1   Decadal Application of WRF/Chem for Regional Air Quality and Climate Modeling over the
2   U.S. under the Representative Concentration Pathways Scenarios. Part 1: Model Evaluation and
3   Impact of Downscaling

4
5   Khairunnisa Yahya, Kai Wang, Patrick Campbell, Ying Chen, Timothy Glotfelty, Jian He,

6   Michael Pirhalla, and Yang Zhang*

7   Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University,

8   Raleigh, NC 27695, U.S.A.

9   * Corresponding author: yang_zhang@ncsu.edu

10   **Key Points**

- 11   1. A comprehensive decadal evaluation of WRF/Chem over the U.S. using surface and satellite
12   datashows an overall good performance
- 13   2. WRF/Chem outperforms WRF in terms of radiative variables due to the inclusion of
14   chemistry feedbacks to climate
- 15   3. WRF/Chem outperforms CESM in terms of domain-average performance statistics, and
16   temporal/spatial variations due to a finer grid scale

17
18   **Abstract**

19   An advanced online-coupled meteorology-chemistry model, i.e., the Weather Research
20   and Forecasting Model with Chemistry (WRF/Chem), is applied for current (2001-2010) and
21   future (2046-2055) decades under the representative concentration pathways (RCP) 4.5 and 8.5
22   scenarios to examine changes in future climate, air quality, and their interactions. In this Part I
23   paper, a comprehensive model evaluation is carried out for current decade toassessthe
24   performance of WRF/Chem and WRFunder both scenarios and the benefits of downscaling the
25   North Carolina State University's (NCSU) version of the Community Earth System Model

(CESM_NCSU) using WRF/Chem. The evaluation of WRF/Chem shows an overall good performance for most meteorological and chemical variables on a decadal scale. Temperature at 2-m is overpredicted by WRF (by ~0.2 to 0.3°C) but underpredicted by WRF/Chem (by ~0.3 to 0.4 °C), due to higher radiation from WRF. Both WRF and WRF/Chem show large overpredictions for precipitation, indicating limitations in their microphysics or convective parameterizations. WRF/Chem with prognostic chemical concentrations, however, performs much better than WRF with prescribed chemical concentrations for radiation variables, illustrating the benefit of predicting gases and aerosols and representing their feedbacks into meteorology in WRF/Chem. WRF/Chem performs much better than CESM_NCSU for most surface meteorological variables and O₃ hourly mixing ratios. In addition, WRF/Chem better captures observed temporal and spatial variations than CESM_NCSU. CESM_NCSU performance for radiation variables is comparable to or better than WRF/Chem performance because of the model tuning in CESM_NCSU that is routinely made in global models.

KEYWORDS: WRF/Chem, WRF, CESM, the Continental U.S., decadal evaluation, air quality-climate interactions

1. Introduction

The Representative Concentration Pathways (RCPs) are a set of four pathways that lead to radiative forcing levels of 2.6, 4.5, 6.0 and 8.5 W m⁻² by 2100 developed as part of the Intergovernmental Panel on Climate Change (IPCC) fifth assessment report (AR5). They contain current and future concentrations and emissions of greenhouse gases (GHGs) and air pollutants, as well as land-use trajectories developed for the climate modeling community for near-term and long-term modeling experiments [van Vuuren *et al.*, 2011]. A major motivation for the near-term RCP scenarios is to understand the effects of emissions on air quality and their trends while the

motivation for the long-term RCP scenarios is to understand risks from climate change impacts [Moss *et al.*, 2008]. A number of model simulations have been completed under the RCP scenarios using various global and regional climate models (GCMs and RCMs, respectively). Regional models are able to account for more detailed vegetation and terrain features as well as smaller-scale atmospheric processes such as mesoscale weather phenomena; they produce model outputs that relate more closely to observations compared to coarser global model predictions [Feser *et al.*, 2011]. Another advantage of RCMs is that they can be applied at a higher resolution (36 km to a few kilometers) compared to traditional GCMs at resolutions of 1 to 2 degrees [Flato *et al.*, 2013]. RCMs are usually used to dynamically downscale GCMs simulations to provide more detailed information over a region of interest [Rummukainen, 2010]. However, in the abovementioned IPCC study, all the RCMs are applied over Japan or Europe [Kanada *et al.*, 2008; Christensen *et al.*, 2010; Kusaka *et al.*, 2010; van Roosmalen *et al.*, 2010; Kendon *et al.*, 2012]; or evaluated in terms of temperature or precipitation on a climactic scale [Kanada *et al.*, 2008; Kusaka *et al.*, 2010; van Roosmalen *et al.*, 2010; Kendon *et al.*, 2012].

In this study, the North Carolina State University's modified Community Earth System Model (CESM_NCSU) is downscaled using the NCSU's modified online-coupled Weather Research and Forecasting Model with Chemistry (WRF/Chem) model to simulate air quality and climate over continental U.S. (CONUS) for current (2001-2010) and future (2046-2055) decades under the RCP4.5 and RCP8.5 scenarios. WRF/Chem simulates detailed tropospheric gas and aqueous-phase chemistry and aerosol microphysics together with meteorology at each model time step. It accounts for the impacts of prognostic gas and aerosol concentrations on radiation and clouds through direct and indirect feedback mechanisms. Model simulations using WRF/Chem version 3.6.1 or older over CONUS have been evaluated for one or multiple year-

long periods using emissions based on the U.S. National Emission Inventories (NEI)[e.g., *Im et al.*, 2014a, b; *Yahya et al.*, 2014, 2015a, b] and a decade-long period(2001-2010) using emissions based on RCP8.5. Decadal current and future climate simulations over CONUS under multiple RCP scenarios have not been performed using advanced online-coupled meteorology-chemistry models such as WRF/Chem, because such simulations are computationally veryexpensive. While it is conventional to perform climate simulations for a period of 30 or more, the current computational constraints do not permit simulations over such long timescale using a RCM with very detailed chemistry and aerosol treatments (e.g., WRF/Chem). As the first step towards to long-term climate simulations using very detailed RCMs, the decadal simulations are performed using WRF/Chem, which can be extended to longer periods when computational resources become available.

Compared to previous regional climate simulations reported in the literature, there are several additional features of this work. First, both CESM and WRF/Chem used in this work are online-coupled climate-chemistry models whereas other future air quality and/or climate simulations use offline-coupled models, or modeling systems that contain either no orhighly-simplified representations of chemistry and aerosol processes. Second, CESM_NCSU is downscaled using WRF/Chem with similar gas-phase chemical mechanisms and aerosol treatments, which reduces uncertainties in regional model predictions due to inconsistent model treatments in both regional and global models. Third, the downscaling at a $36 \text{ km} \times 36 \text{ km}$ resolution with hourly output using WRF/Chem allows for a comprehensive evaluation of a suite of variables against hourly observations that are not possible using CESM or other RCMs at coarser spatial and temporal resolutions. The results are presented in Part 1 paper with a focus on

the evaluation of current decade results and Part 2 paper with a focus on projected changes in climate, air quality, and their interactions for future decade.

2. Model Configuration and Simulation Design

The model used in this study is a modified version of WRF/Chem v3.6.1 described in *Yahya et al.*[2016], which includes several updates implemented by *Wang et al.*[2014]. Compared to the standard version of WRF/Chem v3.6.1 released by NOAA, a major difference in this modified version of WRF/Chem v3.6.1 is the use of the extended Carbon Bond 2005 (CB05) of *Yarwood et al.* (2005) gas-phase mechanism with chlorine chemistry of *Sarwar et al.*[2007] for gas-phase chemistry. More information on the physics and chemistry modules can be found in Section S1 in the Supplementary material. RCP emissions used in this study are first downloaded from the website at (<http://tntcat.iiasa.ac.at:8787/RcpDb/dsd?Action=htmlpage&page=welcome>). They are then processed by (i) mapping the RCP species to CB05 species in WRF/Chem, (ii) re-gridding the RCP emissions from the $0.5 \times 0.5^\circ$ grid resolution to the 36×36 km grid resolution used for regional simulations over CONUS, and (iii) applying species and location-dependent temporal profiles to the regridded RCP emissions. More details regarding the model configurations as well as the preparation of anthropogenic emissions from RCP can be found in *Yahya et al.*[2016]. Biogenic emissions are calculated online using the Model of Emissions of Gases and Aerosols from Nature version 2 [*Guenther et al.*, 2006]. Dust emissions are based on an online dust scheme in WRF/Chem, i.e., the Atmospheric and Environmental Research Inc. and Air Force Weather Agency [AER/AFWA] scheme [*Jones and Creighton*, 2011]. Emissions from sea salt are generated based on the scheme of *Gong et al.*[1997].

The chemical and meteorological initial and boundary conditions (ICONS/BCONs) for WRF/Chem simulations for both current and future decades come from the modified CESM/CAM version 5.3 with updates by *Gantt et al.*[2014], *He and Zhang*[2014], and *Glotfelty et al.*[2016a, b](hereafter CESM_NCSU) [Glotfelty, T., J. He, and Y. Zhang, 2016a, New Organic Aerosol Treatments in CESM/CAM5: Development and Initial Application, in preparation; Glotfelty, T., J. He, and Y. Zhang, 2016b, The Impact of Future Climate Policy Scenarios on Air Quality and Aerosol/Cloud Interactions using an Advanced Version of CESM/CAM5: Part I. Model Evaluation for the Current Decadal Simulations, Atmospheric Environment, Submitted]. The mozbc preprocessor developed by the National Center for Atmospheric Research (NCAR) is modified slightly and is used to process the 6-hourly CESM_NCSU output into WRF/Chem meteorological and chemical ICONs/BCONs. The gas-phase mechanism in CESM_NCSU is based on the 2005 Carbon Bond Mechanism for Global Extension (CB05_GE) from *Karamchandani et al.*[2012] while the aerosol-phase mechanism is based on the 7-mode modal aerosol module (MAM7) with a new SOA module based on the VBS method. While CB05_GE used in CESM-NCSU contains very similar sets of chemical reactions as CB05 used in WRF/Chem, both MAM7 and MADE use the modal representation for aerosol size distributions, and the VBS method for SOA formation used in MADE is similar to that used in MAM7. One major difference between the two VBS methods is that MADE/VBS does not simulate SOA from primary OA, which is treated in MAM7. The CESM_NCSU simulations are conducted over the same current and future decades at a horizontal resolution of $0.9^{\circ} \times 1.25^{\circ}$ and a vertical resolution of 30 layers under the two RCP scenarios. The results of CESM_NCSU simulations under the two RCP scenarios are described in *Glotfelty et al.*[2016b] and *Glotfelty and Zhang*[2016][Glotfelty, T. and Y. Zhang, 2016, The Impact of Future Climate Policy

Scenarios on Air Quality and Aerosol/Cloud Interactions using an Advanced Version of CESM/CAM5: Part II. Future Trend Analysis and Impacts of Projected Anthropogenic Emissions, Atmospheric Environment, submitted].

The meteorological ICONs/BCONs derived from CESM_NCSU results for current decade simulations are bias-corrected using the simple bias correction technique of *Xu and Yang*[2012] before they are used for WRF/Chem downscaling simulations. Temperature, water vapor, geopotential height, wind, and soil moisture variables from the National Center for Environmental Protection Final Reanalysis (NCEP FNL) dataset, which is available every 6 hours, are used to correct the ICONs and BCONs for WRF/Chem simulations generated by CESM_NCSU. The CESM_NCSU ICONs and BCONs are averaged for each 10-year climatological month from 2001 to 2010. There are two sets of NCEP FNL data including the averages of each climatological month from (i) 1995 to 2010, and from (ii) 2001 to 2014. The differences between the first set of the NCEP FNL data and the CESM_NCSU results between 2001 and 2005 are used to correct CESM_NCSU meteorological ICONs/BCONs for each month from 2001 to 2005. Similarly, the differences between the second set of the NCEP FNL data and the CESM_NCSU results between 2006 and 2010 are used to correct CESM_NCSU ICONs/BCONs from 2006 to 2010. The bias correction method used in this study is different from that used in *Yahya et al.*[2016] in two folds: (1) more years of NCEP FNL data are used (1995-2014 vs. 2001-2010) and (2) the bias correction period is split between 2001 to 2005 and 2006 to 2010 in this work, comparing to a single time period of 2001-2010 used in *Yahya et al.*[2016]. The simulations are performed at a horizontal resolution of 36-km with 148×112 horizontal grid cells over the CONUS domain and parts of Canada and Mexico, and a vertical resolution of 34 layers from the surface to 100-mb. Following *Yahya et al.* [2016], the

simulations are spin-up for 1 month to allow the chemical fields to stabilize. The chemical fields are also allowed to change with time continuously during the entire simulation period while the meteorology is reinitialized with the bias-corrected meteorological ICONs based on CESM_NCSU simulations every month to constrain the meteorological fields so that they remain close to the bias-corrected CESM_NCSU fields. For WRF/Chem future decade simulations, the meteorological ICONs and BCONs from CESM future decade simulations are bias-corrected with the same magnitudes as the differences between NCEP data and the meteorological ICONs and BCONs used for the WRF/Chem current year simulations.

In Section 5, simulation results with WRF v3.6.1 (without chemistry) are compared with those of WRF/Chem v3.6.1 to assess the effects of prescribed gas and aerosol concentrations used in WRF v3.6.1 vs. prognostic gas and aerosol concentrations and their feedbacks to meteorology simulated in WRF/Chem. WRF/Chem v3.6.1 prescribes GHGs concentrations but simulates trace gas and aerosol concentrations as prognostic variables, whereas the WRF v3.6.1 prescribes the concentration of GHGs, O₃, aerosol number, and cloud droplet number over land and the ocean. For example, the Morrison scheme in WRF v3.6.1 uses a constant CDNC value of 250 cm⁻³ if there is no coupling with chemistry. Compared to WRF that simulates meteorological processes only, WRF/Chem simulates additional processes such as gas- and aqueous-phase chemistry, aerosol microphysics, and feedback mechanisms of gases and aerosols to clouds, radiation, and other meteorological variables and processes. The WRF only and WRF/Chem simulations use the same physics and dynamics options, the meteorological ICONs/BCONs derived from CESM_NCSU, as well as the reinitialization method.

The model evaluation protocol and analysis methodology can be found in Section S2 in the Supplementary material.

3. Differences in RCP4.5 and RCP8.5 Current Year emissions over the continental U.S.

Glotfelty et al. [2016a, b] compared changes between RCP4.5 and RCP8.5 global emissions for current and future decades. For the current decade simulations, RCP emissions are available for the years 2000, 2005, and 2010. As shown in Figures S1, S2, and S3, there are negligible differences between RCP4.5 and RCP8.5 emissions for the year 2000 over CONUS; however there are relatively significant differences for the current years 2005 and 2010. For the year 2000, the relative percentage differences in the domain-average emissions of species between the RCP4.5 and RCP8.5 are $\leq 0.2\%$ except for SO_2 , whose difference is 2.8%. For 2005, the percentage differences in the domain-average emissions of species between RCP4.5 and RCP8.5 are mostly within 6% except for carbon monoxide (CO) and volatile organic compounds (VOCs) whose differences are $>50\%$. For 2010, while the differences in the domain-average emissions of NO and NO_2 between RCP4.5 and RCP8.5 remain small, large differences exist in the domain-average emissions of most species, e.g., -63.5% for CO, -30.3% for SO_2 , -97.8% for VOCs, -31.6% for EC, and -22.8% for SO_4^{2-} , indicating higher emissions for RCP4.5 compared to RCP8.5. These differences will influence the evaluation and therefore model performance for the current decade simulations. In addition, the differences for the current decade emissions will have to be taken into account when comparing the differences between RCP4.5 and RCP8.5 for future decade simulations.

4. Evaluation for Current Climate Simulations

4.1. Surface meteorological predictions

Yahya et al.[2016] conducted a comprehensive evaluation for decadal climate simulations based on the RCP8.5 emissions, which used a different set of results from CESM_NCSU as ICONs and BCONs and somewhat different bias correction approach comparing to the RCP8.5 simulation in this work. This section will serve to complement the study by *Yahya et al.*[2016] by including the model performance for RCP4.5 simulations and providing updated performance evaluation and additional analyses based on the latest RCP 8.5 simulations. It will focus on the overall general model performance during the current decade, which will serve as a basis to understand the impact of model performance on future climate and air quality predictions. As described in Section 3, there are also slight differences in the model performance for RCP4.5 and RCP8.5 simulations for the current climate due to similar emissions for most species and small impacts of differences in the emissions of several species on simulated meteorology under the two scenarios. Table 1 summarizes the statistics from the climatological evaluation from both RCP4.5 and RCP8.5 current decade simulations for surface meteorological variables, aerosol-cloud-radiation variables, dry and wet deposition, as well as surface and column mass concentrations of chemical variables. From the statistics, the surface meteorological variables perform similarly for RCP4.5 and RCP8.5 simulations, with only a difference of 0.1°C for T2 on a climatological timescale. The model in general performs well for T2, with a slight underprediction (MBs of -0.3 and -0.4°C, and NMBs of -2.3% and -2.9%, respectively). The model tends to slightly overpredict RH2 and WS10 with NMBs within 11%, but largely overpredict precipitation with NMBs of 34.6-189.8%. WD10 is predicted to be more southerly compared to observational data.

Figure 1 shows the scatter plots for climatologically-averaged simulated and observed T2, RH2, and WS10 at all NCDC sites; and precipitation at all NADP sites. While there are

slight differences in the magnitude of the model predictions for the surface meteorological variables, the overall model performance and trends are similar and consistent for RCP4.5 and RCP8.5 current decade simulations. The model performs very well for the range of T2 values. The model tends to overpredict RH2 across the range of RH2 values of 30-90%. For WS10, overpredictions occur at more sites than underpredictions, leading to relatively small net overpredictions in domain-average WS10. As shown in *Yahya et al.*[2016], the large overpredictions occur at sites that are located along the coastlines in the eastern and northern portions of the domain. The large underpredictions occur mainly over the Great Plains in the central U.S. Precipitation is significantly overpredicted against NADP data from the whole range of values for precipitation. However, as shown in Figure 1, the model predicts slightly better precipitation (closer to the 1:1) line for RCP4.5 compared to RCP8.5. The model also performs better when evaluated against PRISM as compared to GPCP and NCDC. The PRISM dataset is only over the land, the evaluation against PRISM thus reflects the model performance over the land. Precipitation in a significant portion of the ocean domain over the eastern side is overpredicted, which is reflected in the precipitation statistics against GPCP. NADP is weekly precipitation, as compared to NCDC which is hourly precipitation. As WRF/Chem tends to overpredict precipitation intensity [*Wang and Kotamarthi, 2015*], this helps to explain the larger overpredictions in hourly precipitation as compared to weekly precipitation. In addition, NCDC sites are more intensive (>1000) compared to NADP sites (~250 sites). According to *Yahya et al.*[2016], the large overpredictions in hourly precipitation are attributed to the choice of cumulus parameterization scheme. In this case, the use of Grell 3D scheme contributes to the large overpredictions over the Atlantic. *Yahya et al.*[2016] showed that the use of a different cumulus

parameterization scheme, e.g., the Multi-scale Kain Fritsch [MSKF] scheme [Zheng *et al.*, 2016] available in WRF/Chem v3.7 and onwards improved the precipitation predictions.

4.2. Air quality predictions

Comparing to differences in simulated meteorology, there are larger differences in simulated chemical concentrations between RCP4.5 and RCP8.5 due to the differences in emissions as described in Section 3. For hourly, maximum 1-hr, and maximum 8-hr O₃ mixing ratios evaluated against CASTNET and AQS, model performances for RCP4.5 simulation are better compared to that for RCP8.5 simulation for the evaluation against CASTNET, which is not the case for the evaluation against AQS. For example, the NMBs for maximum 8-hr O₃ mixing ratio are -9.9% for RCP4.5 and -11.0% for RCP8.5 against CASTNET and 4.9% for RCP4.5 and 3.5% for RCP8.5 against AQS. The model slightly underpredicts against CASTNET but slightly overpredicts against AQS. As mentioned in *Yahya et al.* [2016], the differences in the model performance trends between CASTNET and AQS evaluation are possibly due to the nature of the sites for each network, namely, the AQS network includes urban, suburban, and rural sites (>1000 sites) while the CASTNET network is mainly made up of rural sites (< 100 sites). Figure 2 shows scatter plots of maximum 1-hr and 8-hr O₃ mixing ratios against AQS. The performances for both RCP4.5 and RCP8.5 simulations are very close, however the RCP4.5 simulation seems to be generally slightly more overpredicted across all sites compared to RCP8.5 (i.e., markers for RCP4.5 are slightly shifted to the right of the markers for RCP8.5), likely due to higher emissions of CO and VOCs (see Figures S1, S2 and S3). Figure 3 compares the climatological time series of observed and simulated maximum 8-hr O₃ mixing ratios at various sites from CASTNET and AQS. There are noticeable variations between RCP4.5 and RCP8.5 maximum 8-hr O₃ mixing ratios on a climatological hourly basis, however, on average,

the trends are similar. The model underpredicts O₃ mixing ratios for both RCP4.5 and RCP8.5 in Yellowstone National Park (YEL), a rural CASTNET site during the warmer months. However, at the AQS sites including the suburban site Port Allen in Texas (PAT) and urban sites such as Huntsville in Tennessee (HUN), Los Angeles in California (LAC) and Clearmont in Florida (CLE), the model tends to overpredict or slightly overpredict maximum 8-hr O₃ mixing ratios during the summer. The model performs relatively well for both RCP4.5 and RCP8.5 in Chicago, Illinois (CHI) on average throughout the year. The slight overpredictions for the RCP4.5 simulation compared to the RCP8.5 simulation occur more obviously in the suburban sites PAT, HUN, and CLE from August to November. However, the statistical, spatial, and temporal evaluations show overall minor differences between RCP4.5 and RCP8.5 simulations, with similar magnitudes of underpredictions and overpredictions at each site. This indicates consistency in errors, if any, in input emissions and ICONs/BCONs for both RCP4.5 and RCP8.5, or inherent model limitations such as the tendency of the model for precipitation overpredictions, or very low or very high surface wind speeds, as well as biases in temperature and radiation [Yahya *et al.*, 2016].

Table 1 shows that PM₁₀ is largely underpredicted with NMBs of -58% for both RCP4.5 and RCP8.5, likely due to the treatment of coarse inorganic and organic aerosols as non-reactive tracers in the MADE/VBS module. PM_{2.5} concentrations are slightly overpredicted against IMPROVE but underpredicted against STN for both RCP4.5 and RCP8.5. However, as shown in Figure 2, the differences between PM_{2.5} performance against IMPROVE and STN for RCP4.5 and RCP8.5 do not follow a consistent trend, and there are no clear indications whether the RCP4.5 simulation or the RCP8.5 simulation performs better. Model simulation for RCP8.5 compared to that for RCP4.5 produces higher PM_{2.5} and SO₄²⁻ against IMPROVE and higher

306 SO_4^{2-} against STN, due to the influence from SO_2 and SO_4^{2-} emissions which are higher for
307 RCP8.5 compared to RCP4.5 in 2000 and 2005 (Figures S1 and S2). For the other species,
308 RCP4.5 produces higher concentrations compared to RCP8.5 due to the influence of emissions.
309 Overall, the model tends to overpredict SO_4^{2-} against both IMPROVE and TC. Model
310 performance for NO_3^- is poor as the capability of thermodynamic models to reproduce the
311 partitioning of total nitrates between gas and aerosol phases would differ across the continental
312 U.S. depending on local chemical and meteorological conditions (Yu et al., 2005). The
313 underpredictions of PM_{10} and many $\text{PM}_{2.5}$ species can also be attributed to the overpredictions of
314 WS10 and precipitation.

315   Figure 4 compares the climatological time series of simulated and observed $\text{PM}_{2.5}$
316 concentrations at a number of IMPROVE and STN sites. Similar to the trends of maximum 8-hr
317 O_3 mixing ratios, there are variations in the simulated $\text{PM}_{2.5}$ concentrations between the RCP4.5
318 and RCP8.5 simulations, however, on average the trends are similar for IMPROVE sites such as
319 Yosemite National Park in California (YOSE1) and James River Face Wilderness in Virginia
320 (JARI), as well as for STN sites such as Corpus Christi in Texas (CCT), in Birmingham in
321 Alabama (BAL), and Sacramento in California (SAC). In New Brunswick in New Jersey (NBR),
322 however, the variations especially in the peak $\text{PM}_{2.5}$ concentrations are more obvious compared
323 to those at the other sites. The model largely underpredicts the peak $\text{PM}_{2.5}$ concentrations for
324 both RCP4.5 and RCP8.5 in YOSE1 in June-July, JARI in August-September, with particularly
325 large underpredictions during the winter months for SAC. Such large underpredictions at SAC
326 are likely due to the underpredicted the emissions of primary OA or gaseous precursors of
327 secondary $\text{PM}_{2.5}$, as errors in meteorology or ICONs/BCONs cannot fully explain large
328 underpredictions of $\text{PM}_{2.5}$ concentrations (e.g., $> 20 \mu\text{g m}^{-3}$). Similar to O_3 , the differences

between model performances for the RCP4.5 and RCP8.5 simulations are minor and with consistent, significant underpredictions or overpredictions for both simulations.

The model largely underpredicts dry deposition when compared against the CASTNET MLM model for SO_2 , HNO_3 , SO_4^{2-} , NH_4^+ and NO_3^- , at times by more than one order of magnitude. However, the magnitudes for dry deposition fluxes are much lower than those from wet deposition; therefore the underpredictions in dry deposition should not affect the PM or gas species to a large extent. Wet deposition for SO_4^{2-} is overpredicted against NADP whereas that for NO_3^- is largely underpredicted, possibly related to the model biases in SO_4^{2-} and NO_3^- concentrations which are overpredicted and underpredicted, respectively. The model performs relatively well in predicting wet deposition for NH_4^+ for both RCP4.5 and RCP8.5 with low NMBs. For column variables, the model tends to underpredict column CO and HCHO but overpredict NO_2 and TOR. The correlation coefficient is 0.9 for the evaluation of column CO, NO_2 , and HCHO against MOPITT, SCIAMACHY, and SCIAMACHY, respectively, but it is negative (-0.3) for TOR, indicating a poor spatial correlation. RCP8.5 also predicts a much higher tropospheric NO_2 due to higher NO_2 emissions compared to RCP4.5 during 2005-2010.

4.3. Radiation and cloud predictions

For radiation variables evaluated against CERES, the performances for both RCP4.5 and RCP8.5 simulations are very similar, with similar correlation coefficients and NMB differences of $< 0.5\%$ for SWDOWN, GSW, GLW, OLR, and SWCF. However, for LWCF, the RCP8.5 simulation has a higher bias (MB of 7.5 Wm^{-2} compared to MB of 6.9 Wm^{-2}) for the RCP4.5 simulation. This is likely due to the influence from higher mixing ratios of gaseous species in the troposphere for RCP8.5 compared to RCP4.5. For example, as shown in the previous section, tropospheric NO_2 column mass concentrations predicted by RCP8.5 simulations are higher

compared to RCP4.5. The model reproduces very well SWDOWN, GSW, GLW, OLR, and CF with NMBs within 7%, and reasonably well for SWCF and LWCF with NMBs < 30%. While the NMBs are within 15% for AOD, 20% for CDNC, and 5% for CWP, their correlation coefficients are low(0.1-0.4). Among all cloud variables, COT has the worst performance with an NMB of -44% and a correlation coefficient of 0.3 for both simulations. According to *Yahya et al.*[2016], the biases in aerosol-cloud variables for WRF/Chem RCP4.5 and RCP8.5 simulations are likely due to uncertainties in satellite retrievals from MODIS, biases in aerosol size distribution, aerosol composition, aerosol water content, and reflectances for the calculation of AOD in WRF/Chem, as well as differences in the calculations of cloud variables such as CDNC between the satellite retrievals and WRF/Chem. For example, satellite retrievals for AOD from MODIS and other satellites are subject to uncertainties due to different calibration methods of radiation sensors, differences in cloud screening algorithms to derive clear sky pixels [*Li et al.*, 2009], the use of standard reference models of aerosol size distributions, and the treatment of the surface effect (e.g., for bright surfaces such as deserts). Model performances for aerosol-cloud variables are similar for both RCP4.5 and RCP8.5 simulations, as aerosol-cloud variables are less sensitive to small differences in simulated aerosol concentrations between the two simulations.

5. Effects of Prognostic and Prescribed Gas and Aerosol Concentrations

Table 2 shows the statistical evaluation of surface meteorological variables, radiation variables, CWP, and CF for WRF v3.6.1 simulations without chemistry, which can be compared to the statistics for WRF/Chem v3.6.1 simulations in Table 1. WRF simulations overpredict T2 by 0.2°C and 0.3°C, while WRF/Chem simulations underpredict T2 by 0.3°C and 0.4°C for

RCP4.5 and RCP8.5 scenarios, respectively. This trend is consistent with the performance for radiation variables. For example, compared to the WRF/Chem predictions, WRF predicts ~30 W m⁻² higher SWDOWN and ~25 W m⁻² higher GSW, which contribute to higher T2. The WRF RCP8.5 simulation also predicts higher (by ~5%) RH2 than the WRF/Chem RCP8.5 simulation due to an increased evaporation associated with higher T2. However, this is not the case for the WRF RCP4.5 simulation. Figure 5 shows the spatial distributions of MBs for T2 and RH2 for both WRF RCP4.5 and RCP8.5 simulations based on the climatological 10-year averages. The spatial distributions for T2 for both simulations are similar. However for RH2, RCP8.5 gives much higher RH2 compared to RCP4.5, especially over the eastern U.S. The trend for CWP against MODIS is similar to that for RH2. RCP8.5 predicts an average of ~11 g m⁻² of CWP higher than RCP4.5 (see Table 2 and Figure 6), indicating a larger amount of moisture in the atmosphere. The higher T2 for both WRF RCP4.5 and RCP8.5 simulations also result in slightly higher predicted WS10 compared to that of WRF/Chem simulations.

As shown in Figure 6, The WRF only RCP 8.5 simulation shows similar large overpredictions in precipitation against GPCP, NCDC, and PRISM compared to the WRF/Chem RCP 8.5 simulation, which indicates that the overpredictions are not due to the presence of prognostic aerosols in WRF/Chem (which might have increased convection). Overpredicted precipitation occurs over the eastern portion of the domain for both WRF and WRF/Chem simulations due mainly to the limitations in the microphysics or convective parameterizations in WRF and WRF/Chem. *Wang and Kotamarthi*[2015] conducted 12-km WRF simulations under RCP4.5 and RCP8.5 scenarios for current year 1995-2004 and also showed wet biases in general for WRF over CONUS with the exception of southeastern U.S., where a dry bias occurs. Their WRF simulations used the ICONs and BCONs from the

Community Climate System Model version 4 (CCSM4), which also gave similar wet biases. There are a number of additional measures taken by *Wang and Kotamarthi*[2015] to reduce precipitation bias which are not used in this study. They include (i) the bias correction in sea surface temperature (SST) and (ii) weak interior nudging above 850 hPa to wavelengths around 1200km with a nudging coefficient of $3 \times 10^{-5} \text{ s}^{-1}$. *Wang and Kotamarthi*[2015] also used the Grell-Devenyi convective parameterization that is the older version of the Grell 3D scheme [Grell and Freitas, 2014] used in this work. With the exception of the evaluation against NCDC for RCP4.5 simulations, all other precipitation evaluations show that WRF predicts much more precipitation compared to WRF/Chem. This is because compared to the prescribed aerosols in WRF, WRF/Chem predicts more aerosols with a more accurate spatial pattern that suppress warm-rain processes by reducing the size of cloud droplets, narrowing the size distribution, and increasing the depth of the mixed-phase region [Tao *et al.*, 2012]. The differences in the trend for evaluation against NCDC for the WRF and WRF/Chem RCP4.5 simulations can arise due to different types of observational data in the NCDC dataset compared to other precipitation datasets. The NCDC dataset is made up of a number of surface sites, and is available on an hourly basis, while GPCP and PRISM are monthly-averaged regridded global datasets. Despite the differences in the simulated precipitation between the WRF RCP4.5 and RCP8.5 simulations shown in Table 2, there are negligible differences between simulated spatial distributions (see Figure 5).

As shown in Tables 1 and 2 and Figures 6 and 7, WRF/Chem performs much better compared to the WRF simulations for most radiation variables, with the exception of OLR. WRF/Chem performances for RCP4.5 and RCP8.5 simulations are also closer in terms of statistics than those of the WRF and WRF/Chem simulations. Similarly, the differences between

WRF RCP4.5 and RCP8.5 statistics for radiation are small as compared to the differences between WRF and WRF/Chem results in general. The NMB for SWDOWN against CERES is close to 0% for WRF/Chem but ~17% for WRF. Figure 6 shows that WRF overpredicts SWDOWN, especially over the southwestern portion of the domain and in general over land. This is likely because WRF does not simulate aerosols (e.g., sulfate and sea salt) over the ocean, which helps to reflect and scatter shortwave radiation away from the surface. Figure 7 compares climatologically-averaged monthly radiation variables from the WRF and WRF/Chem RCP8.5 simulations and the SURFRAD dataset at six surface sites. The findings are consistent with those from CERES, namely, WRF shows a clear overprediction for SWDOWN compared to WRF/Chem at almost all sites. As shown in Figure 7, WRF/Chem performs better for GLW at almost all sites as compared to WRF. The NMBs for GLW are 0.4% and 0.7% for WRF/Chem and -3.3% and -2.9% for WRF. The better performance of WRF/Chem is because the prognostic O₃ profile near the surface in WRF/Chem may be more accurate as compared to the default O₃ profile for WRF from the RRTMG scheme. The underpredictions in GLW by WRF are likely due to the absence of absorbing aerosols such as black carbon and prescribed radiative gases such as O₃, which result in less longwave radiation reaching the surface. However, both WRF and WRF/Chem underpredict OLR at the top of the atmosphere with slightly better performance for WRF (NMBs of -5.8% vs. -6.3%). The OLR is likely to be more affected by stratospheric O₃ which uses the default profile in both WRF and WRF/Chem. In addition, *Montornes et al.* [2015] showed that the O₃ profile provided in WRF with the RRTMG scheme has significant limitations due to the simplified representation of spatial and temporal variability of O₃ concentrations. For example, RRTMG uses a single O₃ profile for all latitudes and seasons. Both WRF and WRF/Chem perform similarly for LWCF. SWCF is overpredicted by WRF/Chem but

underpredicted by WRF. The spatial pattern for SWCF is influenced by CWP more than by CF. WRF and WRF/Chem show similar performance statistics for CF. As shown in Figure 6, compared to MODIS, both WRF and WRF/Chem overpredict CWP in the eastern portion of the domain over the ocean, due to large overpredictions in the cloud rain mixing ratio (QRAIN), which also contributes to the large overpredictions of precipitation and COT over the same area. WRF underpredicts CWP over the western portion of the domain. WRF/Chem also underpredicts CWP in western U.S. but over smaller areas and to a lesser extent. Similar to CWP, WRF and WRF/Chem also overpredict SWCF in the eastern portion of the domain over the ocean and underpredict it in the western portion of the domain. The spatial distributions of simulated Precip, CWP, and SWCF are closely related, as they are affected by the limitations in the microphysics or convective parameterizations in WRF and WRF/Chem. As shown in Figure 7, the differences between WRF and WRF/Chem albedo are not too significant, with slightly better performance by WRF/Chem at the Bondville site in IL. For WRF and WRF/Chem, the largest errors also occur when albedo is high, i.e., during the winter months, likely due to the high spatial variability of snow cover during these months. The surface albedo is likely to affect GSW, for which WRF/Chem performs much better than compared to WRF as shown in Tables 1 and 2 for both RCP4.5 and RCP8.5 simulations.

6. Impact of Downscaling on Regional Air Quality and Climate Predictions

Table 3 summarizes the statistics for several major surface meteorological, chemical, and aerosol-cloud-radiation variables for CESM_NCSU results for both RCP4.5 and RCP8.5 over CONUS. These statistics can be compared with those for WRF/Chem shown in Table 1 to determine if downscaling improves model performance for those variables. As shown in Tables 1

and 3, compared to CESM_NCSU, WRF/Chem performs much better in simulating T2, RH2, and WS10 against NCDC observations in terms of Corr, NMB, and NME for both RCP4.5 and RCP8.5 simulations. For T2, the domainwide MBs are -0.3 and -0.4°C for WRF/Chem and -0.9 and -1.2 °C for CESM_NCSU for the RCP4.5 and RCP8.5 simulations, respectively. Figure 8 compares the spatial distributions of T2, Q2, WS10, planetary boundary layer height (PBLH), and precipitation from WRF/Chem and CESM RCP4.5 and RCP8.5 simulations. Under the same RCP scenario, compared to CESM, WRF/Chem predicts slightly warmer and wetter atmosphere with higher wind speeds and lower PBLH over most of the domain. It also shows more spatial variability in the western portion of the domain for T2, over the oceanic areas off the southern and eastern U.S. coasts for Q2, throughout the domain for WS10 and PBLH, and both western U.S. and eastern oceanic areas for Precip. Note that CESM_NCSU uses different physics options compared to WRF/Chem, e.g., it uses the shallow convection scheme and deep convection scheme of *Park and Bretherton*[2009] and *Zhang and McFarlane*[1995], respectively. Different physics options lead to different meteorological predictions shown in Figure 8.

Figure 9 shows the spatial distributions of MBs for the T2, RH2, and WS10 against NCDC data. CESM_NCSU mainly underpredicts T2 over most of CONUS. CESM_NCSU also underpredicts T2 to a larger extent over eastern and western U.S., but performs better than WRF/Chem over central U.S. In this CESM_NCSU version, the simulated aerosols result in less radiation reaching the ground as compared to the default tuned CESM from NCAR, causing underpredictions of T2. The underpredictions of T2 by CESM_NCSU also result in large overpredictions of RH2, by more than 13% domainwide according to Table 3. As shown in Figure 9, the biases in RH2 are inversely related to those in T2 for both WRF/Chem and

CESM_NCSU, i.e., wherever T2 is underpredicted, RH2 is overpredicted and vice versa. WS10 is overpredicted by WRF/Chem by ~10% but underpredicted by CESM_NCSU by ~20%. WRF/Chem overpredicts WS10 over areas in eastern U.S., but underpredicts WS10 over areas in the central U.S. However, for CESM_NCSU, with the exception of the coast, WS10 is underpredicted over the whole U.S. This is likely due to the coarse grid resolution used in the CESM_NCSU simulation, which cannot resolve the surface features and topography on a fine scale, thus affecting wind speed predictions. CESM also gives uniform distributions of WS10 over different terrains or areas. As shown in Tables 1 and 3, in contrast to WRF/Chem that overpredicts precipitation to a large extent, CESM_NCSU performs well for precipitation against PRISM and GPCP, with correlation coefficients of 0.8 and 0.9 and NMBs within 13% for both RCP4.5 and RCP8.5. Such differences in precipitation predictions are attributed to different physics options used in the two models as mentioned previously. WRF/Chem is shown to overpredict precipitation especially along the coast over eastern U.S. as well as over northwestern U.S. In addition, while WRF/Chem simulates hourly precipitation data which can give rise to more errors due to the errors in frequency and magnitude of precipitation, CESM_NCSU produces monthly-averaged precipitation that can smooth out errors occurring on a shorter time scale.

For chemical species, it is difficult to make an apple to apple comparison between the performances of WRF/Chem and CESM_NCSU because of a use of different output frequencies. The outputs from CESM_NCSU are every six hours, and those from WRF/Chem are hourly. Figure 10 compares the spatial distributions of SO₂, NO_x, NH₃, and CO from both models. There are large differences in simulated mixing ratios of those species in terms of magnitudes and spatial distributions. Compared to CESM, WRF/Chem gives overall higher values for SO₂,

NO_x, and CO but lower values for NH₃ resulted from different meteorology and different wet and deposition fluxes. For example, compared to CESM, WRF/Chem gives much higher PBLH over central and eastern U.S. but slightly lower PBLH over western U.S. (see Figure 8), consequently, it gives much lower mixing ratios of NH₃ over central and eastern U.S. and slightly higher mixing ratios of NH₃ over western U.S. WRF/Chem predicts higher mixing ratios for SO₂, NO_x, and CO because of lower wet and dry deposition fluxes than CESM. Similar to predicted meteorological variables shown in Figure 8, WRF/Chem also shows much more spatial variability in simulated chemical concentrations due to a use of a finer horizontal grid scale.

The National Ambient Air Quality Standards (NAAQS) set by the U.S. Environmental Protection Agency (U.S. EPA) include maximum 1-hr and 8-hr O₃ standards following exposure studies to protect human health, which cannot be calculated from every 6-hr outputs by CESM_NCSU. The 6-hr outputs from CESM_NCSU and hourly outputs from WRF/Chem are compared to hourly observations from surface networks such as CASTNET and AQS. Figure 11 shows the overlay plots for WRF/Chem and CESM_NCSU RCP4.5 simulations for O₃ mixing ratios against AQS and CASTNET. As shown in Table 3 and Figure 11, CESM_NCSU predicts much higher (by ~10 ppb on a domain-average) O₃ mixing ratios as compared to WRF/Chem, with NMBs of >40% and >59% as compared to WRF/Chem with -2% and >2% when evaluated against hourly CASTNET and AQS, respectively. This is likely due to problems in the dry deposition of O₃ mixing ratios in CESM_NCSU [Glotfelty *et al.*, 2016b]. Val Martin *et al.* [2014] showed that correcting the dry deposition in CESM by linking the leaf and stomatal resistances to the leaf area index helped to reduce the summertime surface O₃ bias from 30 ppb to 14 ppb over eastern U.S. WRF/Chem, on the other hand, performs very well in predicting O₃ mixing ratios. For PM_{2.5} concentrations, the performance for CESM_NCSU is slightly better against

IMPROVE but has larger underpredictions against STN as compared to WRF/Chem. While IMPROVE sites are located at national parks and wilderness areas (<http://vista.cira.colostate.edu/improve>), STN sites are located in urban, suburban, or rural areas where the coarser grid resolution of CESM_NCSU cannot well capture site-specific $PM_{2.5}$ concentrations. However, as shown in Figure 11, CESM_NCSU does perform relatively well for $PM_{2.5}$, and the spatial distribution of $PM_{2.5}$ concentrations is comparable and similar to WRF/Chem. WRF/Chem performs better than CESM_NCSU on a smaller scale such as on a state level. For example, over Texas, WRF/Chem is able to capture better the observed $PM_{2.5}$ concentrations compared to CESM_NCSU that underpredicts the $PM_{2.5}$ concentrations.

For the radiation variables, both WRF/Chem and CESM_NCSU use the RRTMG scheme. As shown in Tables 1 and 3, the MB for SWDOWN against CERES for WRF/Chem is smaller than that from CESM_NCSU (-0.2 to 0.1 Wm^{-2} vs. 2.0 - 2.3 Wm^{-2}); however, the correlation coefficient for CESM_NCSU is 1.0 as compared to 0.8 for WRF/Chem. The correlation coefficients are 1.0 for SWDOWN, GSW, GLW, and OLR against CERES for CESM_NCSU but range from 0.6 to 1.0 for WRF/Chem. CESM_NCSU performs well for all the evaluated radiation variables, with either consistent or even better (e.g., GSW, OLR) performance as compared to WRF/Chem. Figure 12 shows the spatial plots for several radiation variables simulated by both CESM_NCSU and WRF/Chem for the RCP4.5 simulation; they are compared with those from CERES. The spatial patterns for radiation variables for CESM_NCSU are similar to CERES. For WRF/Chem, the spatial patterns are generally similar to CERES; however, there are several areas over CONUS where radiation is too low or too high, especially over the ocean over the eastern coast of the U.S., which are likely due to poor performance of cloud variables for WRF/Chem resulted from the inherent limitations of the microphysics or

convective parameterizations, as discussed previously. WRF/Chem overpredicts SWCF by $\sim 8 \text{ W m}^{-2}$, and LWCF by $\sim 7 \text{ W m}^{-2}$, while CESM_NCSU overpredicts SWCF by $\sim 5 \text{ W m}^{-2}$ and performs extremely well for LWCF with MB within 0.5 W m^{-2} .

In this study, WRF/Chem performs relatively poorly for cloud variables when evaluated against MODIS. The WRF/Chem plots for aerosol-cloud variables evaluated against MODIS are extracted from 15Z to 20Z periods only according to the MODIS pass over North America. The cloud variables significantly affect radiation performance by WRF/Chem. Unlike WRF/Chem, CESM_NCSU variables cannot be extracted for 15Z to 20Z over continental U.S. CESM_NCSU aerosol-cloud variables are instead averaged on a monthly basis, and compared to MODIS data, which also gives rise to biases. As shown in Figure 13, AOD spatial pattern for the WRF/Chem RCP4.5 simulation is similar to CESM_NCSU, except for the magnitude of AOD, which could arise due to the differences in averaging. MODIS shows higher AOD over western U.S., which is not the case for both WRF/Chem and CESM_NCSU. The spatial patterns of AOD simulated by both WRF/Chem and CESM_NCSU are similar to those of $\text{PM}_{2.5}$. WRF/Chem predicts CF relatively well as compared to MODIS data. The underprediction of CF for CESM_NCSU is due to the averaging of CESM_NCSU monthly output. CDNC is moderately underpredicted for WRF/Chem but largely overpredicted for CESM_NCSU as compared to MODIS data. The overpredictions in CWP and COT over the ocean in the eastern portion of the domain for WRF/Chem result in overpredicted LWCF and SWCF, as well as underpredicted SWDOWN and OLR over the same region. Overall, over land, CDNC, CWP, and COT are underpredicted for WRF/Chem. CWP is reproduced well in terms of domain-average statistics by WRF/Chem (despite large overpredictions over ocean) but significantly underpredicted by CESM_NCSU. CESM_NCSU gives better agreement for COT against MODIS than WRF/Chem. The correlation

coefficients for CESM_NCSU are also poor for AOD and CF (~0.1), moderate for CDNC (0.4) and CWP (0.6), and relatively good for COT at 0.7. Despite these results, CESM_NCSU performs well for radiation variables because they have been tuned to the current and future year scenarios as a standard practice for global models.

7. Summary and Conclusions

Comprehensive model evaluations are carried out for both RCP4.5 and RCP8.5 current decade simulations using a modified version of WRF/Chem v3.6.1 to assess the model's capability in reproducing current atmosphere. Identifying biases in the current decadal simulations enable assessment of the model's capability for future decade simulations. The differences between the RCP4.5 and RCP8.5 emissions are negligible for the year 2000, but not negligible for 2005 and 2010, with larger emissions for RCP4.5 for many species as compared to RCP8.5. As a result, there exist differences in current decade model performance between the RCP4.5 and RCP8.5 simulations with WRF/Chem. For example, there are slight differences in simulated surface meteorological and air quality variables between current decade RCP4.5 and RCP8.5 simulations, e.g., precipitation, O₃ mixing ratios, and PM_{2.5} concentrations. However, as shown in the statistics, scatter plots, and time series plots, the differences are not too significant. Overall, WRF/Chem performs well for T2 with MBs of -0.3 and -0.4 °C for the RCP4.5 and RCP8.5 simulations, respectively, but not as well for precipitation. Precipitation is generally overpredicted over both land and ocean for both RCP scenarios. WRF/Chem also generally performs well for radiation variables (Corr > 0.5 and NMBs < ±20%), but not as well for aerosol-cloud-radiation variables. O₃ mixing ratios and PM_{2.5} concentrations are generally well predicted for both RCP4.5 and RCP8.5 scenarios with slight differences in model performance due to the different emissions, especially for CO, VOCs, and EC.

Most studies from literature use RCMs which do not simulate trace gases, aerosol, and impacts of chemistry and other processes on those species. Instead, they use prescribed gas and aerosol concentrations, such as the WRF model. The WRF only simulations for RCP8.5 in this study overpredict T2 by 0.2°C. This trend is consistent with the performance for radiation variables for both WRF and WRF/Chem. SWDOWN and GSW are overpredicted for WRF as compared to WRF/Chem when evaluated against CERES. The better performance for most radiation variables for WRF/Chem compared to WRF can be attributed to the more accurate representation of radiative gases and aerosols and their effects on cloud and radiation in WRF/Chem. WRF/Chem is therefore a better choice in simulating current and future decadal simulations as compared to WRF, despite its much higher computational costs. The higher SWDOWN and GSW contribute to higher T2 for WRF. Precipitation predictions for WRF and WRF/Chem are similar for RCP8.5. The overpredictions in precipitation are therefore not due to chemistry or the prognostic gas or aerosol concentrations, but rather due to limitations associated with the microphysics or cumulus parameterization schemes in WRF and WRF/Chem and the use of ICONs/BCONs from CESM simulations (rather than ICONs/BCONs derived from NCEP). The monthly reinitialization of meteorology may also cause a build-up of moisture over the eastern portion of the domain, which contributes to a lesser extent to the overpredictions of precipitation.

Downscaling CESM_NCSU using WRF/Chem also helps to improve model performance for most surface meteorological variables (except for precipitation) and air quality variables including O₃ and PM_{2.5}. The biases in T2 are reduced by > 0.5 °C, even though WRF/Chem underpredicts T2 while CESM_NCSU overpredicts T2, consistent with trends in radiation. CESM_NCSU performs very well for radiation, and better than WRF/Chem as a result of tuning.

CESM_NCSU performs poorly for O₃ mixing ratios against AQS and CASTNET, which is due to problems with dry deposition of O₃ in CESM_NCSU. CESM_NCSU performs slightly better against IMPROVE but much worse against STN for PM_{2.5} concentrations. It is difficult to compare WRF/Chem results against CESM_NCSU for aerosol-cloud-radiation variables against MODIS, due to every 6-hr or monthly resolution of CESM_NCSU results and the MODIS passing time of 6-hour across CONUS. While WRF/Chem gives better agreement for CF, CDNC, and CWP, CESM_NCSU shows lower bias in COT. Large differences exist in simulated meteorology between CESM and WRF/Chem, due mainly to the use of different physics options and horizontal grid resolutions for their simulations, which in turn explains in part the differences in simulated chemical concentrations. Different wet and dry deposition fluxes simulated from both models also help explain in part those differences. The evaluation of this work also indicates that WRF/Chem should be further improved, especially in the representations of precipitation, PM composition, PM₁₀, and aerosol-cloud variables which are important climate drivers.

Acknowledgments

This study is funded by the National Science Foundation EaSM program (AGS-1049200) at NCSU. The emissions for chemical species that are not available from the RCP emissions are taken from the 2008 NEI-derived emissions for 2006 and 2010 provided by the U.S. EPA, Environment Canada, and Mexican Secretariat of the Environment and Natural Resources (Secretaría de Medio Ambiente y Recursos Naturales-SEMARNAT) and National Institute of Ecology (Instituto Nacional de Ecología-INE) as part of the Air Quality Model Evaluation International Initiative (AQMEII). The hourly temporal profiles of 2001, 2005, and 2010 RCPemissions are based on the 2002NEI and the AQMEII 2006 and 2010 emissions derived

based on the 2008 NEI. The authors acknowledge use of the WRF-Chem preprocessor tool mozbc provided by the Atmospheric Chemistry Observations and Modeling Lab (ACOM) of NCAR and the script to generate initial and boundary conditions for WRF based on CESM results provided by Ruby Leung, PNNL. The authors acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and Information Systems Laboratory. GPCP data is provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. PRISM data is obtained from the PRISM Climate Group, Oregon State University at <http://www.prism.oregonstate.edu/>. NADP data is obtained from the website at (<http://nadp.sws.uiuc.edu/nadp/>). CERES and MOPITT CO data is obtained from the NASA Langley Research Center Atmospheric Science Data Center (<https://eosweb.larc.nasa.gov/order-data>). The MODIS data used in this study were acquired as part of the NASA's Earth-Sun System Division and archived and distributed by the MODIS Adaptive Processing System (MODAPS) (<https://ladsweb.nascom.nasa.gov/data/search.html>). NCDC QCLCD data is obtained from the NCDC website at (<http://www.ncdc.noaa.gov/orders/qclcd/>). We also acknowledge the free use of tropospheric NO₂ and HCHO column data from the SCIAMACHY sensor from (www.temis.nl). The data used to generate results presented in this paper will be available free of charge upon request, please contact Yang Zhang at yzhang9@ncsu.edu.

References

Ahmadov, R., S.A. McKeen, A.L. Robinson, R. Bareini, A.M. Middlebrook, J.A. De Gouw, J. Meagher, E.-Y. Hsie, E. Edgerton, S. Shaw, and M. Trainer(2012), A volatility basis set

681 model for summertime secondary organic aerosols over the eastern United States in 2006, J.
 682 Geophys. Res., 117, D06301, doi:10.1029/2011JD016831.

683 Boersma, K.F., H.J. Eskes, and E.J. Brinksma, (2004), Error Analysis for Tropospheric NO₂
 684 Retrieval from Space, J. Geophys. Res., 109, D04311, doi:10.1029/2003JD003962.

685 Christensen, J.H., E. Kjellstrom, F. Giorgi, G. Lenderink, and M. Rummukainen(2010), Weight
 686 assignment in regional climate models, Clim. Res., 44, 179 – 194, doi:10.3354/cr00916.

687 De Smedt, I., J.-F. Muller, T. Stavrou, R.J. van der A., H.J. Eskes, and M. Van Roozendael
 688 (2008), Twelve years of global observations of formaldehyde in the troposphere using
 689 GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 16, 4947 – 4963.

690 De Smedt, I., M. Van Roozendael, T. Stavrou, T., J.-F. Muller, C. Lerot, N. Theys, P. Valks, N.
 691 Hao, and R.J. van der A. (2012), Improved retrieval of global formaldehyde columns from
 692 GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos.
 693 Meas. Tech., 5, 2933 – 2949, doi:10.5194/amt-5-2933-2012.

694 Feser, G., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn(2011), Regional climate models
 695 add value to global model data, BAMS, September 2011, 1181 – 1192,
 696 doi:10.1175/2011BAMS3061.1.

697 Fishman, J., A.E. Wozniak, and J.K. Creilson, (2003), Global distribution of tropospheric ozone
 698 from satellite measurements using the empirically corrected tropospheric ozone residual
 699 technique: Identification of the regional aspects of air pollution, Atmos. Chem. Phys., 3, 893
 700 – 907.

701 Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S.
 702 Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and
 703 M. Rummukainen(2013),Evaluation of Climate Models. In: *Climate Change 2013: The*

Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Stocker, T.F., D. Qin, G.-K. Plattner, M.Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 741–866, doi:10.1017/CBO9781107415324.020.

Gantt, B., J. He, X. Zhang, Y. Zhang, and A. Nenes(2014), Incorporation of advanced aerosol activation treatments into CESM/CAM5: model evaluation and impacts on aerosol indirect effects, *Atmos. Chem. Phys.*, 14, 7485 – 7497, doi:10.5194/acp-14-7485-2014.

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

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

Guenther, A., T. Kart, P. Harley, C. Wiedinmyer, P.I. Palmer, and C. Geron(2006) Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature)), *Atmos. Chem. Phys.*, 6, 3181-3210.

He, J., and Y. Zhang(2014), Improvement and further development in CESM/CAM5: gas-phase chemistry and inorganic aerosol treatments, *Atmos. Chem. Phys.*, 14, 9171 – 9200, doi:10.5194/acp-14-9171-2014.

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

726 Hong, S.-Y. (2010), A new stable boundary-layer mixing scheme and its impact on the simulated
 727 East Asian summer monsoon, Q.J.R. Meteorol. Soc., 136, 1481 – 1496, doi:0.1002/qj.665,
 728 2010.

729 Im, U., R. Bianconi, E. Solazzo, I. Kioutsioukis, A. Badia, A. Balzarini, R. Baró, R. Bellasio, D.
 730 Brunner, C. Chemel, G. Curci, J. Flemming, R. Forkel, L. Giordano, P. Jiménez-Guerrero,
 731 M. Hirtl, A. Hodzic, L. Honzak, O. Jorba, C. Knote, J. J.P. Kuenen, P. A. Makar, A.
 732 Manders-Groot, L. Neal, J. L. Pérez, G. Pirovano, G. Pouliot, R. San Jose, N. Savage, W.
 733 Schroder, R. S. Sokhi, D. Syrakov, A. Torian, P. Tuccella, J. Werhahn, R. Wolke, K. Yahya,
 734 R. Zabkar, Y. Zhang, J.-H. Zhang, C. Hogrefe, and S. Galmarini (2014a), Evaluation of
 735 operational online-coupled regional air quality models over Europe and North America in the
 736 context of AQMEII phase 2. Part I: Ozone, 115, 404-420, *Atmospheric Environment*, doi:
 737 10.1016/j.atmosenv.2014.09.042.

738 Im, U., R. Bianconi, E. Solazzo, I. Kioutsioukis, A. Badia, A. Balzarini, R. Baró, R. Bellasio, D.
 739 Brunner, C. Chemel, G. Curci, H. D. van der Gon, J. Flemming, R. Forkel, L. Giordano, P.
 740 Jiménez-Guerrero, M. Hirtl, A. Hodzic, L. Honzak, O. Jorba, C. Knote, P. Makar, A.
 741 Manders-Groot, L. Neal, J. L Pérez, G. Pirovano, G. Pouliot, R. San Jose, N. Savage, W.
 742 Schroder, R. S. Sokhi, D. Syrakov, A. Torian, P. Tuccella, K. Wang, J. Werhahn, R. Wolke,
 743 R. Zabkar, Y. Zhang, J.-H. Zhang, C. Hogrefe, and S. Galmarini(2014b), Evaluation of
 744 operational online-coupled regional air quality models over Europe and North America in the
 745 context of AQMEII phase 2. Part II: Particulate Matter, *Atmospheric Environment*, 115, 421-
 746 441, doi: 10.1016/j.atmosenv.2014.08.072.

747 Jones, S. and G. Creighton. (2011)AFWA dust emission scheme for WRF/Chem-GOCART,
 748 2011 WRF workshop, June 20-24, Boulder, CO, USA.

749 Kanada, S., M. Nakano, S. Hayashi, T. Kato, M. Nakamura, K. Kurihara, and A. Kitoh(2008),
 750 Reproducibility of Maximum Daily Precipitation Amount over Japan by a High-resolution
 751 Non-hydrostatic Model, SOLA, 4, 105-108, doi:10.2151/sola.2008-027.
 752 Karamchandani, P., Y. Zhang, S.-Y.Chen, and R. Balmori-Bronson(2012), Development of an
 753 extended chemical mechanism for global-through-urban applications, Atmos. Pollut. Res.,
 754 3(1), 1 – 24, doi: 10.5094/APR.2011.047.
 755 Kendon, E.J., N.M. Roberts, C.A. Senior, and M.J. Roberts(2012), Realism of rainfall in a very
 756 high resolution climate model, J. Clim., 25, 2791 – 5806,
 757 DOI: <http://dx.doi.org/10.1175/JCLI-D-11-00562.1>.
 758 Kusaka, H. T. Takata, and Y. Takane(2010), Reproducibility of regional climate in central Japan
 759 using the 4-km Resolution WRF Model, SOLA, 6, 113-116.
 760 Li, Z., X. Zhao, R. Kahn, M. Mishchenko, L. Remer, K.-H. Lee, M. Wang, I. Laszlo, T.
 761 Nakajima, and H. Maring(2009). Uncertainties in satellite remote sensing of aerosols and
 762 impact on monitoring its long-term trend: a review and perspective, Ann. Geophys., 27, 2755
 763 – 2770, doi:10.5194/angeo-27-2755-2009.
 764 Montornes, A., B. Codina, and J.W. Zack(2015), Analysis of the ozone profile specifications in
 765 the WRF-ARW model and their impact on the simulation of direct solar radiation, Atmos.
 766 Chem. Phys., 15, 2693 – 2707, doi:10.5194/acp-15-2693-2015.
 767 Morrison, H., G. Thompson, and V. Tatarskii. (2009),Impact of cloud microphysics on the
 768 development of trailing stratiform precipitation in a simulated squall line: Comparison of
 769 One- and Two-Moment Schemes, Mon. Wea. Rev., 137, 991-1007,
 770 DOI: <http://dx.doi.org/10.1175/2008MWR2556.1>.

771 Moss, R., M. Babiker, S. Brinkman, E. Calvo, T. Carter, J. Edmonds, I. Elgizouli, S. Emori, L.
 772 Erda, K. Hibbard, R. Jones, M. Kainuma, J. Kelleher, J.F. Lamarque, M. Manning, B.
 773 Matthews, J. Meehl, L. Meyer, J. Mitchell, N. Nakicenovic, B. O'Neill, R. Pichs, K. Riahi, S.
 774 Rose, P. Runci, R. Stouffer, D. van Vurren, J. Weyant, T. Wilbanks, J.P. van Ypersele, and
 775 M. Zurek(2008),*Towards New Scenarios for Analysis of Emissions, Climate Change,*
 776 *Impacts, and Response Strategies*. Technical Summary. Intergovernmental Panel on Climate
 777 Change, Geneva, 25pp.

778 Park, S. and C.S. Bretherton(2009), The university of Washington shallow convection and moist
 779 turbulence schemes and their impact on climate simulations with the community atmosphere
 780 model, *J. Climate*, 22, 3449–3469, DOI: <http://dx.doi.org/10.1175/2008JCLI2557.1>.

781 Pouliot, G., H.A.C.D. van der Gon, J. Kuenen, J. Zhang, M. Moran, P. and Makar(2015),
 782 Analysis of the Emission Inventories and Model-Ready Emission Datasets of Europe and
 783 North America for Phase 2 of the AQMEII Project, *Atmos. Environ.*, 115, 345 - 360,
 784 doi:10.1016/j.atmosenv.2014.10.061.

785 Rummukainen, M. (2010), State-of-the-art with regional climate models, *Clim. Change*, 1, 82 –
 786 96, doi:10.1002/wcc.8.

787 Sarwar, G., and P.V. Bhave(2007), Modeling the effect of chlorine emissions on ozone levels
 788 over the eastern United States, *J. Appl. Meteorol. Climatol.*, 46, 1009 – 1019, DOI:
 789 <http://dx.doi.org/10.1175/JAM2519.1>.

790 Sarwar, G., K. Fahey, S. Napelenok, S. Roselle, and R. Mathur(2011), Examining the impact of
 791 CMAQ model updates on aerosol sulfate predictions, the 10th Annual CMAS Models-3
 792 User's Conference, October, Chapel Hill, NC.

793 Tao, W.-K., J.-P. Che, Z. Li, C. Wang, and C. Zhang(2012), Impact of aerosols on convective
794 clouds and precipitation, *Rev. Geophys.*, 50, RG2001, doi:10.1029/2011RG000369.

795 Val Martin, M., C.L. Heald, and S.R. Arnold(2014), Coupling dry deposition to vegetation
796 phenology in the Community Earth System Model: Implications for the simulation of surface
797 O₃, *Geophys. Res. Lett.*, 41, 2988 – 2996, doi:10.1002/2014GL059651.

798 Van Roosmalen, L., J. H. Christensen, M. B. Butts, K. H. Jensen, and J.C. Refsgaard(2010), An
799 intercomparison of regional climate model data for hydrological impact studies in Denmark,
800 *J. Hydrol.*, 380, 3-4, 406 – 419.

801 Van Vuuren, D.P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G.C. Hurtt, T.
802 Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S.J. Smith, and
803 S.K. Rose(2011), The representative concentration pathways: an overview, *Climatic Change*,
804 109, 5 – 31, doi:10.1007/s10584-011-0148-z.

805 Wang, J., and V. R. Kotamarthi(2015), High-resolution dynamically downscaled projections of
806 precipitation in the mid and late 21st century over North America, *Earth's Future*, 3,
807 doi:10.1002/2015EF00304.

808 Xu, Z., and Z.-L. Yang(2012), An improved dynamical downscaling method with GCM Bias
809 Corrections and Its Validation with 30 years of climate simulations, *J. Clim.*, 25, 6271 –
810 6286, DOI: <http://dx.doi.org/10.1175/JCLI-D-12-00005.1>.

811 Yahya, K., K. Wang, M. Gudoshava, T. Glotfelty, and Y. Zhang(2014), Application of
812 WRF/Chem over North America under the AQMEII Phase 2. Part I. Comprehensive
813 Evaluation of 2006 Simulation, *Atmospheric*
814 *Environment*, doi:10.1016/j.atmosenv.2014.08.063.

815 Yahya, K., K. Wang, Y. Zhang, and T. E. Kleindienst(2015a), Application of WRF/Chem over
 816 North America under the AQMEII Phase 2-Part 2: Evaluation of 2010 Simulation and
 817 Responses of Air Quality and Meteorology-Chemistry Interactions to Changes in Emissions
 818 and Meteorology from 2006 to 2010, *Geoscientific Model Development*, 8, 2095-2117, 2015,
 819 doi:10.5194/gmd-8-2095-2015.

820 Yahya, K., J. He, and Y. Zhang(2015b), Multi-Year Applications of WRF/Chem over
 821 Continental U.S.: Model Evaluation, Variation Trend, and Impacts of Boundary Conditions,
 822 *Journal of Geophysical Research*, 20(24), 12748-12777, 10.1002/2015JD023819.

823 Yahya, K., K. Wang, P. Campbell, T. Glotfelty, J. He, and Y. Zhang(2016), Decadal evaluation
 824 of regional climate, air quality, and their interactions over the continental U.S. using
 825 WRF/Chem Version 3.6.1, *Geosci. Mod. Dev.*, 9, 671 – 695, doi:10.5194/gmd-9-671-2016. .

826 Yarwood, G., S. Rao, M. Yocke, and G.Z. Whitten(2005), Final Report – Updates to the Carbon
 827 Bond Mechanism: CB05, Rep. RT-04-00675 Yocke and Co., Novato, Calif., p. 246.

828 Yu, S., B. Eder, R. Dennis S.-H. Chuand S. Schwartz(2006), New unbiased symmetric metrics
 829 for evaluation of air quality models, *Atmos. Sci. Lett.*, 7, 26 – 34, DOI: 10.1002/asl.125.

830 Zhang, G.J., and N.A. McFarlane(1995). Sensitivity of climate simulations to the
 831 parameterization of cumulus convection in the Canadian Climate Centre general circulation
 832 model, *Atmos. Ocean*, 33, 407 – 446, DOI: 10.1080/07055900.1995.9649539.

833 Zhang, Y., P. Liu, B. Pun, and C. Seigneur (2006), A Comprehensive Performance Evaluation of
 834 MM5-CMAQ for the Summer 1999 Southern Oxidants Study Episode,Part-I. Evaluation
 835 Protocols, Databases and Meteorological Predictions, *Atmospheric Environment*, 40, 4825-
 836 4838, [doi:10.1016/j.atmosenv.2005.12.043](https://doi.org/10.1016/j.atmosenv.2005.12.043).

Zhang, Y., X.-Y. Wen, K. Wang, K. Vijayaraghavan, K., and M. Jacobson(2009), Probing into regional ozone and particulate matter pollution in the United States: 2. An examination of formation mechanisms through a process analysis technique and sensitivity study, J. Geophys. Res., 114, D22305, doi:10.1029/2009JD011900.

Zheng, Y., K. Alapaty, J.A. Herwehe, A.D. Del Genio, and D. Niyogi(2016), Improving high-resolution weather forecasts using the Weather Research and Forecasting (WRF)) Model with an updated Kain-Fritsch Scheme, Mon. Wea. Rev., 144, March Issue, 833-860,doi:10.1175/MWR-D-15-0005.1.

Table 1. Summary of statistics for surface meteorological and chemical variables, aerosol-cloud-radiation variables against satellite data, dry and wet deposition, and column gases for WRF/Chem RCP4.5 and RCP8.5 simulations.

Species/Network/Units	Mean Obs	RCP4.5					RCP8.5				
		Mean Sim	Corr	MB	NMB (%)	NME (%)	Mean Sim	Corr	MB	NMB (%)	NME (%)
T2 NCDC (°C)	12.9	12.6	1.0	-0.3	-2.3	7.6	12.5	1.0	-0.4	-2.9	7.7
RH2 NCDC (%)	68.7	70.8	0.8	2.1	3.0	6.6	70.8	0.8	2.1	3.0	6.6
WS10 NCDC (m s⁻¹)	3.5	3.9	0.3	0.3	9.4	24.7	3.9	0.3	0.4	10.7	25.2
WD10 NCDC (deg)	155.5	178.8	0.2	23.3	15.0	19.0	178.5	0.2	23.0	14.8	18.7
Precip NADP (mm week⁻¹)	18.0	26.0	0.6	8.1	45.0	50.2	26.8	0.6	8.8	49.3	55.2
Precip GPCP (mm day⁻¹)	2.3	4.0	0.6	1.7	75.9	86.3	4.1	0.6	1.9	82.4	93.0
Precip PRISM (mm day⁻¹)	2.1	2.9	0.7	0.7	34.6	39.7	2.9	0.7	0.8	37.6	43.6
Precip NCDC (mm hr⁻¹)	0.07	0.2	0.4	0.1	189.8	194.3	0.16	0.3	0.1	126.1	133.6
CERES SWDOWN (W m⁻²)	184.1	184.2	0.8	0.1	0.0	8.1	183.9	0.8	-0.2	0.0	8.0
CERES GSW (W m⁻²)	157.5	151.0	0.8	-6.4	-4.0	9.3	150.9	0.8	-6.5	-4.1	9.3
CERES GLW (W m⁻²)	323.3	324.6	1.0	1.2	0.4	1.7	325.8	1.0	2.5	0.7	1.7
CERES OLR (W m⁻²)	240.0	225.0	0.6	-15.0	-6.2	6.3	224.8	0.6	-15.2	-6.3	6.4
CERES SWCF (W m⁻²)	-41.8	-50.3	0.5	8.5	20.3	30.7	-50.2	0.5	8.3	20.0	30.5
CERES LWCF (W m⁻²)	24.8	31.8	0.6	6.9	28.0	32.6	32.3	0.6	7.5	30.3	33.8
MODIS AOD	0.14	0.12	0.1	-0.0	-14.9	37.5	0.12	0.1	-0.0	-13.0	38.3
MODIS CF	58.3	60.9	0.7	2.6	4.5	12.0	60.9	0.7	2.7	4.6	11.2
MODIS CDNC (cm⁻³)	169.8	137.3	0.4	-32.5	-19.1	38.5	136.9	0.4	-33.0	-19.4	38.5
MODIS CWP (g m⁻²)	179.5	170.4	0.3	-9.1	-5.1	57.0	179.9	0.3	0.3	0.2	59.2
MODIS COT	16.5	9.3	0.3	-7.3	-43.9	53.3	9.4	0.3	-7.1	-43.3	53.2
Hourly O₃ CASTNET (ppb)	35.3	33.1	0.6	-2.2	-6.1	16.6	32.8	0.6	-2.5	-7.2	17.1
Max-1hr O₃ CASTNET (ppb)	47.5	39.8	0.6	-7.7	-16.2	17.7	39.3	0.6	-8.2	-17.3	18.8
Max 8-hr O₃ CASTNET (ppb)	43.4	39.2	0.6	-4.3	-9.9	13.6	38.7	0.6	-4.8	-11.0	14.4
Hourly O₃ AQS (ppb)	30.1	32.9	0.4	2.8	9.3	17.4	32.5	0.4	2.4	7.9	17.3
Max-1hr O₃ AQS (ppb)	49.9	50.8	0.6	0.9	1.8	7.9	50.1	0.6	0.2	0.3	7.8
Max 8-hr O₃ AQS (ppb)	44.4	46.6	0.6	2.2	4.9	9.5	46.0	0.6	1.6	3.5	9.2

Table 1. (cont.) Summary of statistics for surface meteorological and chemical variables, aerosol-cloud-radiation variables against satellite data, dry and wet deposition, and column gases for WRF/Chem RCP4.5 and RCP8.5 simulations.

Species/Network/Units	Mean Obs	RCP4.5					RCP8.5				
		Mean Sim	Corr	MB	NMB (%)	NME (%)	Mean Sim	Corr	MB	NMB (%)	NME (%)
Hourly PM10 AQS ($\mu\text{g m}^{-3}$)	24.6	10.3	0.2	-14.3	-58.1	58.4	10.2	0.2	-14.4	-58.4	58.8
PM _{2.5} IMPROVE ($\mu\text{g m}^{-3}$)	5.5	5.8	0.9	0.3	5.6	20.3	5.9	0.9	0.4	6.7	20.6
PM _{2.5} STN ($\mu\text{g m}^{-3}$)	12.0	10.4	0.6	-1.6	-13.2	22.1	10.3	0.6	-1.7	-13.8	22.1
SO ₄ ²⁻ IMPROVE ($\mu\text{g m}^{-3}$)	1.4	2.0	0.9	0.6	41.5	46.2	2.1	0.9	0.7	50.7	53.6
SO ₄ ²⁻ STN ($\mu\text{g m}^{-3}$)	3.8	4.0	0.6	0.1	3.0	29.8	4.2	0.6	0.4	9.3	31.1
NO ₃ ⁻ IMPROVE ¹ ($\mu\text{g m}^{-3}$)	0.5	0.5	0.8	-0.1	-12.5	44.2	0.5	0.7	-0.0	-9.7	48.0
NO ₃ ⁻ STN ($\mu\text{g m}^{-3}$)	1.5	0.8	0.3	-0.8	-49.2	53.1	0.7	0.3	-0.8	-54.4	57.5
NH ₄ ⁺ IMPROVE ¹ ($\mu\text{g m}^{-3}$)	0.6	0.7	0.6	0.1	23.0	60.0	0.8	0.2	0.2	37.5	75.8
NH ₄ ⁺ STN ($\mu\text{g m}^{-3}$)	1.3	1.1	0.6	-0.2	-13.0	29.7	1.1	0.6	-0.2	-15.1	31.1
EC IMPROVE ($\mu\text{g m}^{-3}$)	0.2	0.2	0.6	-0.0	-20.5	46.0	0.2	0.6	-0.1	-23.5	43.2
EC STN ($\mu\text{g m}^{-3}$)	0.6	0.4	0.3	-0.2	-30.6	49.2	0.4	0.3	-0.2	-35.9	48.4
TOC IMPROVE ($\mu\text{g m}^{-3}$)	1.1	1.0	0.6	-0.1	-13.5	30.9	1.0	0.6	-0.1	-5.5	29.7
TC IMPROVE ($\mu\text{g m}^{-3}$)	1.4	1.2	0.6	-0.2	-14.8	31.4	1.2	0.7	-0.1	-8.9	29.4
TC STN ($\mu\text{g m}^{-3}$)	4.3	2.0	0.2	-2.4	-54.5	54.9	1.9	0.2	-2.4	-56.2	56.6

Table 1. (cont.) Summary of statistics for surface meteorological and chemical variables, aerosol-cloud-radiation variables against satellite data, dry and wet deposition, and column gases for WRF/Chem RCP4.5 and RCP8.5 simulations.

Species/Network/Units	Mean Obs	RCP4.5					RCP8.5				
		Mean Sim	Corr	MB	NMB (%)	NME (%)	Mean Sim	Corr	MB	NMB (%)	NME (%)
CASTNET MLM SO₂DD²(kg/ha)	6.3e-2	2.1e-3	0.8	-0.1	-96.7	96.7	2.2e-3	0.8	-0.1	-96.5	96.5
CASTNET MLM HNO₃DD (kg/ha)	9.4e-2	1.7e-3	0.7	-0.1	-98.2	98.2	1.7e-3	0.6	-0.1	-98.1	98.1
CASTNET MLM SO₄²⁻DD (kg/ha)	1.9e-2	5.0e-4	0.7	-0.02	-97.2	97.2	5.6e-4	0.7	-0.02	-97.0	97.0
CASTNET MLM NH₄⁺DD (kg/ha)	6.6e-3	1.0e-4	0.6	-0.01	-98.4	98.4	1.0e-4	0.6	-0.01	-98.5	98.5
CASTNET MLM NO₃⁻DD (kg/ha)	5.3e-3	5.5e-5	0.5	-0.01	-98.9	98.9	5.2e-5	0.5	-0.01	-99.0	99.0
NADP SO₄²⁻WD²(kg/ha)	0.22	0.30	0.8	0.1	39.3	50.1	0.32	0.8	0.1	47.3	57.1
NADP NH₄⁺WD (kg/ha)	0.057	0.057	0.3	0.0	0.3	57.0	0.053	0.3	0.0	-5.6	56.8
NADP NO₃⁻WD (kg/ha)	0.18	0.03	0.1	-0.2	-84.8	85.3	0.03	0.1	-0.2	-83.3	84.1
MOPITT CO (molec cm⁻²)	1.9e18	1.6e18	0.9	-3.4e17	-17.9	18.1	1.5e18	0.9	-4.4e17	-22.8	22.9
OMI TOR (DU)	30.0	67.1	-0.3	37.1	123.4	123.3	66.0	-0.3	35.9	119.7	119.7
SCIAMACHY NO₂ (molec cm⁻²)	1.5e15	2.1e15	0.9	5.1e14	33.1	40.2	3.2e15	0.9	1.6e15	103.7	105.2
SCIAMACHY HCHO (molec cm⁻²)	5.0e15	3.4e15	0.9	-1.5e15	-31.5	31.5	5.0e15	3.4e15	0.9	-31.2	31.2

Table 2. Summary of statistics for surface meteorological variables, cloud-radiation variables against satellite data for WRF only simulations over continental U.S.

Species/Network/Units	Mean Obs	RCP4.5					RCP8.5				
		Mean Sim	Corr	MB	NMB (%)	NME (%)	Mean Sim	Corr	MB	NMB (%)	NME (%)
T2 NCDC (°C)	12.9	13.2	1.0	0.3	1.9	7.2	13.1	1.0	0.2	1.3	7.3
RH2 NCDC (%)	68.7	69.8	0.8	1.1	1.6	6.5	75.4	0.6	6.7	9.8	15.8
WS10 NCDC (m s⁻¹)	3.5	3.9	0.3	0.4	10.5	29.2	4.0	0.3	0.5	13.2	29.9
Precip GPCP (mm day⁻¹)	2.3	4.1	0.6	1.9	83.9	91.5	4.4	0.6	2.1	93.0	100.8
Precip PRISM (mm day⁻¹)	2.1	3.3	0.7	1.1	52.2	54.6	3.3	0.7	1.2	55.3	57.6
Precip NCDC (mm hr⁻¹)	0.07	0.2	0.4	0.1	141.9	147.6	0.2	0.3	0.1	152.2	158.5
CERES SWDOWN (W m⁻²)	184.1	216.5	0.9	32.4	17.6	18.0	214.7	0.8	30.6	16.6	17.5
CERES GSW (W m⁻²)	157.5	178.4	0.8	20.9	13.3	15.1	176.7	0.8	19.2	12.2	14.9
CERES GLW (W m⁻²)	323.3	312.5	1.0	-10.8	-3.3	3.5	313.8	1.0	-9.5	-2.9	3.2
CERES OLR (W m⁻²)	240.0	227.1	0.7	-12.8	-5.3	5.5	226.0	0.7	-14.0	-5.8	5.9
CERES SWCF (W m⁻²)	-41.8	-31.4	0.3	10.4	-24.8	34.4	-33.3	0.3	-8.5	-20.4	35.3
CERES LWCF (W m⁻²)	24.8	30.1	0.6	5.3	21.5	28.9	31.7	0.6	6.9	28.0	33.1
MODIS CWP (g m⁻²)	179.5	148.6	0.4	-30.9	-17.2	60.4	160.0	0.3	-19.6	-10.9	65.2
MODIS CF	58.3	61.4	0.6	3.1	5.3	14.1	62.1	0.6	3.8	6.5	14.7

Table 3. Summary of statistics for surface meteorological and chemical variables, aerosol-cloud-radiation variables against satellite data, and column gases for CESM RCP4.5 and RCP8.5 simulations over continental U.S.

Species/Network/Units	Mean Obs	RCP4.5					RCP8.5				
		Mean Sim	Corr	MB	NMB (%)	NME (%)	Mean Sim	Corr	MB	NMB (%)	NME (%)
T2 NCDC (°C)	12.9	12.0	0.9	-0.9	-7.0	11.1	11.7	0.9	-1.2	-9.3	12.5
RH2 NCDC (%)	68.7	82.7	0.6	14.0	20.4	22.7	81.8	0.6	13.1	19.1	21.8
WS10 NCDC (m s⁻¹)	3.5	2.8	0.3	-0.7	-20.7	29.7	2.8	0.3	-0.7	-19.9	29.4
Precip GPCP (mm day⁻¹)	2.3	2.3	0.9	0.1	2.7	19.2	2.4	0.9	0.1	6.2	22.7
Precip PRISM (mm day⁻¹)	2.1	2.4	0.8	0.2	9.8	25.8	2.4	0.8	0.3	13.2	26.3
CERES SWDOWN (W m⁻²)	184.1	186.1	1.0	2.0	1.1	3.8	186.4	1.0	2.3	1.2	4.0
CERES GSW (W m⁻²)	157.5	159.7	1.0	2.2	1.4	4.2	158.9	1.0	1.5	0.9	4.6
CERES GLW (W m⁻²)	323.3	320.5	1.0	-2.8	-0.9	1.7	319.7	1.0	-3.6	-1.1	1.9
CERES OLR (W m⁻²)	240.0	232.5	1.0	-7.5	-3.1	3.2	232.5	1.0	-7.4	-3.1	3.2
CERES SWCF (W m⁻²)	-41.8	-47.5	0.8	5.7	13.7	18.6	-47.1	0.8	5.3	12.7	18.5
CERES LWCF (W m⁻²)	24.8	24.9	0.9	0.1	0.3	9.7	24.5	0.8	-0.3	-1.2	11.2
MODIS AOD	0.14	0.08	0.1	-0.1	-46.5	47.0	0.07	0.05	-0.1	-48.4	48.7
MODIS CF	58.3	28.3	0.1	-30.0	-51.5	51.5	27.1	0.3	-31.2	-53.5	53.5
MODIS CDNC (cm⁻³)	169.8	243.9	0.4	74.1	43.6	50.6	233.9	0.4	64.1	37.7	44.8
MODIS CWP (g m⁻²)	179.5	79.3	0.6	-100.2	-55.8	55.9	79.2	0.6	-100.4	-55.9	56.0
MODIS COT	16.5	11.9	0.7	-4.6	-28.0	30.6	11.9	0.7	-4.6	-28.1	31.1
Hourly O₃ CASTNET (ppb)	35.3	49.6	0.6	14.3	40.6	40.6	49.5	0.6	14.2	40.3	40.3
Hourly O₃ AQS (ppb)	30.1	47.9	0.3	17.8	59.2	59.2	48.0	0.3	17.9	59.3	59.3
PM_{2.5} IMPROVE (µg m⁻³)	5.5	5.7	0.8	0.3	4.6	22.3	5.6	0.7	0.1	1.1	25.6
PM_{2.5} STN (µg m⁻³)	12.0	9.0	0.5	-3.0	-25.3	29.6	8.4	0.4	-3.6	-29.8	34.3
MOPITT CO (molec cm⁻²)	1.9e18	1.5e18	0.9	-3.8e17	-19.7	19.9	1.5e18	0.9	-4.4e17	-22.8	22.9
OMI TOR (DU)	30.0	71.6	-0.3	41.5	138.2	138.2	72.4	-0.3	42.3	140.9	140.9
SCIAMACHY NO₂ (molec cm⁻²)	1.5e15	2.4e15	0.9	8.9e14	57.5	61.3	2.5e15	0.9	9.2e14	59.2	63.3
SCIAMACHY HCHO (molec cm⁻²)	5.0e15	4.0e15	0.9	-9.7e14	-19.4	20.9	4.0e15	0.9	-1.0e15	-20.1	21.2

List of figures

Figure 1. Scatter plots of observed vs. simulated surface meteorological variables T2, RH2 and WS10 against NCDC data and precipitation against NADP data.

Figure 2. Scatter plots of observed vs. simulated surface O₃ against AQS and PM_{2.5} against IMPROVE and STN.

Figure 3. Climatological averaged timeseries of maximum 8-hr O₃ concentrations at selected CASTNET and AQS sites.

Figure 4. Climatological averaged 24-hr PM_{2.5} concentrations at selected IMPROVE and STN sites.

Figure 5. Spatial distributions of MBs for T2 and RH2 against NCDC observational data for WRF RCP4.5 and RCP8.5 simulations over continental U.S.

Figure 6. Spatial distributions of current climatological averages simulated by WRF and WRF/Chem under the RCP8.5 scenario versus satellite-derived data over continental U.S.

Figure 7. Scatter plots of WRF and WRF/Chem RCP8.5 average 10-year simulated radiation variables against SURFRAD observations.

Figure 8. Spatial distributions of current climatological averages for meteorological species simulated by WRF/Chem and CESM under the RCP4.5 and RCP8.5 scenarios over continental U.S.

Figure 9. Spatial distributions of MBs for surface meteorological variables against NCDC observational data for WRF/Chem and CESM RCP4.5 simulations over continental U.S.

Figure 9. Spatial distributions of current climatological averages for gas-phase species simulated by WRF/Chem and CESM under RCP4.5 and RCP8.5 scenarios over continental U.S.

Figure 11. Overlay plots from WRF/Chem and CESM RCP4.5 for O₃ against AQS and CASTNET, and PM_{2.5} against STN and IMPROVE over continental U.S.

Figure 12. Spatial distributions of radiation variables over continental U.S. under the RCP4.5 scenario simulated by WRF/Chem and CESM versus derived from CERES satellite retrieval over continental U.S.

Figure 13. Spatial distributions of aerosol-cloud variables over continental U.S. under the RCP4.5 scenario simulated by WRF/Chem and CESM versus derived from CERES satellite retrieval over continental U.S.

Figure 1. Scatter plots of observed vs. simulated surface meteorological variables T2, RH2 and WS10 against NCDC data and precipitation against NADP data.

Figure 2. Scatter plots of observed vs. simulated surface O_3 against AQS and $PM_{2.5}$ against IMPROVE and STN.

Figure 3. Climatological averaged timeseries of maximum 8-hr O_3 concentrations at selected CASTNET and AQS sites.

Figure 4. Climatological averaged 24-hr $PM_{2.5}$ concentrations at selected IMPROVE and STN sites.

Figure 5. Spatial distributions of MBs for T2 and RH2 against NCDC observational data for WRF RCP4.5 and RCP8.5 simulations over continental U.S.

Figure 6. Spatial distributions of current climatological averages simulated by WRF and WRF/Chem under the RCP8.5 scenario versus satellite-derived data over continental U.S.

WRF RCP8.5

WRF/Chem RCP8.5

Figure 7. Scatter plots of WRF and WRF/Chem RCP8.5 average 10-year simulated radiation variables against SURFRAD observations.

Figure 8. Spatial distributions of current climatological averages for meteorological species simulated by WRF/Chem and CESM under the RCP4.5 and RCP8.5 scenarios over continental U.S.

Figure 9. Spatial distributions of MBs for surface meteorological variables against NCDC observational data for WRF/Chem and CESM RCP4.5 simulations over continental U.S.

Figure 10. Spatial distributions of current climatological averages for gas-phase species simulated by WRF/Chem and CESM under RCP4.5 and RCP8.5 scenarios.

Figure 11. Overlay plots from WRF/Chem and CESM RCP4.5 for O_3 against AQS and CASTNET, and PM_{2.5} against STN and IMPROVE over continental U.S.

Figure 13. Spatial distributions of aerosol-cloud variables over continental U.S. under the RCP4.5 scenario simulated by WRF/Chem and CESM versus derived from CERES satellite retrievals over continental U.S.