel) provides us with concentration time series for all species in all mechanisms at all sampling locations. Note that these results are not directly comparable with full 3-D simulations as transport and variability in

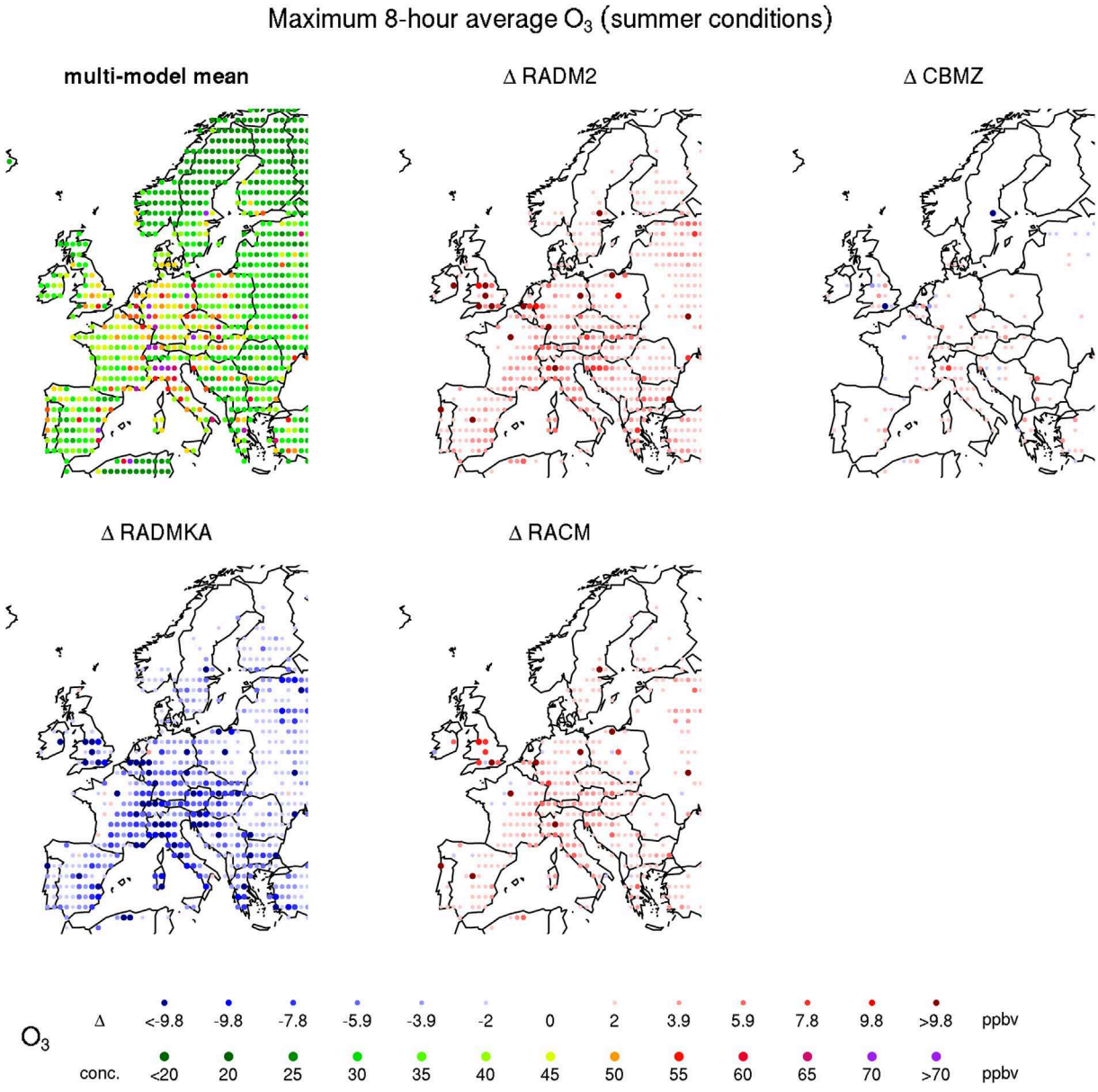

Fig. 4. Same as Fig. 3, but for the European domain and participating models.

meteorological conditions are not considered in this analysis. Our results do however provide insight into the different chemical regimes that exist across the two continents due to varying emissions. Finally, we conducted these simulations using three different sets of temperature, PBLH and photolysis rates, and are hence able to sample the annual evolution of meteorology of the original 3-D simulations.

For Figs. 3 and 4 we calculated the maximum 8-h averaged O_3 concentrations over the days 2–7, a quantity that closely resembles the metrics used in air quality standards in North America and Europe. We show the multi-mechanism average in the top left plot, and the biases from this average for each mechanism in the other plots. The reader can therefrom infer at which locations the highest maximum 8-h average O_3 is predicted, and also what the variability in this quantity will be depending on the choice of gas-phase mechanism. Figs. 3 and 4 show the results for 'summer' conditions, the corresponding figures for 'spring/autumn' and 'winter' conditions, as well as the results using the error-corrected mechanisms (for North America) can be found in Sections S1 And S4 in the Supplementary material respectively.

In North America under 'summer' conditions we find that, as expected, the eastern, and especially north eastern United States exhibit the highest levels of maximum 8-h average O_3 due to the strong anthropogenic emissions (Fig. 2). Also, the Gulf coast and major parts of Texas show high maximum 8-h average O_3. When analyzing mechanism-specific deviations from the mean we find that RADM2 gives consistently higher levels (up to +8 ppbv) of maximum 8-h average O_3 than the other mechanisms, especially over the Southeastern US. It is very instructive to consult Fig. 2, as there is a clear spatial correlation between the areas of highest biogenic emissions and the magnitude of the bias in RADM2 maximum 8-h average O_3. MOZART-4 predicts levels of maximum 8-h average O_3 that are slightly higher (+1 ppbv) than the multi-mechanism mean in the southeast and along the Atlantic coast, but considerably lower (up to −8 ppbv) around the Great Lakes area and the Midwest. We can show (Fig. S19) that the low bias of MOZART-4 over the Midwest/Great lakes area is almost exclusively due to an error found in the reaction rate constant of $NH_3 + OH$. Both CB05 mechanisms exhibit a slightly lower-than-average maximum 8-h average O_3 in general, with the strongest low bias in the region of highest biogenic VOC emissions (approximately over the states of Missouri, Arkansas and Louisiana). The CBMZ results are closest to the multi-model mean.

Under 'spring' conditions, the low bias in MOZART-4 maximum 8-h average O_3 is most pronounced (Fig. S2, up to −9 ppbv). The observations under 'summer' conditions that RADM2 gives higher concentrations (up to +9 ppbv) and that the Carbon Bond mechanisms are closer to the multi-model mean still holds. Again the mistake found in the MOZART-4 implementation was responsible for this underestimation. In the simulation using the corrected version of MOZART-4 (Fig. S18) we now find much smaller overall deviations from the multi-model mean for all mechanisms (maximum of ±5 ppbv), and different spatial patterns: large differences in maximum 8-h average O_3 only appear at grid points where urban areas are included in the averaged emissions, with RADM2 responding almost universally with higher maximum 8-h average O_3 than the remaining mechanisms, CBMZ again at the center of the distribution, and both CB05 mechanisms tend to predict lower maximum 8-h average O_3. MOZART-4 (corrected) predicts lower than average maximum 8-h average O_3 (up to −5 ppbv) over the Great Lakes area and grid points with urban influence, but slightly higher maximum 8-h average O_3 over the more rural grid points of the Southeast. Results from the simulations using 'winter' conditions are very similar to those for 'spring' conditions over North America.

Looking at the results of the simulations under 'summer' conditions in Europe (Fig. 4) we find the expected patterns of multi-model mean maximum 8-h average O_3, with highest concentrations over the polluted regions of northern Italy (Po valley), northern Switzerland, western Germany and Belgium, Netherlands, Luxembourg. The CBMZ mechanism is closest to the multi-model mean. RACM and RADM2 tend to predict higher-than-multi-model-mean concentrations of maximum 8-h average O_3, with the strongest deviations from the mean at grid points with strong influence from urban centers (Paris, Berlin, Madrid, Milan, etc.). RADMKA results are universally at the lower end of the mechanisms, with underestimations up to −10 ppbv over most of Central and Southern Europe. Only over the west of France and in parts of Eastern Europe is RADMKA close to the multi-model mean. It has become apparent in our work that RADMKA exhibits a low bias in O_3 compared to the other mechanisms under 'summer' conditions, but we could not identify the exact cause of this finding. The negative bias of RADMKA versus the multi-model mean we find is consistent with the negative overall bias when comparing the results of the 3-D model simulations against station data as it is reported by Im et al. (2015). A correlation is apparent between areas of lower biogenic VOC emissions (Fig. 2) and the magnitude of the underestimation of maximum 8-h average O_3, suggesting that the reason for the differences is to be found in the part of RADMKA describing the oxidation of biogenic precursors.

We do not observe the same low bias of RADMKA against the multi-mechanism mean in the simulations under 'spring/autumn' and 'winter' conditions (Figs. S3 and S4), possibly due to the lower biogenic emissions. Under 'spring' conditions, RADM2 and RACM predict the highest maximum 8-h average O_3, up to 8 ppbv higher than the multi-model mean at grid points with urban-influenced emission conditions. CBMZ is closest to the multi-model mean, with a tendency to predict lower-than-mean maximum 8-h average O_3 under polluted conditions (Berlin, western Germany, Po Valley, etc.). RADMKA results now indicate both higher- and lower-than-mean maximum 8-h average O_3, with highest negative biases versus the multi-model mean over Italy, southeastern Europe and the British Isles, but higher-than mean concentrations over northern Germany and parts of France. Results look similar under 'winter' conditions (Fig. S3). Simulations including the corrected RACM mechanism showed no observable effect on the results in any season.

3.2. Diurnal patterns of select compounds

In contrast to the previous section where we conducted box model simulations using average emissions at regular intervals, we will now sample the emission datasets at the locations of a given station network. Im et al. (2015) evaluated 3-D model performance during AQMEII Phase 2 through comparison against surface observations. They obtained a statistically valid comparison through selecting only stations with high data availability and spatial representativeness (only background sites), but their selection was not intended to get a set of comparisons that are equally distributed across the range of NO_x and VOC concentrations. Hence, a mechanism with a bias within this NO_x-VOC plane might have its bias exacerbated or decreased based on station selection. To understand mechanism performance for a comparison against exactly this particular range of conditions, we conducted box model simulations using emission conditions at the locations of these observations (511 in Europe, 219 in North America; see Fig. S22 in the Supplementary material for a map of station locations).

We first calculate the average diurnal cycle over days 2–7 for compounds of interest from a simulation using the emission situation at a station location. Doing so for all station locations in the

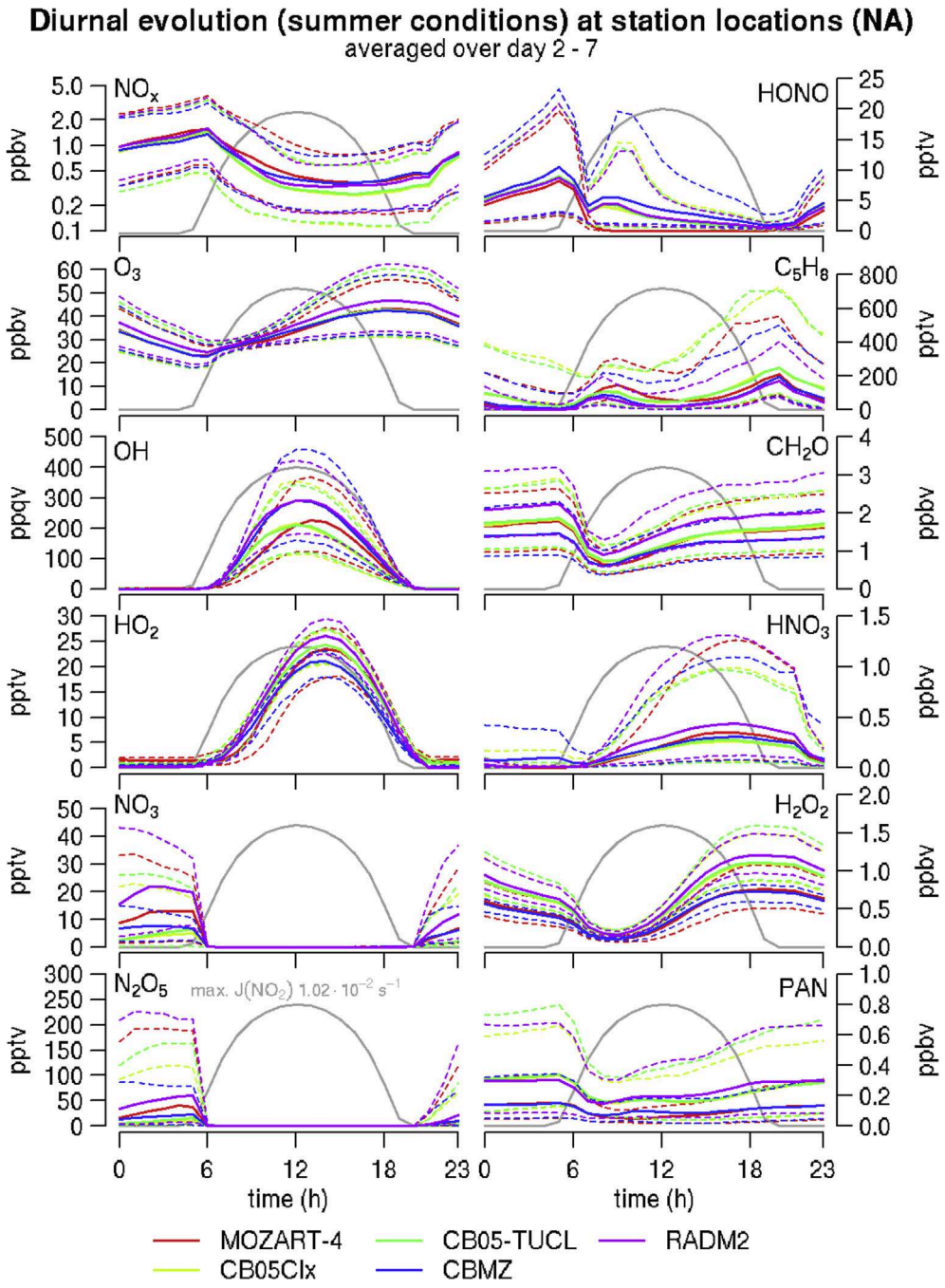

Fig. 5. 25 (dashed), 50 (solid) and 75 (dashed)% values for select species as result of the box model runs under 'summer' conditions at the observation locations used in Im et al. (2015) for the North American domain. Different colors represent different mechanisms. Gray line is J (NO_2) (units not shown). Note logarithmic scale for NO_x. PAN is sum of all PAN species (MOZART-4: PAN + MPAN, CB05Clx: PAN + PANX, CB05-TUCL: PAN + PANX + OPAN, CBMZ: PAN, RADM2: PAN + TPAN).

network we derive the hourly 25%, 50% (i.e., the median) and 75% values, which we present in Figs. 5 and 6 for the simulations under 'summer' conditions. The corresponding figures for the 'spring/autumn' and 'winter' conditions, as well as the ones showing the relative and absolute deviations of the 50% value from the multi-mechanism mean 50% value can be found in the Supplementary material (Sections S2 And S3).

Looking at the results over North America under 'summer' conditions, we find that the mechanisms agree on the diurnal evolution of the median concentrations of O_3 within 4 ppbv (5%) (Figs. 5 and S11), with RADM2 predicting up to 3 ppbv more O_3 than the multi-mechanism average. The remaining mechanisms (MOZART-4, CB05Clx, CB05-TU and CBMZ) are within 1 ppbv. Im et al. (2015) highlighted differences in predicted O_3 between the simulations of U.S. EPA (tagged as 'US6' in Im et al. (2015)) and North Carolina State University ('US8'). Both groups use a similar CB05 mechanisms but differ in the way photolysis is approximated for certain VOCs. We find that these differences did not result in differences in predictions of O_3 (or other compounds investigated), concluding that we can rule out the gas-phase mechanism as the

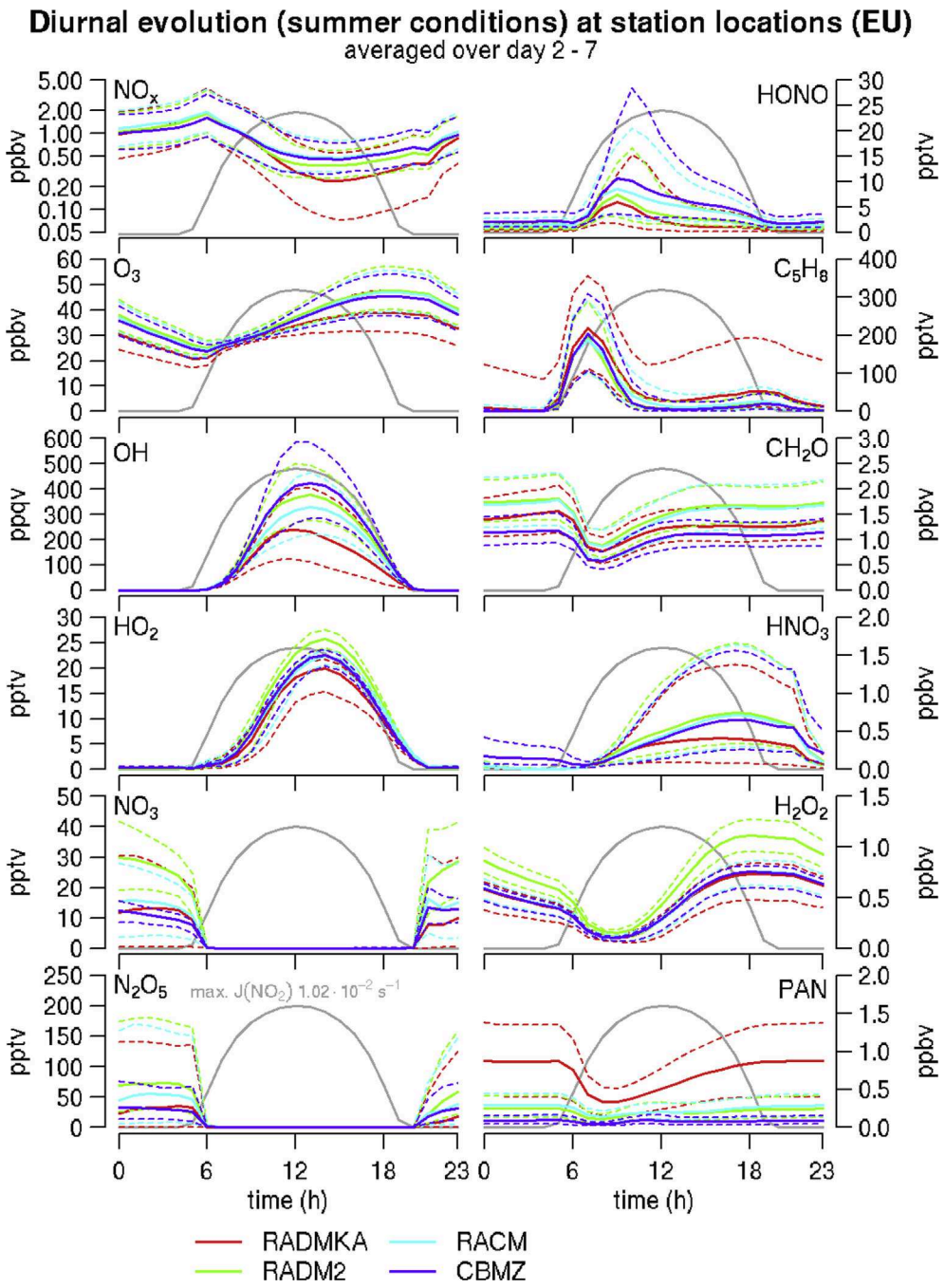

Fig. 6. Same as Fig. 5 but for the European domain and mechanisms. PAN is sum of all PAN species (RADMKA: PAN + TPAN + MPAN, RADM2: PAN + TPAN, RACM: PAN + TPAN + MPAN, CBMZ: PAN).

responsible process for the differences found between those two mechanisms by Im et al. (2015).

The mechanisms under investigation achieve these very similar concentrations of O_3 even though their levels of NO_x deviate from the multi-mechanism mean up to 25%, and daytime levels of key radicals OH and HO_2 by up to 40%/20% respectively. We find a distinctly different diurnal evolution of OH radicals in the MOZART-4 mechanism which we can again attribute to the erroneous rate constant for NH_3 + OH (Fig. S15). Key nighttime species like NO_3 and N_2O_5 exhibit a similar evolution over time across mechanisms, but vary greatly (several tens of pptv, −75 to above +100%) in the absolute concentrations. In our simulations we did not consider a heterogenous reaction of N_2O_5 on (aerosol) surfaces as a possible sink for the NO_3 radical which would introduce additional uncertainty due to model differences in aerosol surface area density.

Isoprene (C_5H_8), a major precursor of secondary organic aerosol formation, and compound of major interest in current research due to its potential influence on radical concentrations, varies ± 40 pptv (−90 to more than +100%). This is despite identical biogenic emissions. We note that a large fraction of this discrepancy will be due to the differences in predicted concentrations of OH/NO_3 radicals amongst the mechanisms.

Formaldehyde (CH_2O), a major oxidation product of anthropogenic and biogenic VOCs and often used to evaluate modeling

results against station measurements as well as satellite observations, has again a very similar diurnal evolution, but concentrations vary by ± 20%. H_2O_2 and HNO_3 are important endpoints for OH and HO_2 radicals in the atmosphere, and hence are an indicator for the amount of radical cycling. We find that the amount of H_2O_2 formed varies by ±25% between the mechanism. HNO_3 concentrations vary less during daytime (±10% during the day, with RADM2 up to 50% more than the multi-mechanism mean), but considerably at night (−75 to above +100%). Finally we find that peroxyacetyl nitrates (PAN), an important reservoir species for NO_2 and hence responsible for remote O_3 production, differs by up to ± 50% (±0.1 ppbv) between mechanisms, with CBMZ and MOZART-4 producing less, and the CB05 mechanisms and RADM2 more.

Examining the simulations under 'spring/autumn' conditions (Figs. S6, S10), we find a pronounced overestimation of NO_x, underestimation of O_3, and strong suppression of radicals OH and HO_2 as well as differences in secondary pollutants by the MOZART-4 mechanism. Again we can trace this error back to the erroneous rate constant of NH_3 + OH, and we can show that fixing this mistake will lead to results comparable to that of the other mechanisms (Fig. S15). In the simulations under 'winter' conditions we find very similar patterns in terms of deviation from the multi-model mean compared to the simulations under 'spring/autumn' conditions.

We also evaluated the set of mechanisms that have been employed over the European domain. When analyzing the results under 'summer' conditions (Figs. 6 and S14) we find that mechanisms agree within 10 ppbv (3 ppbv without RADMKA) on the diurnal evolution of median O_3. Overall the diurnal evolution of species investigated as well as the variability across mechanisms is very similar to the results over North America. Two species exhibit considerably different diurnal cycles, though: firstly, HONO shows a much stronger build-up during nighttime for the mechanisms employed over North America than over Europe. We attribute this to different direct emissions of HONO. And secondly, the diurnal evolution of isoprene concentrations predicted over North America shows a much stronger secondary peak at nightfall (once the boundary layer collapses) than the mechanisms evaluated over the European domain. We attribute this to stronger emissions of isoprene over the North American domain (Fig. 2).

For the mechanisms that we evaluated over Europe it becomes evident that under 'summer' conditions RADMKA is the mechanism that stands apart. While RADM2, RACM and CBMZ agree on the diurnal evolution of median O_3 within 3 ppbv (Fig. S14), RADMKA predicts up to 6 ppbv less O_3, especially during afternoon and evening hours. This is accompanied by distinctly lower concentrations of NO_x during afternoon and evening (−0.15 ppbv, −25%), lower OH (−100 ppqv, −25%) and HNO_3 (−0.2 ppbv, −25%). Striking are also concentrations of PAN that are 3–4 times higher in RADMKA than the prediction of the next-highest mechanism.

The biases found in RADMKA predictions compared to its peers begin to disappear when looking at the results under 'spring/autumn' and 'winter' conditions (Figs. S7, S8). Mechanisms agree on the diurnal evolution of median O_3 within 3 ppbv (Figs. S12, S13) and the differences in NO_x are reduced to ± 10% (±0.1 ppbv). Differences in nighttime species NO_3 and N_2O_5, formaldehyde,

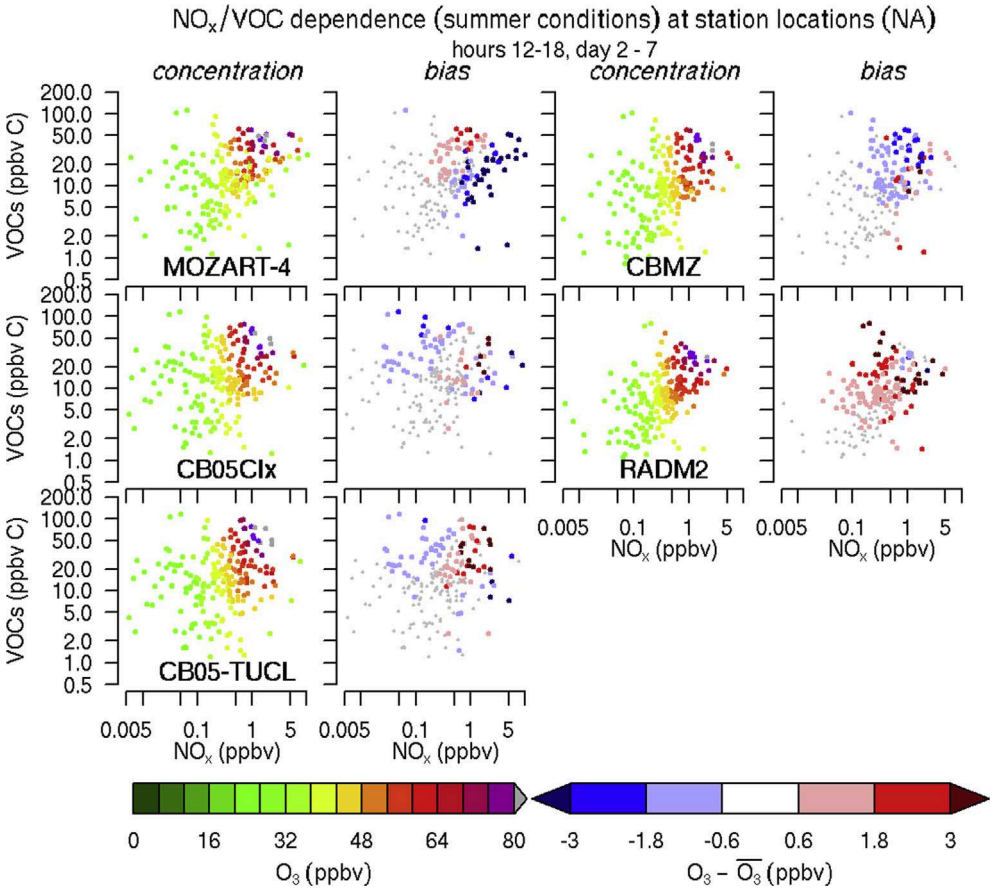

Fig. 7. Predicted O_3 concentrations and biases as a function of NO_x and VOC concentrations for the stations approach over the NA domain. Biases are relative to the multi-model mean at each location. Used are only afternoon values (12–18 LT).

isoprene and PAN are reduced as well. Differences in HNO_3 and H_2O_2 remain similar to 'summer' conditions, still indicating comparable differences in radical cycling numbers. Results from the simulation under 'winter' conditions are similar to the 'spring/summer' results.

3.3. Response to varying NO_x/VOC conditions

In the following we look at the dependence of O_3 on NO_x and VOC levels using the same set of simulations at the locations of stations employed in the last section. Figs. 7 and 8 show O_3 concentrations plotted as a function of NO_x and VOCs. We present each point color-coded by the averaged afternoon (12–18 LT, days 2–7) concentration of O_3 at one of the sampling points, located at the respective afternoon-averaged concentrations of NO_x and total VOCs. Also plotted are the relative biases towards the multi-model mean at each location. This hence provides the reader with an understanding of the pollution situations in which a mechanism might have a certain bias for O_3 compared to its peers.

It is obvious from Figs. 7 and 8 that all mechanisms are able to represent the NO_x/VOC dependence of O_3 in general. The CB05 mechanisms in North America (Fig. 7) tend to be biased low in O_3 under low NO_x/high VOC conditions (e.g. the biogenic emissions-rich southeastern US, Fig. 2) as well as under very high NO_x conditions. CBMZ is low biased over polluted conditions in general, with some erratic high-biased points interspersed. MOZART-4 predicts higher O_3 concentrations than the multi-model average for moderately NO_x-polluted regions under high VOC loads (e.g. rural areas within a region with high biogenic emissions), and a strong low bias under high NO_x conditions. This underestimation is reduced when using the error-corrected MOZART-4 mechanism, but the general pattern still stays the same (not shown). The RADM2 mechanism-predicted O_3 is often higher than the multi-mechanism average, except for situations where we have both high VOC as well as NO_x concentrations. Again, the southeastern US is a prime example of an area where mechanisms differ for O_3 (see Fig. 3).

For the mechanisms applied over the European domain, we find RACM and CBMZ close to the multi-model mean, with slight overestimations at average to high levels of NO_x (0.5 ppbv) and high VOCs. The RADM2 mechanism tends to predict higher levels of O_3 especially under high VOC load conditions across all NO_x levels. For RADMKA finally, we find that the low bias in O_3 observed before is most pronounced under high NO_x and/or high VOC conditions, again indicating that the isoprene and monoterpene oxidation chemistry might be a key factor in explaining these differences.

3.4. NO_x/VOC limited regimes

Indicators can serve as another useful measure to understand a chemical regime. Martin et al. (2004) proposed the use of the ratio of formaldehyde (CH_2O) over nitrogen dioxide (NO_2) to evaluate whether the conditions are VOC- or NO_x-limited (i.e., $\partial O_3/\partial VOC > \partial O_3/\partial NO_x$ and vice versa). Such a measure is complementary to what we presented in the last section, as it represents the slope of the O_3 surface in Figs. 7 and 8 at a certain point. The indicator of Martin et al. (2004) is especially useful as both quantities can be observed via satellite instruments and are hence eminently useful for 3-D model evaluation as is done in Zhang et al. (2010) or Campbell et al. (2015). CH_2O/NO_2 indicates VOC-limited conditions if the ratio is below 1, and NO_x-limited conditions if above 1.

We calculated this indicator based on the result of the box model simulations using the emission situation at station locations, averaging ratios of CH_2O to NO_2 over the local hours 8–12 (similar to Campbell et al. (2015), to match overpass times of the satellite instrument observing these variables) of days 2–7. Fig. 9 shows the resulting histogram of NO_x/VOC limitation for each season and the two continents. Clearly visible is the predominance of VOC-limited conditions during winter, caused probably by the much lower emissions of biogenic VOCs. Spring/autumn conditions are marked

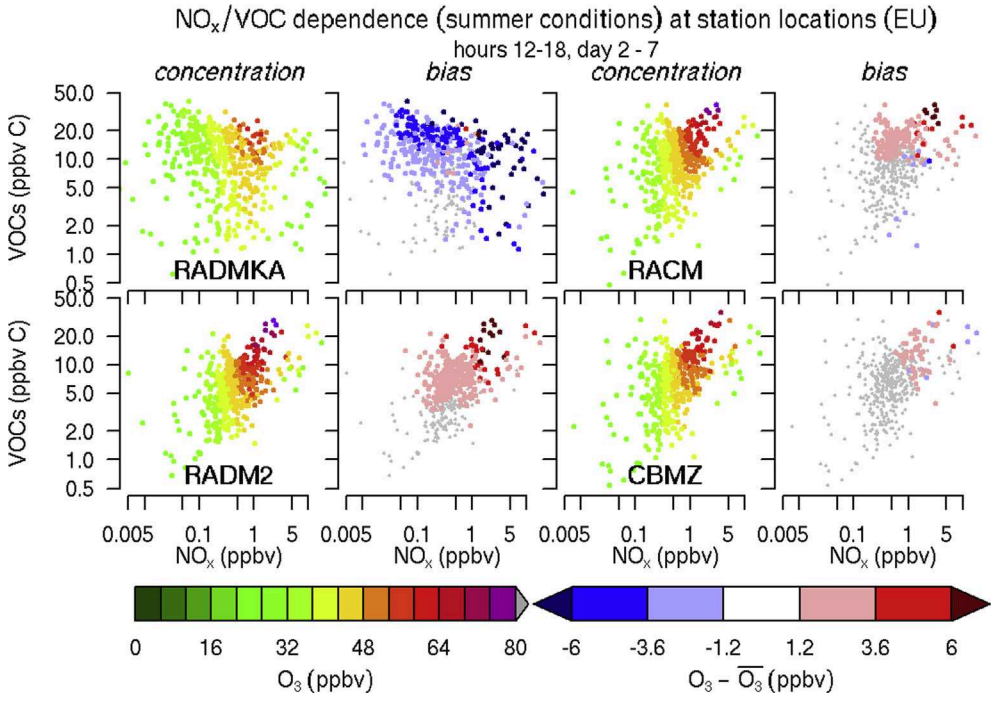

Fig. 8. Same as Fig. 7, but for O_3 over the EU domain.

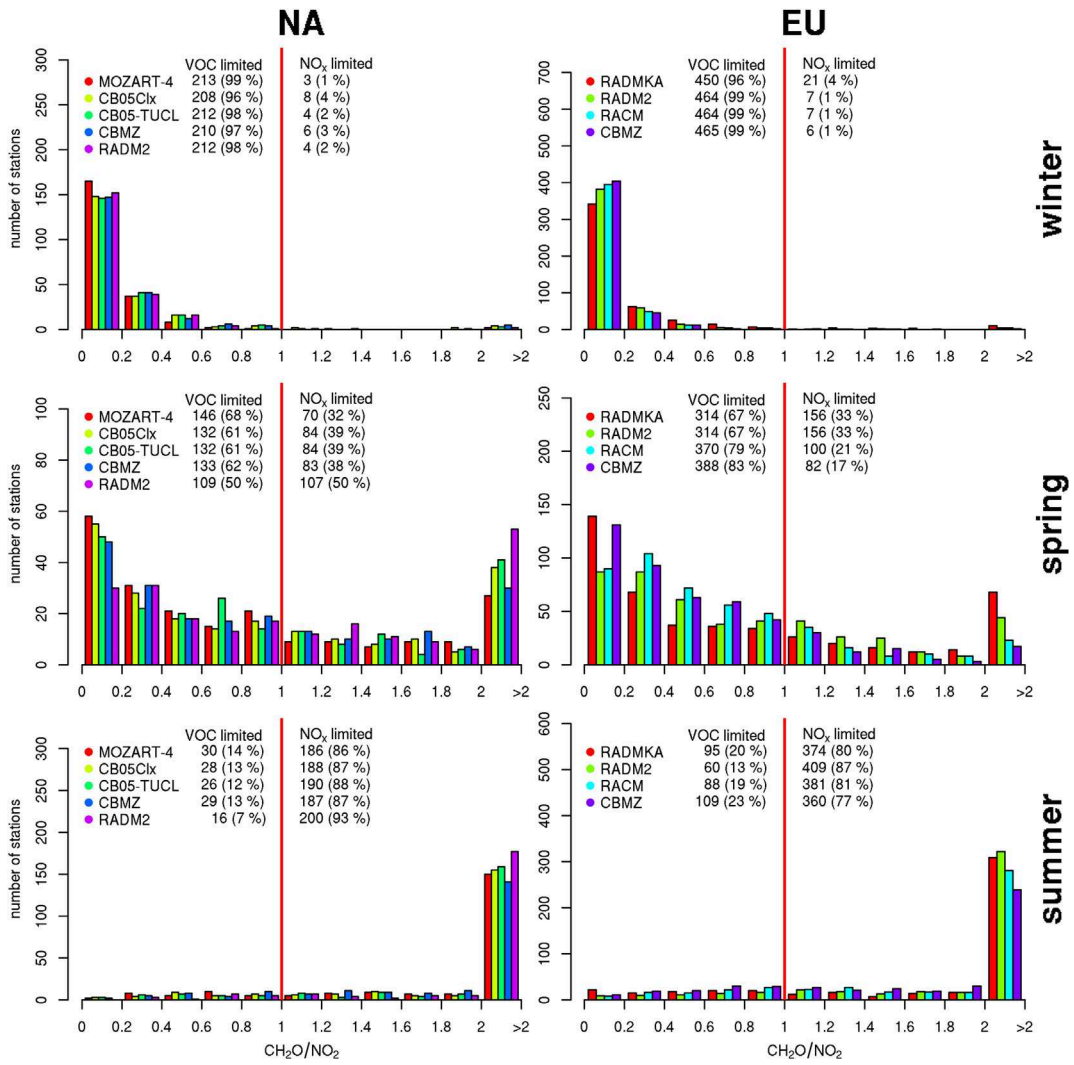

Fig. 9. Histogram of ratios of CH_2O/NO_2 based on box model simulations using emission conditions at station locations. Shown are ratios based on values averaged over the local hours 8–12 of days 2–7. Ratios above 2 are omitted in the plot for clarity, but included in the calculations of statistics and percentiles. A value of CH_2O/NO_2 below 1 indicates VOC-limited, values above 1 NO_x-limited conditions. The turnover value is marked with a red line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

by a transition to more NO_x-limited conditions, and under summer conditions, over 80% of station locations are NO_x-limited.

This evolution is represented in all mechanisms on both continents. The MOZART-4 mechanism predicts the highest fraction of VOC-limited locations, especially under spring/autumn conditions where additional 7%/6% of locations would be considered VOC-limited compared to the CB05 mechanisms and CBMZ respectively, and almost 20% more stations compared to RADM2. The CB05 mechanisms and CBMZ are very similar over the North American domain in predicting NO_x/VOC-limited regimes. RADM2 and its variant RADMKA predict much lower fractions of VOC-limited conditions on both continents, with considerable differences (>10% classified differently) especially under spring/autumn conditions.

Assuming one would attempt a study to investigate the effect of changing emissions on tropospheric O_3 concentrations, this would lead to different answers depending on the chemical mechanism employed. Especially under spring/autumn conditions, RADM2/RADMKA-based simulations would indicate that reducing NO_x will be 'more effective' in reducing O_3 concentrations, in the sense that an additional 10% of stations will exhibit larger changes in O_3 when reducing NO_x than under a comparable change in VOCs. RACM, CBMZ and MOZART-4 on the other hand would favor changes in VOCs at these station locations. Those results would have important policy implications, in particular, they indicate a need to perform ensemble simulations using different gas-phase mechanisms to support the development of robust emission control strategies for O_3 pollution.

3.5. Differences in major inorganic reaction rates

We briefly analyzed individual rate constants in an effort to understand the differences in concentrations found across the mechanisms. In Fig. 10 we show the rate constants for reactions important for OH and O_3 production and loss. We find that primary O_3 production reactions ($NO + HO_2 \rightarrow NO_2 + OH$ and $NO + CH_3O_2 \rightarrow NO_2 + CH_2O + HO_2$) as well as $O^3P + O_2 \rightarrow O_3$ are consistent across mechanisms, while O_3 loss reactions ($O_3 + OH \rightarrow O_2 + HO_2$, $O_3 + HO_2 \rightarrow 2\ O_2 + OH$) show minor differences. The reaction $OH + NO_2 \rightarrow HNO_3$ is instrumental as a radical termination reaction in the troposphere, however there has been considerable uncertainty in the past in the reaction rate and products formed (e.g.,

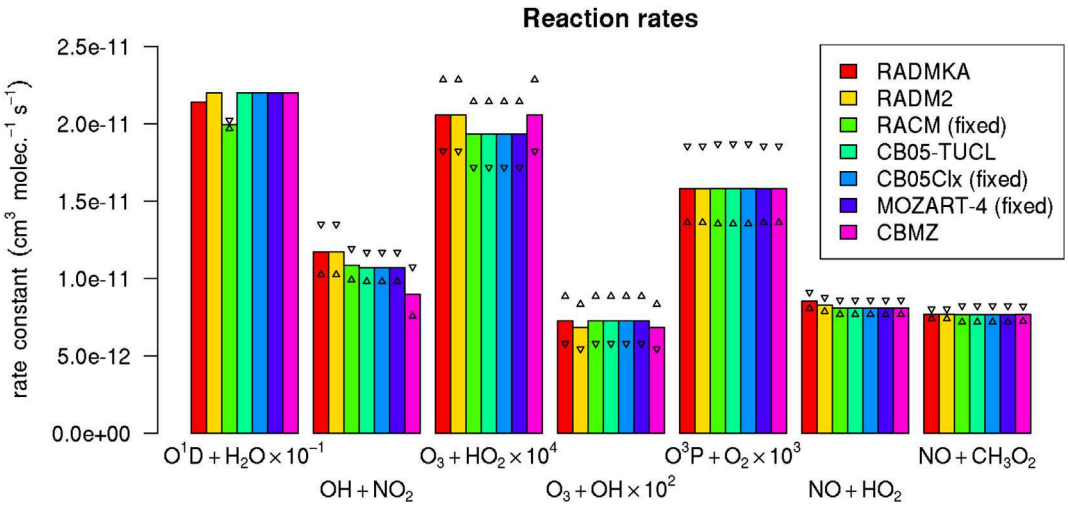

Fig. 10. Comparison of reaction rates important for the formation and loss of OH and O_3. Bars show reaction rates at 298.15 K, upward pointing triangles at 318.15 K, downward pointing triangles at 278.15 K. If triangles are omitted, no temperature dependence is considered. Note that rates for $O^1D + H_2O$, $O^3P + O_2$, $O_3 + OH$, and $O_3 + HO_2$ are scaled for presentation purposes.

Zhang and Donahue, 2006, and references therein). We find that mechanisms differ in both the value at 298 K as well as in the magnitude of the temperature dependence of this reaction, which could partly explain differences in oxidant levels observed as well as the amount of HNO_3 formed. Overall the reaction rates investigated are – apart from $OH + NO_2 \rightarrow HNO_3$ – very similar, indicating that the differences in resulting concentrations that we observe are due to differences in the VOC part of the mechanisms.

4. Discussion

The largest variability of predicted O_3 across mechanisms was found over regions with strong biogenic influence. In particular, isoprene chemistry is a topic of intense current scientific investigation, as recent findings suggest that the loss of OH radicals through oxidation of isoprene is lower than previously estimated and the cycling of NO_x may be larger than predicted by older mechanisms (e.g., Paulot et al., 2009; Taraborrelli et al., 2012). In this intercomparison we used mechanisms with very different descriptions of isoprene chemistry. The RADM2 mechanism uses the original formulation from Stockwell (1986). CB05 represents isoprene chemistry as a condensation of the detailed mechanism of Carter (1996). RADMKA and RACM use updated formulations based e.g. on the work of Geiger et al. (2003). MOZART-4 includes a fairly detailed representation of isoprene chemistry, including chemistry of first-generation (e.g., MVK, MACR and hydroxycarbonyls) and subsequent generation products (e.g., glycolaldehyde, hydroxyacetone, methylglyoxal, glyoxal). Also included is a representation of isoprene hydroxyl-peroxy radical isomerization (e.g., Crounse et al., 2011), leading to formation of an isoprene-derived hydroperoxyaldehyde. Furthermore, not all mechanisms include a description of monoterpene chemistry. RADM2 does not consider monoterpenes at all, and CBMZ only includes reactions of monoterpenes with radicals to form condensable vapors for SOA production.

Mechanisms were found to differ more strongly in their predictions of O_3 levels and other pollutants in regions with strong biogenic VOC emissions, hence suggesting that these are regions where predictions are more uncertain. We did not compare against measurements and hence cannot determine which mechanism matches observations best, but we found that updates to the oxidation chemistry for biogenic VOCs seem to have had strong influence on predicted concentrations. Isoprene chemistry is a rapidly evolving field and future refinements to the mechanisms should reflect our increased understanding of the relationship between isoprene oxidation and HO_x and NO_x cycling. The reader is referred to the literature for an in-depth discussion of differences in isoprene oxidation mechanisms (e.g., Pöschl et al., 2000; Archibald et al., 2010; Zhang et al., 2011).

Our results further suggest that processes and parameterizations based on secondary products/radicals will strongly be affected by the choice of gas-phase mechanism, even when a comparison e.g. against observations of O_3 would suggest excellent agreement. While radical concentrations themselves are not of major importance for regulatory questions, they are key to a number of processes like the formation of secondary organic aerosols (SOA). SOA is formed through the continuous oxidation of biogenic and anthropogenic gas-phase precursors like isoprene, aromatics or alkanes by oxidants like O_3 or OH, but also NO_3. In the modeling systems investigated here, SOA formation would typically be parameterized in the form of additional products that are added to the oxidation reaction of a precursor in the gas-phase mechanism. Often, aging reactions of these products with OH are included as well to consider further oxidation reactions that continue to lower a substance's volatility. The differences in oxidant levels found here would directly influence any parameterization of aerosol formation through the oxidation of gas-phase precursors, adding yet another source of uncertainty in our ability to represent these types of aerosols in state-of-the-art modeling systems.

Results from the calculation of the CH_2O/NO_2 indicator are provocative, as they classify a considerable fraction of stations (up to 20%) differently. This would then also mean that their response to changes in emissions would be very different, and that the choice of gas-phase mechanism is also crucial in simulations made to derive efficient emission reduction strategies.

Our study was designed to constrain mechanisms as well as possible, so as to remove any differences in the input to the different mechanisms, hence allowing to compare the results in terms of the performance of the gas-phase mechanism itself. We intentionally did not investigate the influence of different photolysis schemes, but the reader is reminded that accurate photolysis rates are a prerequisite of any successful prediction. Neither did we investigate the effect of different numerical solvers on the results.

We note, however, that for example in WRF-Chem, all mechanisms use the same solver as we have used in our study.

5. Conclusions

We intercompared the majority of gas-phase mechanisms employed in the AQMEII phase 2 model intercomparison in order to understand mechanism-specific biases, guide other evaluations and provide the reader with an insight into the uncertainties resulting from the choice of gas-phase mechanism for their 3-D model simulations. Our analysis methods ensured that the mechanisms were compared under tight constraints for all processes but the lumping of anthropogenic VOC emissions (which was done mechanism-specific) so that the resulting differences can only be discussed in terms of the gas-phase mechanism itself or the lumping of anthropogenic VOC emissions. Simulations were made under three different sets of environmental constraints to represent meteorological conditions of all seasons.

There are a number of implications from our analysis:

- An uncertainty in predicted O_3 of 4 ppbv (5%) solely due to the choice of gas-phase mechanism should be considered in the analysis of 3-D model results of O_3. For NO_x we found an uncertainty up to 25% across mechanisms.
- Predicted concentrations of peroxyacetyl nitrates (PAN) are found to vary by 50%, highlighting that also remote production of O_3 can be directly affected by the choice of gas-phase mechanism.
- Predictions of key VOCs have higher uncertainty (±100% for isoprene, ±20% for formaldehyde), which suggests that biases of this magnitude e.g., in the comparison against satellite data, could be solely due to the choice of the gas-phase mechanism.
- Differences in daytime OH radical concentrations of up to 40% (20% for HO_2) imply that parameterizations that depend on this concentration (e.g., secondary organic aerosol formation) have an inherent uncertainty of this magnitude.
- Concentrations of compounds central to nighttime chemistry (NO_3, N_2O_5) vary by up to 100% between mechanisms, indicating considerable uncertainty in our knowledge of this potentially important part of tropospheric chemistry (aerosol formation, e.g., also depends on reaction with NO_3)
- A variability of 25/10% in the radical termination species H_2O_2 and HNO_3 suggest substantially different radical cycling numbers.
- Regions with the highest biogenic VOC emissions tend to produce the largest variability in predicted O_3, hence suggesting larger uncertainty in the chemistry of biogenic VOCs.
- Classification of stations into chemical regimes differs by up to 20%, which will lead to a likewise uncertainty in the answer to the questions which the most efficient emission reduction strategy would be.

A number of subtle errors have been discovered in both the implementation of mechanisms as well as the preparation of emissions that so far went unnoticed in the evaluation of the 3-D simulations. MOZART-4 exhibited a strong low bias in O_3 under 'spring' conditions over the Great Lakes/Midwest area, which was found to be due to an erroneous rate constant (NH_3 + OH) in the WRF-Chem implementation of MOZART-4. Simulations with the corrected mechanism resulted in MOZART-4 being much closer to the multi-model mean. Errors in the implementations of the RACM and CB05Clx mechanisms that were found did not result in notable changes of the results when employing a corrected version. We observed a strong low bias in O_3 (versus the multi-model mean) by RADMKA during summer months which we attribute to an error in the mechanism, with the magnitude of the bias anti-correlated with the amount of biogenic emissions over Europe. This bias was not found under 'spring'/'winter' conditions (with lower biogenic emissions). Analysis of the reason for this finding is ongoing. All these findings further underline the value of assessing complex modeling systems by disassembling them into core components.

When connecting our results with the results from the 3-D model intercomparison, we found that the two variants of the CB mechanism do not differ in their predicted O_3 concentrations, which rules out the gas-phase mechanism as the responsible model component for differences found by Im et al. (2015) in the 3-D model evaluation. On the other hand we can confirm a strong negative bias in O_3 predictions by the RADMKA mechanism under 'summer' conditions, hence suggesting that efforts should be undertaken to improve this mechanism under these conditions. A spatial correlation of the magnitude of this bias with biogenic emission strength suggests the error to be found in this part of the mechanism.

We compare mechanisms that span two decades of research into tropospheric chemistry, from unaltered RADM2 implementations to current mechanisms like the variants of CB05 or MOZART. While it is out of scope of this work to show which mechanism performs the best compared to measurements, we do presume that advances in our understanding of tropospheric chemistry should be considered in mechanisms used for state-of-the-art modeling efforts, and groups should hence strive to update their gas-phase mechanisms accordingly.

Most importantly, our work shows that the choice of gas-phase mechanism introduces non-negligible uncertainty in predictions made using state-of-the-art modeling systems. This uncertainty is not limited to regulated gaseous pollutants, but extends to the predictions of radical concentrations as well as secondary products, including the ones central to aerosol formation.

Acknowledgments

Alessandra Balzarini (RSE) is thanked for providing emissions speciation information for the CBMZ mechanism. The group of Bernhard Vogel (IMK-TRO, KIT) is thanked for providing the KPP files of the RADMKA mechanism. Geoff Tyndall (NCAR) kindly provided assistance in updating the MOZART mechanism. Lea Giordano was supported by the Swiss State Secretariat for Education, Research and Innovation, project C11.0144. Y. Zhang and K. Yahya acknowledge funding support from the NSF Earth System Program (AGS-1049200) and high-performance computing support from Yellowstone by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and Stampede, provided as an Extreme Science and Engineering Discovery Environment (XSEDE) digital service by the Texas Advanced Computing Center (TACC). The UPM authors gratefully acknowledge the computer resources, technical expertise and assistance provided by the Centro de Supercomputacion y Visualizacion de Madrid (CESVIMA) and the Spanish Supercomputing Network (BSC). The UMU group acknowledges the funding from the project CGL2013-48491-R, Spanish Ministry of Economy and Competitiveness. G. Curci and P. Tuccella were supported by the Italian Space Agency (ASI) in the frame of PRIMES project (contract n.I/017/11/0). This research was supported by the National Center for Atmospheric Research, which is operated by the University Corporation for Atmospheric Research on behalf of the National Science Foundation. Any opinions, findings and conclusions or recommendations expressed in the publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the U.S. Environmental Protection Agency (EPA) or any other organization participating in the AQMEII project.

This paper has been subjected to EPA review and approved for publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2014.11.066.

References

Alapaty, K., Mathur, R., Pleim, J., Hogrefe, C., Rao, S.T., Ramaswamy, V., Galmarini, S., Schaap, M., Makar, P., Vautard, R., Baklanov, A., Kallos, G., Vogel, B., Sokhi, R., 2012. New directions: understanding interactions of air quality and climate change at regional scales. Atmos. Environ. 49, 419–421.

Archibald, A.T., Jenkin, M.E., Shallcross, D.E., 2010. An isoprene mechanism intercomparison. Atmos. Environ. 44 (40), 5356–5364.

Brasseur, G., Hauglustaine, D., Walters, S., Rasch, P., Müller, J.-F., Granier, C., Tie, X., 1998. MOZART, a global chemical transport model for ozone and related chemical tracers: 1. model description. J. Geophys. Res. Atmos. (1984–2012) 103 (D21), 28265–28289.

Campbell, P., Zhang, Y., Yahya, K., Wang, K., Hogrefe, C., Pouliot, G., Knote, C., Hodzic, A., San Jose, R., Perez, J.L., Jimenez-Guerrero, P., Baro, R., Makar, P., 2015. A multi-model assessment for the 2006 and 2010 simulations under the air quality model evaluation international initiative (AQMEII) phase 2 over North America: part I. Indicators of the sensitivity of O_3 and $PM_{2.5}$ formation regimes. Atmos. Environ. 115, 569–586.

Carter, W.P., 1996. Condensed atmospheric photooxidation mechanisms for isoprene. Atmos. Environ. 30 (24), 4275–4290.

Carter, W.P., 2010. Development of the SAPRC-07 chemical mechanism. Atmos. Environ. 44 (40), 5324–5335.

Computational and Information Systems Laboratory, 2012. Yellowstone: IBM iDataPlex System (NCAR Community Computing). National Center for Atmospheric Research, Boulder, CO. http://n2t.net/ark:/85065/d7wd3xhc.

Crounse, J.D., Paulot, F., Kjaergaard, H.G., Wennberg, P.O., 2011. Peroxy radical isomerization in the oxidation of isoprene. Phys. Chem. Chem. Phys. 13 (30), 13607–13613.

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., Irwin, J., Rao, S.T., Scheffe, R., Schere, K., Steyn, D., Venkatram, A., 2010. A framework for evaluating regional-scale numerical photochemical modeling systems. Environ. Fluid Mech. 10 (4), 471–489.

Emmons, L., Walters, S., Hess, P., Lamarque, J.-F., Pfister, G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., et al., 2010. Description and evaluation of the model for ozone and related chemical tracers, version 4 (MOZART-4). Geosci. Model Dev. 3 (1), 43–67.

Geiger, H., Barnes, I., Bejan, I., Benter, T., Spittler, M., 2003. The tropospheric degradation of isoprene: an updated module for the regional atmospheric chemistry mechanism. Atmos. Environ. 37 (11), 1503–1519.

Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C., Eder, B., 2005. Fully coupled online chemistry within the WRF model. Atmos. Environ. 39 (37), 6957–6975.

Gross, A., Stockwell, W.R., 2003. Comparison of the EMEP, RADM2 and RACM mechanisms. J. Atmos. Chem. 44 (2), 151–170.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P., Geron, C., 2006. Estimates of global terrestrial isoprene emissions using MEGAN (Model of emissions of gases and aerosols from Nature). Atmos. Chem. Phys. 6.

Hauglustaine, D., Granier, C., Brasseur, G., Megie, G., 1994. The importance of atmospheric chemistry in the calculation of radiative forcing on the climate system. J. Geophys. Res. Atmos. (1984–2012) 99 (D1), 1173–1186.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baro, R., Bellasio, 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., Makar, P., Manders-Groot, A., Neal, L., Perez, J., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R., Syrakov, D., Torian, A., Tuccella, P., Werhahn, K., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S., 2015. 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.

Karamchandani, P., Zhang, Y., Chen, S.-Y., Balmori-Bronson, R., 2012. Development of an extended chemical mechanism for global-through-urban applications. Atmos. Pollut. Res. 3 (1).

Kim, Y., Sartelet, K., Seigneur, C., 2011. Formation of secondary aerosols over europe: comparison of two gas-phase chemical mechanisms. Atmos. Chem. Phys. 11 (2), 583–598.

Knote, C., Brunner, D., Vogel, H., Allan, J., Asmi, A., Äijälä, M., Carbone, S., van der Gon, H., Jimenez, J., Kiendler-Scharr, A., Mohr, C., Poulain, L., Prevot, A.S., Swietlicki, E., Vogel, B., 2011. Towards an online-coupled chemistry-climate model: evaluation of trace gases and aerosols in COSMO-ART. Geosci. Model Dev. 4 (4), 1077–1102.

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

Kuhn, M., Builtjes, P., Poppe, D., Simpson, D., Stockwell, W., Baart, A., Das, M., Fiedler, F., Hov, Ø., Kirchner, F., et al., 1998. Intercomparison of the gas-phase chemistry in several chemistry and transport models. Atmos. Environ. 32 (4), 693–709.

Luecken, D., Phillips, S., Sarwar, G., Jang, C., 2008. Effects of using the CB05 vs. SAPRC99 vs. CB4 chemical mechanism on model predictions: ozone and gas-phase photochemical precursor concentrations. Atmos. Environ. 42 (23), 5805–5820.

Madronich, S., Flocke, S., 1997. Theoretical Estimation of Biologically Effective UV Radiation at the Earth's Surface. In: NATO ASI Series, vol. 52. Springer-Verlag, Berlin.

Martin, R.V., Fiore, A.M., Van Donkelaar, A., 2004. Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions. Geophys. Res. Lett. 31 (6).

Paulot, F., Crounse, J., Kjaergaard, H., Kürten, A., St Clair, J., Seinfeld, J., Wennberg, P., 2009. Unexpected epoxide formation in the gas-phase photooxidation of isoprene. Science 325 (5941), 730–733.

Pouliot, G., Denier van der Gon, H., Kuenen, J., Makar, P., Zhang, J., Moran, M., 2015. Analysis of the emission inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project. Atmos. Environ. 115, 345–360.

Pöschl, U., von Kuhlmann, R., Poisson, N., Crutzen, P.J., 2000. Development and intercomparison of condensed isoprene oxidation mechanisms for global atmospheric modeling. J. Atmos. Chem. 37 (1), 29–52.

Sandu, A., Sander, R., 2006. Technical note: simulating chemical systems in Fortran90 and matlab with the kinetic preprocessor KPP-2.1. Atmos. Chem. Phys. 6, 187–195.

Sarwar, G., Luecken, D., Yarwood, G., 2006. Developing and implementing an updated chlorine chemistry into the community multiscale air quality model. In: Presented at the 28th NATO/CCMS International Technical Meeting, Leipzig, Germany, May 15–19.

Stockwell, W., Kirchner, F., Kuhn, M., Seefeld, S., 1997. A new mechanism for regional atmospheric chemistry modeling. J. Geophys. Res. 102 (D22), 25847–25879.

Stockwell, W.R., 1986. A homogeneous gas phase mechanism for use in a regional acid deposition model. Atmos. Environ. 20 (8), 1615–1632.

Stockwell, W.R., Middleton, P., Chang, J.S., Tang, X., 1990. The second generation regional acid deposition model chemical mechanism for regional air quality modeling. J. Geophys. Res. Atmos. (1984–2012) 95 (D10), 16343–16367.

Stull, R.B., 1988. An Introduction to Boundary Layer Meteorology. Atmos. Oceanogr. Sci. Lib. Vol. 13. Springer.

Taraborrelli, D., Lawrence, M.G., Crowley, J.N., Dillon, T.J., Gromov, S., Gross, C.B.M., Vereecken, L., Lelieveld, J., 2012. Hydroxyl radical buffered by isoprene oxidation over tropical forests. Nat. Geosci. 5, 190–193.

University of North Carolina, Chapel Hill, N, 2003. Sparse Matrix Operator Kernel Emission (SMOKE) Modelling System. Environmental Programs, Carolina.

Vogel, B., Vogel, H., Bäumer, D., Bangert, M., Lundgren, K., Rinke, R., Stanelle, T., 2009. The comprehensive model system COSMO-ART – radiative impact of aerosol on the state of the atmosphere on the regional scale. Atmos. Chem. Phys. 9, 8661–8680.

Whitten, G.Z., Heo, G., Kimura, Y., McDonald-Buller, E., Allen, D.T., Carter, W.P.L., Yarwood, G., 2010. A new condensed toluene mechanism for carbon bond: Atmos. Environ. 44, 5346–5355.

Whitten, G.Z., Hogo, H., Killus, J.P., 1980. The carbon-bond mechanism: a condensed kinetic mechanism for photochemical smog. Environ. Sci. Technol. 14 (6), 690–700.

Wyat Appel, K., Gilliland, A.B., Sarwar, G., Gilliam, R.C., 2007. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance: part I—ozone. Atmos. Environ. 41 (40), 9603–9615.

Yarwood, G., Rao, S., Yocke, M., Whitten, G., December 2005. Updates to the Carbon Bond Mechanism: CB05. Report to the U.S. Environmental Protection Agency. http://www.camx.com/publ/pdfs/CB05_Final_Report_120805.pdf.

Zaveri, R.A., Peters, L.K., 1999. A new lumped structure photochemical mechanism for large-scale applications. J. Geophys. Res. 104 (D23), 30387–30415.

Zhang, L., Brook, J., Vet, R., 2003. A revised parameterization for gaseous dry deposition in air-quality models. Atmos. Chem. Phys. 3 (6), 2067–2082.

Zhang, J., Donahue, N.M., 2006. Constraining the mechanism and kinetics of $OH + NO_2$ and $HO_2 + NO$ using the multiple-well master equation. J. Phys. Chem. A 110 (21), 6898–6911.

Zhang, Y., Liu, P., Pun, B., Seigneur, C., 2006. A comprehensive performance evaluation of MM5-CMAQ for the summer 1999 southern oxidants study episode, Part III. diagnostic and mechanistic evaluations. Atmos. Environ. 40, 4856–4873.

Zhang, Y., Liu, X.-H., Olsen, K., Wang, W.-X., Do, B., Bridgers, G., 2010. Responses of future air quality to emission controls over North Carolina, part II: analyses of future-year predictions and their policy implications. Atmos. Environ. 44 (23), 2767–2779.

Zhang, H., Rattanavaraha, W., Zhou, Y., Bapat, J., Rosen, E.P., Sexton, K.G., Kamens, R.M., 2011. A new gas-phase condensed mechanism of isoprene-NO_x photooxidation. Atmos. Environ. 45 (26), 4507–4521.

Zhang, Y., Chen, Y., Sarwar, G., Schere, K., 2012. Impact of gasphase mechanisms on weather research forecasting model with chemistry (WRF/Chem) predictions: mechanism implementation and comparative evaluation. J. Geophys. Res. Atmos. (1984–2012) 117 (D1).