Reply to Reviewer #1

Anonymous Referee #1

Received and published: 4 September 2015

The subject is appropriate to GMD. This manuscript presents results of the first decadal application of WRF/Chem v3.6.1 with CB05 from 2001 to 2010 over the continental US using the Representative Concentration Path- way (RCP 8.5) emissions. The capability and appropriateness for long term climatological simulations are assessed on the basis of meteorological, chemical, and aerosol-cloud-radiation variables against data from surface networks and satellite retrievals. The results showed that the model performs very well for the 2m temperature (T2) for the 10 year period with only a small cold bias of -0.3 0C. They also found that in general, the model performs relatively well for chemical and meteorological variables, and not as well for aerosol-cloud-radiation variables. A lot of model evaluations have been done with tremendous observational data. Therefore I recommend clearly the acceptance for publication of this manuscript after minor revisions.

Reply:

We thank the reviewer for careful review of this manuscript and valuable comments to improve the quality of manuscript.

We have carefully addressed all the comments raised by the reviewer to improve the presentation quality and organization of our paper. Please see below our point-by-point replies. All page and line numbers in this reply refer to those in the revised manuscript in the track mode.

Several editorial comments for improving the information content C1877 and presentation of the paper are listed as follows:

1. Title: It should be "Decadal evaluation of regional climate, air quality, and their interactions over the continental U.S. using WRF/Chem Version 3.6.1" because this is your study area.

Reply:

The title has been modified as suggested.

2. Abstract: Please summarize the results quantitatively instead of qualitatively such as what do you mean by "slightly overpredicted"?

Reply:

The abstract has been modified to summarize the results quantitatively by including more statistical measures such as values of NMBs and MBs.

3. P6709, L20-24-61: Regarding the online-coupled models, please add discussions about the recent work for the two-way coupled WRF-CMAQ (such as Yu, Shaocai, R.Mathur, J. Pleim, D. Wong, R. Gilliam, K. Alapaty, C. Zhao, and X. Liu, 2014. Aerosol indirect effect on the grid-scale

1

clouds in the two-way coupled WRF-CMAQ: model description, development, evaluation and regional analysis. Atmos. Chem. Phys. 14, 11247–11285, doi:10.5194/acp-14-1-2014.)

Reply:

The above paper has been added to reference and a brief discussion regarding the work is described in lines 70-72 of page 4 as follows:

"For example, the WRF model has been coupled online to the CMAQ model with the inclusion of aerosol indirect effects to study chemistry and climate interactions (Yu et al., 2014)."

4. P6715, L180-21: Please cite the definitions of MB, NMB, RMSE etc for some references (such as Yu, Shaocai, Brian Eder, Robin Dennis, Shao-hang Chu, Stephen Schwartz, 2006. New unbiased symmetric metrics for evaluation of air quality models. Atmospheric Science Letter, 7, 26-34.)

Reply:

The above reference has been added.

5. P6727, L25-25-593: Regarding the bad performance of NO3-, one of the reasons is because of partition of total (HNO3+NO3) between gas and aerosol phases as discussed by Yu et al. (Yu, Shaocai, Robin Dennis, Shawn Roselle Athanasios Nenes, John Walker, Brian Eder, Kenneth Schere, Jenise Swall, Wayne Robarge, 2005. An assessment of the ability of 3-D air quality models with current thermodynamic equilibrium models to predict aerosol NO3- Journal of Geophysical Research, 110, D07S13, doi:10.1029/2004JD004718.). Please add this discussion.

Reply:

The additional discussion regarding the performance of NO_3^- has been added as suggested by the reviewer to lines 566 to 571 of pages 25-26, as follows: "Other possible reasons for the underpredictions of NO_3^- concentrations include both prediction and measurement errors associated with SO_4^{2-} and TNH_4 that can greatly affect the performance of NO_3^-, inaccuracies in the assumptions used in the thermodynamic model (e.g., the assumption that inorganic ions are internally mixed and the equilibrium assumption might not be representative, especially for particles with larger diameters), as well as inaccuracies in T2 and RH predictions (Yu et al., 2005)".

6. Regarding the captions of Figures 1, 7 and 8: Please one sentence to say "the observations are represented by diamonds in the figures".

Reply:

The markers in Figure 1 are not just the observational data, but rather are the spatial distribution of mean biases (MBs) as stated in the figure caption. To avoid the confusion, we added the following sentence into the caption of Figure 1 "Each marker represents

the MB of each variable at each observational site". For Figures 7 and 8, we indicated in the captions that the observation is represented by markers and simulation is represented by the background.

Interactive comment on Geosci. Model Dev. Discuss., 8, 6707, 2015.

C1878

Reply to Reviewer #2

Anonymous Referee #2

Received and published: 15 September 2015

The authors present for the first time a decadal regional chemistry climate simulation including a full coupling of chemistry-aerosol-radiation feedbacks. For this they use the model WRF/Chem. So far WRF/Chem was mainly used for short term studies. The authors analyse some meteorological variable (2m temperature, 10 m wind speed and precipitation), ozone, PM 2.5 and aerosol-cloud-radiation variables and conclude that the performance of the model is good for the meteorological and chemical variables whereas the aerosol-cloud-radiation results should be improved for long-term climate simulations. Alltogether, most of the results are not fully comprehensible as the authors provide not enough details about the procedures used.

Reply:

We thank the reviewer for careful review of this manuscript and valuable comments to improve the quality of manuscript.

We have carefully addressed all the comments raised by the reviewer to improve the presentation quality and organization of our paper. We have also included more details about the methodology in our study. Please see below our point-by-point replies. All page and line numbers in this reply refer to those in the revised manuscript in the track mode.

Especially, more details should be provided about
• the re-initialisation procedure and how this interacts with the ICs/BCs from CESM/CAM5 (including a more quantitive assessment how much the reinitializing frequency changes the results) and

Reply:

Sensitivity simulations for 1 month (July 2005) have been carried out to quantify the differences in the reinitialization frequency, meteorological ICs/BCs and cumulus parameterization subroutines. The results are documented in the last part of the supplementary material. In summary, the monthly reinitialization frequency gives the highest correlation with observational data GPCP and PRISM, however, it also gives large values of normalized mean bias (NMB) and normalized mean error (NME). The use of a 5-day reinitialization helps to reduce both NMB and NME with slight to moderate improvements, it also reduces the R value. Overall, there are no substantial changes in results generated using a 5-day versus a 1-month reinitialization. More discussions regarding this have also been included in our reply in the Scientific Question part.

• about the way the statistics presented in Table 2 has been calculated. Is this really a point-to-point / date-to-date comparison?

Reply:

This has been addressed below in our reply in the scientific question part.

Therefore the article is subject to major revisions from my point of view. A list of the scientific and content related questions follows as well as a list of required technical corrections.

Scientific questions and content-related remarks:
• page 6711, line 2, p. 6714, l. 20: What do you mean by "similar gas-phase chemistry and aerosol treatment"? Which are the differences if they are only "similar" and not identical? Do you still have to map species (if yes, which one and how), or are you using identical species? Please provide more details.

Reply:

Both WRF/Chem and CESM use the CB05 gas-phase mechanism (Yarwood et al., 2005). However, WRF/Chem includes chlorine chemistry from Sarwar et al. (2007), while CESM_NCSU uses a modified version of CB05, the CB05 Global Extension (CB05GE) (Karamchandani et al., 2012). CB05GE includes more bromine associated chemical reactions for the stratosphere, reactions involving mercury species, and additional heterogeneous reactions on aerosol particles, cloud droplets and on polar stratospheric clouds (PSCs), which are more important for global simulations. Both WRF/Chem and CESM_NCSU also use a modal aerosol size representation, rather than a sectional size representation. MADE/VBS is used in WRF/Chem while a 7-mode prognostic Modal Aerosol Model (MAM7) (Liu et al., 2012) is used in CESM_NCSU. Both aerosol modules include sulfate, nitrate, ammonium, black carbon, organic carbon, dust, and sea salts. For gas-phase species, no species mapping are needed at all and for aerosol species only minimal mapping is require (i.e., mapping of the same species for the same aerosol modes). The above information has been added into the revised paper (lines 184-195 of page 9).

• page 6712, line 1-13: What are these re-initialisations good for? First of all, what are you re-initialising? Meteorology? Chemistry? The whole model? From what you write in the paper I understand that you only re-initialise the meteorological, but not the chemical fields. Is this done in order to keep the model near the observed weather? But in this case 1 month should be much too long.

Additionally, in this way the chemical and meteorological variables are not consistent any more.

Please give reasons for this procedure! Personally, I have my doubts, that you can use a model setup including such a procedure for climate applications at all. From what you say later on, the results depend on this re-initialisation frequency what just strengthens my reservations against this procedure. (Especially the "buildup of storm systems, especially over the warm Atlantic" (page 6717, line 27-28) makes me wary.) But I think I cannot really judge until I get more information about the reasons for this procedure and about how this re-initialisation works. Additionally, I do not understand, how this re-initialisation with NCEP data fits with the statement in section 2.2 that you are using ICs/BCs from CESM/CAM5 for meteorology and chemical fields.

Reply:

The reviewer is correct that the reinitialization has only been done for meteorology (it has been stated explicitly in the revised paper). The reinitialization technique was recommended by the original developer of WRF/Chem at NOAA and has been used in the

past extensively for both climate/air quality studies that focus on meteorology-chemistry feedbacks (e.g., Chen et al., 2013; Glotfelty et al., 2014; Penrod et al., 2014; Berg et al., 2015; Forkel et al., 2015; Ritter et al., 2013). In such studies, nudging or FDDA techniques cannot be used as they may quench the feedback effects to a large extent. The use of the reinitialization technique is to provide reasonable meteorological fields while allowing chemistry-meteorology feedbacks within the system. From this perspective, reinitialization technique serves similarly to the nudging technique to constrain the meteorological fields (e.g., wind fields or precipitation) from getting too large discrepancies due to the accumulation of small numerical errors over a long time period, and also to ensure more accurate meteorological fields (typically by comparing with observations) to drive the chemical calculations.

There were some confusions in our original paper regarding reinitialization. The model was reinitialized towards the bias-correct CESM/CAM5 meteorology, instead of NCEP data itself. The biases in CESM/CAM5 predicted meteorology were first corrected using the NCEP data before their use to derive initial and boundary conditions for WRF/Chem simulations (which are referred to as biased-corrected CESM/CAM5 BCs and ICs). We have clarified such confusions in the revised paper. We also added some more details regarding the bias-correction approach used in this work in lines 201 - 212 of pages 9-10, as follows: "Temperature, water vapor, geopotential height, wind, and soil moisture variables available every 6 hours from the NCEP Final Reanalyses (NCEP FNL) dataset are used to correct the ICs and BCs derived based on results of CESM_NCSU for WRF/Chem simulations. In this bias-correction approach, monthly climatological averages for ICs and BCs are first derived from both NCEP and CESM_NCSU cases. The differences between the ICs and BCs from the NCEP and CESM_NCSU climatological averages are then added onto the CESM_NCSU ICs and BCs to generate bias-corrected CESM_NCSU ICs/BCs."

From our past experience by conducting many years of simulations over various geographical locations, to run an online-couple meteorology/chemistry model freely without any reinitialization could generate erroneous meteorological fields and further deteriorate the simulation of air quality. Thus, reinitialization of meteorology is an alternative method to the commonly-used nudging technique to ensure satisfactory meteorological fields and to drive the chemical systems, which eventually make the simulation results of both meteorology/air quality credible and scientifically sound.

The reinitialization frequency may affect the simulation results. More frequent model reinitializations can give predictions of meteorology that are closer to the reference data that provide the ICs and BCs (however, this does not mean necessarily better predictions). We have conducted a few sensitivity simulations to further test the impacts of reinitialization frequency on the simulation results. The comparison of predicted precipitation against GPCP and PRISM shows that the 1-month reinitialization gives the best correlation coefficients (R), 0.5 and 0.7 respectively, compared to the 5-day reinitialization with R values of 0.4 and 0.3, respectively. However, the 5-day reinitialization gives lower NMB and NME compared to the 1-month reinitialization. The WRF/Chem simulation with 1-month reinitialization also gives slightly better spatial distribution of precipitation and other cloud related variables than those using the 5-day reinitialization. Therefore, we chose the 1-month reinitialization for our final production simulations. The above comparison also shows that the reinitialization frequency was not the main reason for the buildup of storm systems over the warm Atlantic as previously thought. Based on additional sensitivity simulations that we carried out, the

cause of the buildup of storm systems is more likely due to the choice of cumulus parameterization scheme in our model. By comparing the sensitivity simulations using the Grell 3D (in this work) and the multi-scale Kain-Fritsch (MSKF) cumulus scheme (which is available in WRF/Chem v3.7 and later), we found that the simulations with MSKF give much lower precipitation amounts, as well as much lower NMB and NME. However, the R value is not as good as for the simulations with the Grell 3D. . In addition, the MSKF scheme does not include aqueous-phase chemistry in convective clouds, which is currently only available in the Grell cumulus parameterization scheme. The results of our sensitivity analysis for precipitation have been included in the Supplementary material (see Section A4).

Therefore, based on our sensitivity simulations and findings, we have revised our conclusions regarding the reasons for the precipitation biases in Section 3.1 in our manuscript.

• page 6712, line 18-19: Why are you using a discrete and not a linear distribution of the emissions over the years? An assumption that the emissions changed linear seems to be more realistic. Especially, as for the first period the emission data is "valid" for the year before the actual period and for the last period for the last year. Only the middle period is centered around the given emission year.

Reply:

First, RCP emissions are discrete, and only available for the years 2000, 2005, and 2010. As we are conducting a "climatological simulation", using the emissions from the representative years should be sufficient to represent the current state of emissions, since different years of our simulation really present more "current" or "future" years instead of a specific year. In addition, all our model evaluations are conducted based on a climatological timescale (i.e., decadal), rather than on individual years. The distribution of emissions might be more important if we are conducting air quality type of studies for a specific year, such as 2001 or 2010.

• page 6712, line 24-25: The resolution of the emission is very similar to that of the model grid. Following the publications by Valari und Menut (2014) this should be assumed to be too coarse to expect really good results.

Reply:

We agree with the reviewer that emissions at a grid resolution that is very similar to that of the model grid may introduce some errors into the chemical modeling. However, the original resolution of the RCP emissions are coarse, i.e., 0.5° × 0.5°, which is the finest grid resolution for RCP emissions We have regridded the RCP emissions to our model resolution, at 36-km by 36-km. 36-km is a reasonable horizontal resolution and well used for many other regional studies over the continental U.S., which is much finer than most of other global climate/air quality applications. In this study, using a 36-km by 36-km horizontal grid resolution yields 148 × 112 grid cells and considering the multiple decadal simulations of WRF/Chem in this work, even with such a resolution, it is already extremely computationally expensive. Reducing further the horizontal grid space will not be feasible. The publication by Valari and Menut (2008) conducted simulations of up to 6-

km by 6-km resolution, is carried out over a much smaller area as compared to the continental U.S. – over a highly urbanized area of 180-km by 180-km.

• page 6713, l. 7-8: "other RCP groups are used to approximate these emissions (Table S1)". Please be more precise: which species are approximated with which RCP group and how?

Reply:

The RCP species used to approximate the CB05 emissions have indeed been listed in Table S1 in our originally-submitted paper (now Table S2 in the revised paper). An example has been given in the previous sentence, "Some of the CB05 species are directly available in RCP; however, others are lumped into RCP groups, for example, the "other alkanals" and "hexanes and higher alkanes" in the RCP groups can be considered to approximately represent the acetaldehyde and higher aldehydes emissions required by CB05, respectively (Table S2)".

• page 6713, l. 15: Is the "simple inverse distance weighting" mass or better flux conserving? Otherwise the amount of emitted substance would be artificially modified due to your choice of model domain.

Reply:

The "simple inverse distance weighting" method is mass-conserving.

• Sect. 3.2.1 / Table 2: More information about how this statistic was calculated would be desirable.

Reply:

Additional details on how the statistics were calculated have been added in lines 225 to 233, pages 10-11, as follows: "For surface networks with hourly data, e.g., NCDC, the observational data are paired up with the simulated data on an hourly basis for each site. The observational data and simulated data are averaged out for each site. The statistics are then calculated based on the site-specific data pairs. The satellite-derived data is usually available on a monthly basis, and the simulated data are also averaged out on a monthly basis. The satellite-derived data are regridded to the same domain and the total number of grid cells is similar to that of the model outputs. The statistics are calculated based on the grid cell pairs (satellite-derived and simulated data pairs). The time dimension is removed for the climatological evaluation, the statistics are based on a site-specific average or a grid cell average."

Technical corrections:
• p. 6711, l. 15: This sentence is unclear. Maybe just a word or two are missing?

Reply:

This sentence has been revised to make it more clear. The revised sentence is as follows: "The main updates include the implementation of an extended version of Carbon Bond 2005 (CB05) (Yarwood et al., 2005) gas-phase mechanism with the chlorine chemistry (Sarwar et al., 2007) and its coupling with the Modal for Aerosol Dynamics in Europe/Volatility Basis Set (MADE/VBS) (Ahmadov et al., 2012)".

• p. 6712, l. 1: "mb"? Better use SI-Units "hPa".

Reply:

The unit has been changed to hPa.

• p. 6712, l. 12: add degree-sign after first 0.5

Reply:

The degree sign has been added.

• Table S1: please use consistent annotations, i.e., if more then one species / modes are named give (Yes, No ,Group) for each individually.

Reply:

The table has been updated to keep the annotations consistent, see Table S2 (which is original Table S1).

• p. 6714, Eq (1): use larger brackets

Reply:

The bracket size has been changed.

• p. 6716, l. 2: consistently write "sulfate (SO2− 4)"

Reply:

SO_4^{2-} has been defined previously in page 11, line 240, so no change was made.

• p. 6716, l.10: "systemetic" ! "systematic"

:

This typo has been corrected.

• p. 6716, l.26 "0 to −3_ C ! −3_ to 0_ C**

Reply:

This has been corrected.

• p. 6717, l. 22: It is unusual to start with Fig. 3d instead of Fig. 1.

Reply:

The discussion on Fig 3d has been moved back towards the end of Section 3.1, so that the discussion starts from Fig 1. now.

• p. 6717, l. 29: "at the coast" not "in".

Reply:

This has been corrected.

• p. 6719, l. 29: "Corr" not introduced.

Reply:

Corr is introduced earlier in line 222 of page 10, it has been renamed as "R".

• p. 6722 ff.: Obviously you replaced something with "AIRS-AQS", because the space in front of "AIRS-AQS" is missing everywhere.

Reply:

This issue seems to be caused by the typeset of the journal since it didn't show in in the word version of our original manuscript. We will make sure the typeset will be done correctly this time.

• p. 6724, l. 3 (and below): It is very unusual to refer to Winter as JFD instead of DJF. Why are you using this notation?

Reply:

We are using JFD as we are averaging January, February and December from the same year.

• p. 6724, l. 15: AIRS in front of AQS missing.

Reply:

This has been fixed.

• p. 6728, l. 10: remove "the" before isoprene.

Reply:

It has been removed.

• p. 6729, l. 14-17: reformulate this sentence. It is not understandable without thinking a long time about grammar and what you like to say.

Reply:

The sentence has been revised as follows "The MODIS AOD, however, shows slightly elevated AOD over eastern U.S., but the magnitudes are not as high as the simulated AOD over eastern U.S. MODIS-derived AOD is also higher over western U.S. compared to eastern U.S., and this trend is not found in the simulated AOD."

• p. 6730, l. 10-15: repetition of p. 6729, l. 9-14 ?

Reply:

The second part has been removed.

• Fig. 1: What are the dots for? Is it mean bias per measurement station? Please be more precise.

Reply:

Yes. An additional sentence has been added to the figure caption, "Each marker represents the MB of each variable at each observational site."

• Fig. 4 - 7: in caption and y axis labels: "AIRS" missing in front of "AQS".

Reply:

"AIRS" has been added to the captions and y-axis labels.

• Fig. 7 / 8: explain what are the dots. I assume the model results are the 2d plot and the observations are the dots, but you never write that.

Reply:

We have indicated in the captions that the observation is represented by markers and simulation is represented by the background in the revised paper.

• Fig. 9: Colourbar scale is not readable.

Reply:

Figure 9 has been resized to make the color bar scale more readable.

• Fig. 9: What does the "(MODIS)" below the AOD, CDNC, CWP and COT annotation mean?

Reply:

It means MODIS-derived satellite data. However, the (MODIS) has been removed as there is already a header for the MODIS plots, with the exception of CWP, as the CWP is further derived by Bennartz (2007) from MODIS data.

Literatures cited in this reply:

Valari, M. and Menut, L.: Does an Increase in Air Quality Models' Resolution Bring Surface Ozone Concentrations Closer to Reality?, JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY, 25, 1955–1968, doi:10.1175/2008JTECHA1123.1, 2008. Interactive comment on Geosci. Model Dev. Discuss., 8, 6707, 2015.
C2031

Berg, L.K., Shrivastava, M., Easter, R.C., Fast, J.D., Chapman, E.G., Liu, Y., Ferrare, R.A.: A new WRF-Chem treatment for studying regional-scale impacts of cloud processes on aerosol and trace gases in parameterized cumuli, Geosci. Mod. Dev., 8, 409 – 429, doi:10.5194/gmd-8-409-2015.

Chen, D., Li, Q., Stutz, J., Mao, Y., Zhang, L., Pikelnaya, O., Tsai, J.Y., Haman, C., Lefer, B., Rappengluck, B., Alvarez, S.L., Neuman, J.A., Flynn, J., Roberts, J.M., Nowak, J.B., de Gouw, J., Holloway, J., Wagner, N.L., Veres, P., Brown, S.S., Ryerson, T.B., Warneke, C., Pollack, I.B.: WRF-Chem simulation of NOx and O3 in the L.A. basin during CalNex-2010, Atmos. Environ., 81, 421 – 432, 2013.

Forkel, R., Balzarini, A., Baro, R., Bianconi, R., Curci, G., Jimenez-Guerrero, P., Hirtl, M., Honzak, L., Lorenz, C., Im, U., Perez, J.L., Pirovano, G., San Jose, R., Tuccella, P., Werhahn, J., Zabkar, R.: Analysis of the WRF-Chem contributions to AQMEII phase 2 with respect to aerosol radiative feedbacks on meteorology and pollutant distributions, Atmos. Environ., 115, 630 – 645, 2015, doi:10.1016/j.atmosenv.2014.10.056.

Glotfelty, T., Zhang, Y., Karamchandani, P., Streets, D.G.: Will the role of intercontinental transport change in a changing climate? Atmos. Chem. Phys., 14, 9379 – 9402, 2014. Doi:10.5194/acp-14-9379-2014.

Penrod, A., Zhang, Y., Wang, K., Wu, S.-Y., Leung, R.: Impacts of future climate and emission changes on U.S. air quality, Atmos. Environ., 89, 533 – 547, 2014. Doi:10.1016/j.atmosenv.2014.01.001.

Ritter, M., Muller, M.D., Jorba, O., Parlow, E., Sally Liu, L.-J.: Impact of chemical and meteorological boundary and initial conditions on air quality modeling: WRF-Chem sensitivity evaluation for a European domain, Meteorol. Atmos. Phys., 119, 1, 59 -70, 2013.

1   **Decadal Evaluation of Regional Climate, Air Quality, and Their Interactions over the**

2   **Ceontinental U.S. using WRF/Chem Version 3.6.1**

3   Khairunnisa Yahya, Kai Wang, Patrick Campbell, Timothy Glotfelty, Jian He, and Yang Zhang*

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

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

6   Corresponding author: Yang Zhang, Email: yang_zhang@ncsu.edu

7

8   **ABSTRACT**

9   The Weather Research and Forecasting model with Chemistry (WRF/Chem) v3.6.1 with

10   the Carbon Bond 2005 (CB05) gas-phase mechanism is evaluated for its first decadal application

11   during 2001 - 2010 using the Representative Concentration Pathway (RCP 8.5) emissions to assess

12   its capability and appropriateness for long-term climatological simulations. The initial and

13   boundary conditions are downscaled from the modified Community Earth System

14   Model/Community Atmosphere Model (CESM/CAM5) v1.2.2. The meteorological initial and

15   boundary conditions are bias-corrected using the National Center for Environmental Protection's

16   Final (FNL) Operational Global Analysis data. Climatological evaluations are carried out for

17   meteorological, chemical, and aerosol-cloud-radiation variables against data from surface

18   networks and satellite retrievals. The model performs very well for the 2-m temperature (T2) for

19   the 10-year period with only a small cold bias of -0.3 °C. Biases in other meteorological variables

20   including relative humidity at 2-m, wind speed at 10-m, and precipitation tend to be site- and

21   season-specific; however, with the exception of T2, consistent annual biases exist for most of the

22   years from 2001 to 2010. Ozone mixing ratios are slightly overpredicted at both urban and rural

23   locations with an NMB of 9.7% but underpredicted at rural locations with an NMB of -8.8%. $PM_{2.5}$

concentrations are ~~slightly~~ moderately overpredicted with an NMB of 23.3% at rural sites ~~with an NMB of 23.3%~~, but slightly underpredicted with an NMB of -10.8% at urban/suburban sites ~~with an NMB of -10.8%~~. In general, the model performs relatively well for chemical and meteorological variables, and not as well for aerosol-cloud-radiation variables. Cloud-aerosol variables including aerosol optical depth, cloud water path, cloud optical thickness, and cloud droplet number concentration are generally underpredicted on average across the continental U.S. Overpredictions of several cloud variables over eastern U.S. result in underpredictions of radiation variables (such as GSW with an MB of -5.7 W m^{-2}) and overpredictions of shortwave and longwave cloud forcing (MBs of ~7 to 8 W m^{-2}) which are important climate variables. While the current performance is deemed to be acceptable, improvements to the bias-correction method for CESM downscaling and the model parameterizations of cloud dynamics and thermodynamics, as well as aerosol-cloud interactions can potentially improve model performance for long-term climate simulations.

KEYWORDS: Online-Coupled WRF/Chem; Climate, Air Quality, the Representative Concentration Pathway Scenarios, Climatological Evaluation; Chemistry-Climate Interactions

1. Introduction

Regional atmospheric models have been developed and applied for high resolution climate, meteorology, and air quality modeling in the past few decades. Comparing to global models with a coarser domain resolution (Leung et al., 2003) those regional models have advantages over global models because they can more accurately represent mesoscale variability (Feser et al., 2011), and also better predict the local variability of concentrations of specific species such as black carbon and sulfate (Petikainen et al., 2012). General circulation models (GCMs) and global chemical transport models (GCTMs) are usually downscaled to regional meteorological models such as the Weather Research and Forecasting model (WRF) (Caldwell et al., 2009; Gao et al.,

47   2012), regional climate models such as REMO-HAM (Petikainen et al., 2012), the regional

48   modeling system known as Providing Regional Climates for Impacts Studies (PRECIS) (Jones et

49   al., 2004; Fan et al., 2014), and a number of European models described in Jacob et al. (2007), as

50   well as regional CTMs such as the Community Multiscale Air Quality Model (CMAQ) (Penrod et

51   al., 2014; Xing et al., 2015). These regional models are used for climate/meteorology or air quality

52   simulations. Some are applied for more than ten years (Caldwell et al., 2009; Warrach-Sagi et al.,

53   2013; Xing et al., 2015). However these regional models either lack the detailed treatment of

54   chemistry (e.g., in WRF), or use prescribed chemical concentrations (e.g., REMO-HAM uses

55   monthly mean oxidant fields for several chemical species), or do not have online-coupled

56   meteorology and chemistry (e.g., in CMAQ). In addition, the past regional model simulations and

57   analyses have mainly focused on meteorological parameters such as surface temperature and

58   precipitation, cloud variables such as net radiative cloud forcing, and chemical constituents such

59   as ozone. Regional climate model simulations tend to focus on significant climatic events such as

60   extreme temperatures (very cold or very hot) (Dasari et al., 2014), heat waves, heavy precipitation,

61   drought, and storms (Beniston et al., 2007), rather than the important air quality and climate

62   interactions. In addition, the impacts of complex chemistry-aerosol-cloud-radiation-climate

63   feedbacks on future climate change remain uncertain, and these feedbacks are most accurately

64   represented using online-coupled meteorology and chemistry models (Zhang, 2010; IPCC, 2013).

65   An online-coupled meteorology and chemistry model, however, is more computationally

66   expensive compared to an offline-coupled model (Grell et al., 2004), and thus requires significant

67   computing resources for their long-term (a decade or longer) applications. With rapid increases in

68   the availability of high performance computing resources on the petaflop scale, however, long

69   term simulations using online-coupled models have become possible in recent years. For example,

70　recently, the WRF model has been ~~online~~ coupled online to the CMAQ model with the inclusion

71　of aerosol indirect effects ~~which is useful for the~~to study ~~of air quality~~chemistry and climate ~~change~~

72　~~mitigation~~interactions (Yu et al., 2014).

73　　　　The online-coupled WRF model with Chemistry (WRF/Chem) has been updated with a

74　suite of physical parameterizations from the Community Atmosphere Model version 5 (CAM5)

75　(Neale et al., 2010) so that the physics in the global CAM5 model is consistent with the regional

76　model for downscaling purposes (Ma et al., 2014). There are also limited applications of dynamical

77　downscaling (Gao et al., 2013) under the new Intergovernmental Panel on Climate Change (IPCC)

78　Fifth Assessment Report's Representative Concentration Pathway (RCP) scenarios (van Vuuren

79　et al., 2011). Gao et al. (2013) applied dynamic downscaling to link the global-climate-chemistry

80　model CAM-Chem with WRF and CMAQ using RCP 8.5 and RCP 4.5 emissions to study the

81　impacts of climate change and emissions on ozone (O_3). Molders et al. (2014) downscaled the

82　Community Earth System Model (CESM) (Hurrell et al., 2013) to drive the online-coupled

83　WRF/Chem model over Southeast Alaska using RCP 4.5 emissions; however, their study did not

84　address the feedback processes between chemistry and meteorology. This study evaluates the

85　online-coupled regional WRF/Chem model, which takes into account gas and aerosol-phase

86　chemistry, as well as aerosol direct and indirect effects. WRF/Chem is used to simulate the

87　"current" climate scenario for 10 years, from 2001 to 2010 using the RCP 8.5 emissions and

88　boundary conditions from an updated version of CESM with advanced chemistry and aerosol

89　treatments over continental U.S. (CONUS) (He at al., 2015; Glotfelty et al., 2015) with a focus on

90　air-quality and climate interactions. Both CESM and WRF/Chem include similar gas-phase

91　chemistry and aerosol treatments. To our best knowledge, this study is the first to report the

92　WRF/Chem simulation, evaluation, and analyses over a period of 10 years (i.e., 2001-2010) to

93 assess if the model is able to accurately simulate decadal long air quality and climatology by taking

94 into account feedback processes between chemistry and meteorology. This study also assesses

95 whether the RCP8.5 emissions for the 10-year period are robust enough to produce satisfactory

96 performance against observations with WRF/Chem.

97 **2. Model Set-up and Evaluation Protocol**

98 **2.1 Model Configurations and Simulation Design**

99   The model used is the modified WRF/Chem v3.6.1 with updates similar to those

100 implemented into WRF/Chem v3.4.1 as documented in Wang et al. (2014). The main updates

101 include the implementation of an extended version of Carbon Bond 2005 (CB05) (Yarwood et al.,

102 2005) gas-phase mechanism with the chlorine chemistry (Sarwar et al., 2007) and its coupling with

103 the Modal for Aerosol Dynamics in Europe/Volatility Basis Set (MADE/VBS) (Ahmadov et al.,

104 2012)~~The main updates in the default WRF/Chem v3.6.1 include the implementation of an~~

105 ~~extended Carbon Bond 2005 (CB05) of Yarwood et al. (2005) gas-phase mechanism~~, as well as

106 ~~with chlorine chemistry of Sarwar et al. (2007).~~, The gas-phase mechanism, which is then coupled

107 with the Modal for Aerosol Dynamics in Europe / Volatility Basis Set (MADE/VBS) (Ahmadov

108 et al., 2012). MADE/VBS incorporates a modal aerosol size distribution, and includes an

109 advanced secondary organic aerosol (SOA) treatment based on gas-particle partitioning and gas-

110 phase oxidation in volatility bins. The CB05-MADE/VBS option has also been coupled to existing

111 model treatments of various feedback processes such as the aerosol semi-direct effect on photolysis

112 rates of major gases, and the aerosol indirect effect on cloud droplet number concentration (CDNC)

113 and resulting impacts on shortwave radiation. The main physics and chemistry options used in this

114 study as well as their corresponding references can be found in Table 1. The simulations are

115 performed at a horizontal resolution of 36-km with 148×112 horizontal grid cells over the

116   CONUS domain and parts of Canada and Mexico, and a vertical resolution of 34 layers from the

117   surface to 100-~~mb~~hPa. Considering the decadal applications of WRF/Chem in this work which is

118   much longer than many past WRF/Chem applications, the simulations are reinitialized monthly

119   (rather than 1-4 days used in most past WRF/Chem applications to short-term episodes that are on

120   an order of months up to 1-year, e.g., Zhang et al., 2012a, b; Yahya et al., 2014, 2015b) to constrain

121   meteorological fields toward National Centers for Environmental Prediction (NCEP) reanalysis

122   data while allowing chemistry-meteorology feedbacks within the system. As discussed in Sections

123   3.1 and 3.3, the reinitialization frequency of 1-month may be too large to constrain some of the

124   meteorological fields such as moistures, which in turn affect other parameters, and a more frequent

125   reinitialization may be needed to improve the model performance. The impact of the frequency of

126   the reinitialization on simulated meteorological and cloud parameters will be further discussed in

127   Sections 3.1 and 3.,2. A list of acronyms used in this paper can be found in Table S~3~1.

128   **2.2 Processing of Emissions and Initial Conditions (ICs)/Boundary Conditions (BCs)**

129   Global RCP emissions are available as monthly-average emissions for 2000, 2005, and for

130   every 10 years between 2010 and 2100, at a grid resolution of $0.5^\circ \times 0.5^\circ$ (Moss et al., 2010; van

131   Vuuren et al., 2011). The RCP emissions in 2000, 2005, and 2010 are used to cover the 10-year

132   emissions needed for WRF/Chem simulations, i.e., the periods of $2001 - 2003$, $2004 - 2006$, and

133    $2007 - 2010$, respectively. Processing global RCP emissions in 2000, 2005, and 2010 into regional,

134   hourly emissions needed for the 10-year WRF/Chem simulations requires essentially three main

135   tasks. These include 1) mapping the RCP species to CB05 speciation used in WRF/Chem; 2) re-

136   gridding the RCP emissions from $0.5 \times 0.5^\circ$ grid resolution to the 36×36 km grid resolution used

137   for regional simulation over North America; and 3) applying species and location dependent

138   temporal allocations (i.e., emissions variation over time) to the re-gridded RCP emissions. Table

139   ~~S1~~ S2 shows the species mapping between RCP species and CB05 species. To map the RCP

140   species to CB05 speciation, some assumptions are made due the relatively detailed speciation

141   required by CB05. Some of the CB05 species are directly available in RCP; however, others are

142   lumped into RCP groups, for example, the "other alkanals" and "hexanes and higher alkanes" in

143   the RCP groups can be considered to approximately represent the acetaldehyde and higher

144   aldehydes emissions required by CB05, respectively (Table ~~S1~~S2). For the CB05 species such as

145   ethanol, methanol, internal and terminal olefin carbon bonds in the gas-phase, and elemental and

146   organic carbon in the accumulation mode of the aerosol particles, other RCP groups are used to

147   approximate these emissions (Table ~~S1~~S2). For the remaining CB05 species that are not available

148   in RCP, the 2000 emissions are based on the 2002 National Emission Inventory (NEI) (version 3,

149   http://www.epa.gov/ttn/chief/emch/), while the 2005 and 2010 emissions are based on the 2008

150   NEI (version 2), with year-specific updates for on/off road transport, wildfires and prescribed fires,

151   and Continuous Emission Monitoring-equipped point sources (Pouliot et al., 2014). To re-grid the

152   RCP emissions, the RCP rectilinear grid is first interpolated to a WRF/Chem curvilinear grid using

153   a simple inverse distance weighting (NCAR Command Language Function – rgrid2rcm), and a

154   subset of the RCP grid that covers the WRF/Chem CONUS domain is then extracted. To derive a

155   temporal allocation for monthly-averaged RCP emissions, hourly emissions profiles are taken

156   from in-house WRF/Chem simulations over CONUS during 2001 (Yahya et al., 2015a), and 2006

157   and 2010 (Yahya et al., 2014, 2015b). For those existing in-house simulations, the emissions were

158   generated with the Sparse Matrix Operator Kernel Emissions (SMOKE) model version 2.3 for

159   2002 NEI and SMOKE version 3.4 for 2008 NEI with year-specific sector emissions for 2006 and

160   2010, which prepare the spatially, temporally, and chemically speciated "model-ready" emissions.

161   Since NEI is updated and released every three years, the temporal profiles of emissions used in

162　SMOKE for 2002, 2006 and 2010 are assumed to be valid for 3-4 years around the NEI years, i.e.,

163　2001-2003, 2004-2006, and 2007-2010, respectively. The temporal allocations applied to the RCP

164　emissions are therefore based on the SMOKE model's profiles for each species and source

165　location, and include non-steady-state emissions rates (i.e., seasonal, weekday or weekend, and

166　diurnal variability) that are valid for the entire simulation periods of 2001-2010. Specifically, the

167　hourly re-gridded RCP emission rates for each species E, or E_{hr}^{RCP} are calculated by

168
$$E_{hr}^{RCP}(t,z,\text{lat},\text{lon}) = \text{E}_{mon}^{RCP}(z,lat,lon)*\left[\frac{E_{hr}^{WRF}(t,z,\text{lat},\text{lon})}{E_{mon}^{WRF}(z,\text{lat},\text{lon})}\right] \quad (1)$$

169　where E_{mon}^{RCP}, E_{mon}^{WRF}, and E_{hr}^{WRF} represent the original monthly-averaged RCP emissions rates, the

170　monthly-averaged WRF/Chem emissions rates, and the hourly WRF/Chem emission rates,

171　respectively, which are valid at each model time t, layer z, and lat and lon grid points. The RCP

172　elevated source emissions for sulfur dioxide (SO_2), sulfate (SO_4^{2-}), elemental carbon (EC) and

173　organic carbon (OC) were also incorporated into the model-ready emissions for WRF/Chem using

174　steps 1) − 3) and Eq. (1) above. Lastly, RCP aircraft source emissions for EC, nitric oxide (NO),

175　and nitrogen dioxide (NO_2) are directly injected into the closest model layers. No temporal

176　allocations are applied to the RCP aircraft source emissions.

177　　　Biogenic emissions are calculated online using the Model of Emissions of Gases and

178　Aerosols from Nature version 2 (MEGAN2) (Guenther et al., 2006). Emissions from dust are based

179　on the online Atmospheric and Environmental Research Inc. and Air Force Weather Agency

180　(AER/AFWA) scheme (Jones and Creighton, 2011). Emissions from sea salt are generated based

181　on the scheme of Gong et al. (1997).

182　　　The chemical and meteorological ICs/BCs come from the modified CESM/CAM5 version

183　1.2.2 with updates by He et al. (2014) and Glotfelty et al. (2015) developed at the North Carolina

184 State University (CESM_NCSU). WRF/Chem and CESM both use the CB05 gas-phase

185 mechanism (Yarwood et al., 2005), however, WRF/Chem includes additional chlorine chemistry

186 from Sarwar et al. (2007), wh~~ile~~ereas CESM_NCSU uses a modified version of CB05, the CB05

187 Global Extension (CB05GE) by Karamchandani et al. (2012). In addition to original reactions in

188 CB05 and chlorine chemistry of Sarwar et al. (2007), CB05GE~~which~~ includes chemistry on the

189 lower stratosphere, reactions involving mercury species, and additional heterogeneous reactions

190 on aerosol particles, cloud droplets and on polar stratospheric clouds (PSCs). Both WRF/Chem

191 and CESM_NCSU ~~also~~ use a modal aerosol size representation~~treatments~~, rather than a sectional~~.~~

192 size representation. While WRF/Chem includes MADE/VBS with 3 prognostic modes (Ahmadov

193 et al., 2012)~~, is used in WRF/Chem~~ CESM_NCSU includes~~while~~ the Modal Aerosol Model with

194 ~~a~~ 7 prognostic ~~-~~modes ~~prognostic Modal Aerosol Model (MAM7)~~ (Liu et al., 2012) is used in

195 CESM_NCSU. In addition to similar gas-phase chemistry and aerosol treatments, CESM_NCSU

196 and WRF/Chem use the same shortwave and longwave radiation schemes (i.e., the Rapid and

197 accurate Radiative Transfer Model for GCM (RRTMG)), though they use different cloud

198 microphysics parameterizations, PBL, and convection schemes. As GCMs generally contain

199 systematic biases which can influence the downscaled simulation, the meteorological ICs/BCs

200 predicted by CESM_NCSU are bias corrected before they are used by WRF/Chem using the

201 simple bias correction technique based on Xu and Yang (2012). Temperature, water vapor,

202 geopotential height, wind, and soil moisture variables available every 6 hours from the NCEP Final

203 Reanalyses (NCEP FNL) dataset~~, which is available every 6 hours,~~ are used to correct the ICs and

204 BCs derived based on results from CESM_NCSU for WRF/Chem simulations ~~generated by~~

205 ~~CESM_NCSU~~. In this bias-correction approach, ~~M~~monthly climatological averages for ICs and

206 BCs are first derived from both NCEP and CESM_NCSU cases. The differences between the

207   NCEP FNL data and the CESM_NCSU derived ICs/BCs are obtained from 2001 to 2010, and are

208   then averaged to produce 10-year average differences of 6-hourly meteorological ICs/BCs. The

209   10-yr average differences of 6-hourly meteorological ICs/BCs are used to correct CESM_NCSU

210   meteorological ICs/BCs for each of the 10 yearsICs and BCs from the NCEP and CESM_NCSU

211   climatological averages are then added onto the CESM_NCSU ICs and BCs to generate new bias-

212   corrected CESM_NCSU ICs/BCs. This bias correction technique can also be applied to future year

213   simulations where NCEP FNL data is not available.

214   **2.3 Model Evaluation Protocol**

215   The focus of the model evaluation is mainly to assess whether the model is able to

216   adequately reproduce the spatial and temporal distributions of key meteorological and chemical

217   variables as compared to observations on a climatological time scale. A scientific question to be

218   addressed in this work is, is WRF/Chem sufficiently good for regional climate and air quality

219   simulations on a decadal scale? A climatological month refers to the average of the month for all

220   the 10 years. For example, January refers to the average for all the months of January from 2001

221   to 2010. Statistical evaluations such as mean bias (MB), Pearson's correlation coefficient

222   (RCorrR), normalized mean bias (NMB), normalized mean error (NME) (The definition of those

223   measures can be found in Yu et al. (,2006) and Zhang et al. (2006)) and Index of Agreement (IOA)

224   ranging from 0 to 1 (Willmott et al., 1981) for major chemical and meteorological variables are

225   included. For surface networks with hourly data, e.g. NCDC, the observational data areis paired

226   up with the simulated data on an hourly basis for each site. The observational data and simulated

227   data are averaged out for each site respectively. The statistics are then calculated based on the site-

228   specific data pairs. The satellite-derived data areis usually available on a monthly basis, and the

229   simulated data areis also averaged out on a monthly basis. The satellite-derived data areis regridded

230   to the same domain and number of grid cells similar to the simulated data. The time dimension is

231   removed for the climatological evaluation, the statistics are based on a site-specific average or a

232   grid cell average. The statistics are then calculated based on the paired satellite-derived vs.

233   simulated grid cell values. The spatial and temporal analyses include spatial plots of MB over

234   CONUS, spatial overlay plots of averaged simulated and observational data, monthly

235   climatologically-averaged time series of major meteorological and chemical variables, annual

236   average time series; probability distributions of major meteorological and chemical variables, and

237   spatial plots of major aerosol and cloud variables compared with satellite data. A summary of the

238   observational data from surface networks and satellite retrievals can be found in Table ~~S2~~S3. The

239   variables that are analyzed in this study include O_3, particulate matter with diameter less than and

240   equal to 2.5 and 10 μm ($PM_{2.5}$ and PM_{10}, respectively), and $PM_{2.5}$ species including sulfate (SO_4^{2-}

241   ), ammonium (NH_4^+), nitrate (NO_3^-), EC, OC, and total carbon (TC = EC + OC), temperature at 2-

242   m (T2), relative humidity at 2-m (RH2), and wind speed at 10-m (WS10), wind direction at 10-m

243   (WD10), precipitation, aerosol optical depth (AOD), cloud fraction (CLDFRA), cloud water path

244   (CWP), cloud optical thickness (COT), CDNC, cloud condensation nuclei (CCN), downward

245   shortwave radiation (SWDOWN), net shortwave radiation (GSW), downward longwave radiation

246   (GLW), outgoing longwave radiation at the top of atmosphere (OLR), and shortwave and

247   longwave cloud forcing (SWCF and LWCF). While uncertainties exit in all the observational data

248   used, ~~systemetic~~ systematic uncertain~~ty~~ analysis/quantification is beyond the scope of this work.

249   In this work, all observational data are considered to be the true values in calculating the

250   performance statistics. The information on the accuracy of most data used in the model evaluation

251   has been provided in Table 2 of Zhang et al. (2012a). Uncertainties associated with some of the

252   observational data are discussed in Section 3.

253 **3. Model Performance Evaluation**

254 **3.1 Meteorological Predictions**

255   Table 2 summarizes the statistics for T2, RH2, WS10, WD10, and precipitation. The model

256 performs very well for a 10-year average T2 with a slight underprediction (an MB of -0.3 °C).

257 This is better or consistent with other studies which tend to report underpredictions in simulated

258 T2. Brunner et al. (2014) reported a range of monthly MBs for T2 of -2 to 1 °C for simulations

259 using a number of CTMs over individual years for 2006 and 2010 with reanalysis meteorological

260 ICs/BCs. Seasonal temperature biases of -1.8 to -2.3 °C were reported from an ensemble of

261 regional climate models (RCMs) for a simulation period of 1971 to 2000 over the Northeast

262 (Rawlins et al., 2012). He et al. (2015) also showed biases of ~~0 to~~ -3 <u>to 0</u> °C over CONUS when

263 compared against NCEP reanalysis data. Kim et al. (2013) compared the results of a number of

264 RCMs over CONUS over a climatological period of 1980 to 2003 against Climatic Research Unit

265 (CRU) surface analysis data at a 0.5º resolution and reported T2 biases of -5 to 5 °C. Figure 9.2

266 from Flato et al. (2013) shows that the Coupled Model Intercomparison Project Phase 5 (CMIP5)

267 models tend to underpredict T2 for the period of 1980 to 2005 over western U.S. by up to -3 °C.

268 The slight bias in T2 can be attributed to errors in soil temperature and soil moisture (Pleim and

269 Gilliam, 2009) or errors in the green vegetation fraction in the National Center for Environmental

270 Prediction, Oregon State University, Air Force and Hydrologic Research Lab (NOAH) Land

271 Surface Model (LSM) (Refslund et al., 2013). RH2 and WS10 are slightly overpredicted.

272 Precipitation is largely overpredicted, consistent with overpredictions in precipitation from WRF

273 and WRF/Chem simulations reported in literatures. For example, Caldwell et al. (2009) attributed

274 the overprediction in precipitation to overprediction in precipitation intensity but underprediction

275 in precipitation frequency. Otte et al. (2012) also reported that the precipitation predicted by WRF

12

276   is too high compared to the North American Regional Reanalyses (NARR) data throughout the

277   whole CONUS domain over a period of 1988 – 2007. ~~Nasrollahi et al. (2012) examined 20~~

278   ~~combinations of microphysics and cumulus parameterization schemes available in WRF and found~~

279   ~~that most parameterization schemes overestimate the amount of rainfall and the extent of high~~

280   ~~rainfall values. In this study, while Grell 3D Ensemble cumulus parameterization contributes in~~

281   ~~part to the overpredictions of precipitation, most overpredictions occur at high thresholds as shown~~

282   ~~in Figure 3 (d) and they are attributed to possible errors in the Morrison two-moment scheme~~

283   ~~because the overpredictions of non-convective precipitation dominate the overpredictions of total~~

284   ~~precipitation.~~ Nudging and reinitialization have been most commonly used methods to control

285   such errors. ~~In this case, the precipitation bias can be attributed to the use of bias corrected~~

286   ~~CESM_NCSU data instead of NCEP reanalysis data, as well as the choice of cumulus~~

287   ~~parameterization scheme in this study (Grell 3D). A number of~~Three sensitivity simulations ~~has~~

288   ~~been~~are conducted for a summer month (July 2005) to ~~pinpoint~~analyze likely causes of the

289   precipitation biases. The baseline simulation (**Base**) uses a monthly reinitialization frequency,

290   CESM_NCSU ICs/BCs, and the Grell 3D cumulus parameterization. The sensitivity simulations

291   include (1) **Sen1**, which is similar to the Base case except with a 5-day reinitialization period; (2)

292   **Sen2**, which is similar to Base except using NCEP for the meteorological ICs/BCs; and (3) **Sen3**,

293   which is similar to Base except using WRF/Chem v3.7 with the Multi-Scale Kain Fritsch (MSKF)

294   cumulus parameterization, instead of Grell 3D. The differences in configuration setup in those

295   sensitivity simulations are given in Table S4. The evaluation and comparison of the baseline and

296   sensitivity results in July 2005 ~~as shown by~~are summarized in Tables S4~~5~~ ~~to~~and S6, and Figure S1

297   in the ~~S~~supplementary material. As shown in Tables S5-S6 and Figure S1, the precipitation bias

298   can be attributed to several factors including the use of Grell 3D cumulus parameterization scheme,

the use of bias-corrected CESM_NCSU data (instead of NCEP reanalysis data), and the use of an

reinitialization frequency of 1-month, among which the first factor dominates the biases in

precipitation predictions. The simulated precipitation is very sensitivity to different cumulus

parameterizations. Compared to scale-aware parameterizations such as the multiple-scale Kain-

Fritsch (MSKF) cumulus scheme, the Grell 3D parameterization has a tendency to overpredict

precipitation, particularly over ocean.inusing the latest version of WRF/Chem (version 3.7.1) that

contains MSKF (which is not available in WRF/Chem version 3.6.1). This work uses

reinitialization but the frequency of reinitialization is monthly rather than every 1-4 days used in

other studies (e.g., Zhang et al., 2012a,b; Yahya et al. 2014, 2015b) , which led to a buildup of

storm systems, especially over the warm Atlantic. This buildup in turn influences mainly non-

convective precipitation over land, especially in the east coast. Simulations with a more frequent

reinitialization tend to perform better than those with less frequent reinitialization (Lo et al., 2008).

Figure 1 shows the spatial distributions of MB for 10-year average predictions of T2, RH2,

WS10, and precipitation. Figure 2 shows the time series of 10-year average monthly and annual

average T2, WS10, RH2, precipitation, O_3, and $PM_{2.5}$ against observational data and IOA statistics.

T2 (Figure 1a) tends to be underpredicted over eastern and western U.S. and overpredicted over

the central U.S. The bias correction method itself may also contribute to the slight biases in T2. A

single temporally averaged (2001 – 2010) NCEP reanalysis file is applied to the 6-hourly BCs for

each individual year, which would in some cases contribute to the biases in the climatological 10-

year evaluation. T2 also tends to be underpredicted during the cooler months but overpredicted

during the warmer months (Figure 2a). While the bar charts in Figure 2 show domain- average

mean observed and mean simulated T2, IOA performance takes into account the proportion of

14

321 differences between mean observed and mean simulated values at different sites. IOA can be

322 calculated as,

$$IOA = 1 - \frac{\sum_{i}^{N}(O_i - S_i)^2}{\sum_{i}^{N}(|O_i - \overline{O}| + |S_i - \overline{S}|)^2} \qquad (2)$$

324 where O_i and S_i denote time-dependent observations and predictions at time and location i,

325 respectively, N is the number of samples (by time and/or location), $\overline{O}$ denotes mean observation

326 and $\overline{S}$ denotes mean predictions over all time and locations, they can be calculated as:

$$\overline{O} = (1/N)\sum_{i=1}^{N} O_i , \ \overline{S} = (1/N)\sum_{i=1}^{N} S_i ,$$

328 IOA values range from 0-1, with a value of 1 indicating a perfect agreement.

329 The model performance in terms of IOA for T2 is slightly worse during the warmer months

330 as compared to the cooler months; however, IOA values for all months are ≥ 0.9. The poorer IOA

331 statistics for the warmer months are possibly influenced to a certain extent by the fact that the IOA

332 tends to be more sensitive towards extreme values (when temperatures are maximum) due to the

333 squared differences used in calculating IOA (Legates and McCabe, 1999). As shown in Figures 1b

334 and 2b, the spatial distributions of MBs for RH2 follow closely the spatial distributions of MBs

335 for T2, where T2 is underpredicted, RH2 is overpredicted and vice versa. Unlike T2, the IOA for

336 RH2 is the highest during the warmer months and the lowest during the winter months, but IOA

337 for RH2 is generally high (> 0.7) for all months. WS10 is also generally overpredicted along the

338 coast, over eastern U.S. and some portions over the western U.S. (Figure 1c), consistent with

339 overpredictions of T2 over the coast, and partially due to unresolved topographical features. In this

340 case the topographic correction for surface winds used to represent extra drag from sub-grid

341 topography (Jimenez and Dudhia, 2012) is used as an option in the 10-yr WRF/Chem simulations;

342   however, WS10 is still overpredicted except for the areas of flat undulating land in the central U.S.

343   Jimenez and Dudhia (2012) also suggested that the grid points nearest to the observational data

344   might not be the most appropriate or most representative, and that the selection of nearby grid

345   points can help to reduce errors in surface wind speed estimations. In this study, as the evaluation

346   is conducted over the whole CONUS, the nearest grid points are used for evaluation, which could

347   also result in errors in wind speed evaluation. The positive T2 and WS10 bias along the coast could

348   be due to the fact that the model grids for temperatures and wind speeds are located over the ocean,

349   however, the observation points are located slightly inland. As shown in Figure 2, WS10 performs

350   well on average for the months of April, May, and June, and is overpredicted for the other months.

351   Nonetheless the climatological NMB for WS10 overall is low at 7.7% (Table 2). WS10 has higher

352   IOA values during the spring months and the lowest IOA during the summer months and in

353   November. The model performs relatively well in predicting WD10 variability with a Corr of 0.6,

354   indicating overall a more southerly direction domain-wide predicted by the model compared to

355   observations. Precipitation is overpredicted for all months except for June, especially during the

356   summer months of July to August. Even with the inclusion of radiative feedback effects from the

357   subgrid-scale clouds in the radiation calculations, precipitation is still overpredicted with the Grell

358   3D scheme, which is consistent with the results shown by Alapaty et al. (2012).   Precipitation

359   mainly has lower IOAs during the summer compared to other months, except in June which

360   actually exhibits the largest IOA of all months. Even though June is considered a summer month,

361   it does not show overprediction in precipitation compared to the other summer months. It is

362   possible that in June, the overall atmospheric moisture content is low. This is consistent with

363   simulated RH2 as June is the only month where RH2 is underpredicted compared to observations.

364   In general the model is able to reproduce the monthly trends in meteorological variables;

365   for example, the predicted trend in T2 closely follows the observed trends by the National Climatic

366   Data Center (NCDC)-. The observed RH2 decreases from January to a minimum in April, and

367   then increases from April to December. Although the model predicts a similar pattern in RH2,

368   there is a lag in the RH2 minimum occurring two months later in June (Figure 2b). For WS10, the

369   observation peaks in April, as compared to the simulated peak in March. The model correctly

370   predicts the observed WS10 minimum occurring in August. The model trend in precipitation is

371   similar to observations, except during the summer months of July through September, where a

372   large overprediction leads to a sharp increase in July, followed by a gradual decrease through

373   December.

374   Figures 2e – 2h show the annual time series trends for T2, RH2, WS10, and precipitation.

375   The model performs relatively well in predicting the annual mean T2 for most years (with MBs of

376   < 0.5 °C; Figure 2e). T2 also does not show an obvious decreasing or increasing T2 trend between

377   2001 and 2010. The IOA for annual T2 for all years are > 0.95. However for 2002, mean simulated

378   T2 is ~0.7 °C higher than the observational data. IOA is still high for 2002 which indicates

379   probably good performance of T2 at most sites, however with large overpredictions at a few sites

380   which could skew the mean observed and mean simulated value but not influence IOA

381   significantly. RH2 is consistently overpredicted by the model with the largest overprediction in

382   2009. With the exception of 2009, observed RH2 is rather steady (65 – 70 %) from 2001 to 2010.

383   IOA is also steady for RH2, except for 2009. As mentioned earlier, WRF tends to overpredict

384   WS10 in general. Figure 2g shows that observations indicate weaker wind speeds from 2001 to

385   2007. Model performance is better from 2007 to 2010 with higher IOAs compared to previous

386   years. WRF has worse performance especially at weaker wind speeds as is the case from 2001 to

387 2007. Model performance for precipitation is more variable year-to-year, with IOAs ranging from

388 0.4 to 0.7; however, there is a systematic positive bias during the 10 year period.

389 Figure 3 shows the probability distributions of T2, RH2, WS10, and precipitation against

390 NCDC and NADP for 10 years. The observed and simulated variables are averaged at each site

391 for the 10-year period, and the pairs are then distributed into a probability distribution over 30 bins

392 of observed and simulated values of T2. For T2, the simulated and observed probability

393 distributions are very similar (Figure 3a), consistent with the statistics for T2 which shows only a

394 small cold bias. The model overpredicts T2 at sites where temperatures are very low. The

395 probability distribution curve for simulated RH2 is also shifted to the right of the observed RH2

396 (Figure 3b), with an observed and modeled peak 74% and 78% respectively. The probability

397 distribution of simulated WS10 is narrower (between 2 and 6 m s^{-1}) compared to that of observed

398 WS10 (between 1 and 7 m s^{-1}). The model thus overpredicts when near-surface wind speeds are

399 low, but underpredicts when wind speeds are very high. This suggests that the surface drag

400 parameterization is still insufficient to help predict low wind speeds; however, it might have

401 contributed to the reduction in the simulated high wind speeds (Mass, 2012). The probability

402 distribution for simulated precipitation against NADP also shows a shift to the right, consistent

403 with the statistics for overpredicted precipitation and also with the probability curve of RH2.

404 Nasrollahi et al. (2012) examined 20 combinations of microphysics and cumulus parameterization

405 schemes available in WRF and found that most parameterization schemes overestimate the amount

406 of rainfall and the extent of high rainfall values. In this study, while Grell 3D Ensemble cumulus

407 parameterization contributes in part to the overpredictions of precipitation, most overpredictions

408 occur at high thresholds as shown in Figure 3 (d) and they are attributed to possible errors in the

409 Morrison two moment scheme because the overpredictions of non-convective precipitation

410 dominate the overpredictions of total precipitation.

411 **3.2 Chemical Predictions**

412 **3.2.1 Ozone**

413   Table 2 summarizes the statistics for major chemical species. The model overpredicts

414 hourly O_3 mixing ratios on average against the Aerometric Information Retrieval System (AIRS)

415 − Air Quality System (AQS) with an NMB of 9.7% and an NME of 22.4%, but underpredicts O_3

416 mixing ratios against the Clean Air Status and Trends Network (CASTNET) with an NMB of -

417 8.8% and an NME of 19.8%. The O_3 mixing ratios are overpredicted at AIRS-AQS sites for all

418 climatological months except for April and May (Figure 4a) but underpredicted at CASTNET sites

419 for all months except for October with the largest underpredictions occurring in April and May

420 where IOA statistics are the lowest (Figure 4b). IOA statistics for all climatological months range

421 from 0.5 to 0.6 for AIRS-AQS and from 0.4 to 0.9 for CASTNET. In general, IOA values tend to

422 be higher for CASTNET compared to AIRS-AQS during the fall and winter months of October to

423 March. The IOA values for AIRS-AQS are rather steady on average over the 12 months compared

424 to CASTNET. This can be attributed to the larger dataset of AIRS-AQS (> 1000 stations)

425 compared to CASTNET (< 100 stations), the high and low undulations in O_3 averages at the

426 CASTNET sites tend to be smoothed or averaged out in O_3 averages at the AIRS-AQS sites given

427 larger AIRS-AQS dataset. The observed data from AIRS-AQS and CASTNET also show the

428 highest monthly O_3 mixing ratios over April and May. This result is consistent with the findings

429 of Cooper et al. (2014), who reported the highest mass of tropospheric O_3 for the northern

430 hemisphere in April and May based on the Ozone Monitoring Instrument (OMI) measurements in

431 2004, which suggested that the column mass of O_3 is not necessarily proportional to nitrogen oxide

19

432   (NO$_x$) emissions that peak during the summer. In addition, Cooper et al. (2014) attributed a shift

433   in the seasonal O$_3$ cycle observed at many rural mid-latitude monitoring sites to emissions

434   reductions in the U.S. The same study also reported that the summertime O$_3$ mixing ratios were

435   lower in eastern U.S. between 2005 and 2010 when compared to previous years, while remaining

436   relatively constant in spring. Thus the summer O$_3$ maximum during 2001- 2004 was replaced by

437   a broad spring/summer peak in 2005 - 2010. Both the observed and simulated O$_3$ mixing ratios do

438   not decrease for AIRS-AQS and CASTNET from 2001 to 2010 (Figures 4e and 4f). This is

439   somewhat consistent with Cooper et al. (2014) which showed that surface and lower tropospheric

440   O$_3$ has a decreasing trend over eastern U.S. but an increasing trend over the western U.S. from

441   1990-1999 to 2010. The predicted annual average O$_3$ mixing ratios are consistent from 2001 to

442   2010, with overpredictions and IOAs of ~0.6 at the AIRS-AQS sites, and underpredictions and

443   IOAs of ~0.6 to 0.8 at the CASTNET sites.

444   Figure 5 shows the probability distributions of maximum 1-hour and 8-hour O$_3$ mixing

445   ratios against CASTNET and AIRS-AQS. The probability distributions of the observed and

446   simulated O$_3$ mixing ratios are very similar. The model is able to simulate the range and

447   probabilities of O$_3$ mixing ratios relatively well at both CASTNET and AIRS-AQS sites. At the

448   CASTNET sites as shown in Figures 5a and b, the model accurately predicts the peak maximum

449   1-hour O$_3$ mixing ratio centered at ~60 ppb, however, slightly underpredicts the peak maximum

450   8-hour O$_3$ mixing ratio by a few ppb. At the AIRS-AQS sites as shown in Figures 5c and d, the

451   predicted probability distribution curve is slightly shifted to the right of the observations for both

452   maximum 1-hour and 8-hour O$_3$ mixing ratios. It is also interesting to note that the probability

453   distributions for CASTNET and AIRS-AQS are quite different. O$_3$ at the AIRS-AQS sites has a

454   unimodal normal distribution, while O$_3$ at the CASTNET sites has a bi-modal distribution, with a

455   tail of the distribution extending toward lower O_3 mixing ratios (0 – 20 ppb). The peak distribution

456   occurs at around 10 ppb, because the O_3 mixing rations are low at most CASTNET sites. The

457   second peak at ~60 ppb for CASTNET occurs mainly around the summer months during which

458   O_3 is produced through photochemistry involving its precursors. These distributions are attributed

459   to the nature of the sites' locations, where the AIRS-AQS network includes a mixture of urban,

460   suburban and rural sites, leading to a normal distribution of O_3 mixing ratios centered at relatively

461   higher O_3 mixing ratios, while the CASTNET network includes mostly rural sites that exhibit a

462   low maximum 1-hour and 8-hour O_3 mixing ratios, thus leading to a distribution with a tail skewed

463   towards the lower O_3 mixing ratios.

464   Figure 6 shows the diurnal variation of O_3 concentrations and IOA statistics for the four

465   climatological seasons against CASTNET (Figures a to d) and AIRS-AQS (Figures e to h) (Winter

466   - January, February and December (JFD); Spring - March, April, and May (MAM); Summer -

467   June, July, and August (JJA); Fall - September, October, and November (SON). Figure 6a shows

468   that in more rural sites (CASTNET) in winter O_3 tends to be underpredicted during the morning

469   (01:00 – 09:00 local standard time (LST)) and evening hours (18:00 – 24:00 LST). However,

470   Figure 6b shows that in general for all AIRS-AQS sites including urban sites, O_3 is systematically

471   overpredicted for all hours of the day. The diurnal trends for CASTNET and AIRS-AQS are

472   completely opposite for winter. As CASTNET sites are located in areas where urban influences

473   are minimal, most of these sites are likely to be NO_x-limited sites (Campbell et al., 2014).

474   Underpredicted NO_x emissions in rural areas can lead to underpredictions in O_3 concentrations in

475   NO_x-limited areas. As shown in Figure 2a), T2 is generally overpredicted during the winter

476   months, which explains the overpredictions in O_3 for most sites against AIRS-AQS. As shown in

477   Figures 6a, b and c, for CASTNET, the diurnal variations of O_3 in MAM and JJA are similar to

478   that in JFD. As shown in Figure 6d, slight overpredictions during the daylight hours of 10:00 to

479   17:00 LST occur in SON at the CASTNET sites, however the trends are similar for morning and

480   evening hours as compared to the other seasons. Similar to SON at the CASTNET sites, for AIRS-

481   AQS sites, overpredictions during daylight hours occur in JJA and SON (Figures 6 g and h), and

482   also to a much lesser extent in MAM (Figure 6f). This is probably due to the overpredictions of

483   T2, which are the smallest during MAM compared to other months as shown in Figure 2a.

484   Figure 7 compares the spatial distributions of 10-year average of the predicted and

485   observed hourly O_3 mixing ratios. The O_3 mixing ratios tend to be underpredicted in eastern and

486   northeastern U.S., where most of the CASTNET sites are located (Figure 7a). This is consistent

487   with the diurnal trends from Figures 6a to d which also show underpredictions for CASTNET sites.

488   From Figure 1a, T2 is underpredicted on average over northeastern U.S., which results in

489   underpredictions in biogenic emissions in the rural areas from MEGAN2. This would in turn

490   reduce O_3 mixing ratios in VOC-limited areas. O_3 photochemical reactivities would also be

491   reduced due to reduced T2. O_3 mixing ratios are, however, overpredicted over northwestern U.S.,

492   and also near the coastline of western U.S. The overprediction of O_3 mixing ratios in northwestern

493   U.S. can be attributed to an overprediction in the chemical BCs from CESM, as indicated by the

494   high O_3 mixing ratios near the northwestern region of the domain boundary.

495   **3.2.2 Particulate Matter**

496   The 10-year average $PM_{2.5}$ concentrations are overpredicted with an NMB of 23.3 %

497   against IMPROVE, and underpredicted with an NMB of -10.8 % against the Speciated Trends

498   Network (STN) (Table 2). In addition, the IOA trend in Figure 4c shows very good performance

499   for $PM_{2.5}$ against the Interagency Monitoring of Protected Visual Environments (IMPROVE) with

500   IOA values > 0.8. IOA values for $PM_{2.5}$ against STN are high (~ 0.6 − 0.8) during the spring and

501 summer months, but lower (~ 0.4) during the winter months (Figure 4d). The IMPROVE surface

502 network covers generally rural areas and national parks while the STN surface network covers

503 urban sites. The horizontal resolution of 36×36 km^2 used in this study may be too coarse to resolve

504 the locally high PM$_{2.5}$ concentrations at urban sites in STN which are in proximity of significant

505 point sources, especially during the fall and winter. During these colder seasons, PM$_{2.5}$

506 concentrations over the U.S. in general tend to be higher due to an extensive use of woodstove and

507 cold temperature inversions, which trap particulates near the ground (EPA, 2011). As shown in

508 Table 2, the concentrations of PM$_{2.5}$ species such as SO$_4^{2-}$, OC, and TC are overpredicted at the

509 IMPROVE sites, while the concentrations of the other main PM$_{2.5}$ species NO$_3^-$, NH$_4^+$, and EC are

510 underpredicted at both IMPROVE and STN sites. TC concentrations, which are the sum of OC

511 and EC, are overpredicted due to larger overpredictions of OC compared to the underpredictions

512 of EC. The model also simulates both primary organic aerosol (POA) and secondary organic

513 aerosol (SOA). OC is calculated as the sum of POA and SOA divided by the ratio of OA/OC,

514 which is assumed to be a constant of 1.4 (Aitken et al., 2008). This calculation of OC using a

515 constant of 1.4 is an approximation, which is subject to uncertainties when comparing simulated

516 OC against observational data, as the ratio of OA/OC can be different in different environments

517 (Aitken et al., 2008).

518   As shown in Table 2, at the STN sites, the model slightly overpredicts the concentrations

519 of SO$_4^{2-}$, while underpredicting those of NO$_3^-$, NH$_4^+$, and EC. The overpredictions of SO$_4^{2-}$ are

520 likely due to the uncertainties that arise from processing of the RCP SO$_2$ emissions. The RCP SO$_2$

521 emissions are only available as a total emission flux, and they are not vertically distributed to the

522 important point sources such as furnaces and stacks. In this work, two steps are taken to resolve

523 the RCP elevated SO$_2$ emissions in each emission layer. First, a set of factors are derived from the

524   fraction of the elevated emissions in each layer to the vertical sum of emissions for NEI used by

525   default in the SMOKE model with the NEI data. Second, these factors are applied to the total RCP

526   emissions to obtain SO_2 emissions in each emission layer. The total RCP SO_2 emissions were

527   higher than the total NEI emissions, resulting in higher surface and elevated SO_2 emissions.

528   Figures 4g and 4h compare the modeled annual average time series for $PM_{2.5}$ against IMPROVE

529   and STN observations, respectively. In general, the model performs well for $PM_{2.5}$ at the

530   IMPROVE (IOA > 0.8) and STN (IOA ~ 0.5 – 0.7) sites. A declining trend in $PM_{2.5}$ observed and

531   simulated concentrations are also observed over the years. For the later years (2007 to 2010), the

532   model performs significantly better against IMPROVE compared to STN. As 2010 NEI emissions

533   are used for the years 2007 to 2010, there are not many variations in the simulated $PM_{2.5}$

534   concentrations over these 4 years.

535   Figures 7 and 8 show the spatial plots of 10-yr average of simulated 24-hour average ,

536   PM_{10}, $PM_{2.5,}$ and $PM_{2.5}$ species concentrations, overlaid with observations from both STN and

537   IMPROVE. The underpredictions of PM_{10} are dominated by an underprediction in the wind-blown

538   dust emissions, especially in western U.S. (Figure 7b). This is confirmed in Table 2, which shows

539   an MB of -11.5 $\mu g\ m^{-3}$ and an NMB of -51.2% against PM_{10} observations at AIRS-AQS sites. The

540   observational data indicate the elevated concentrations of dust over portions of Arizona and

541   California (> 50 $\mu g\ m^{-3}$), which are not reproduced by the simulations (the simulated

542   concentrations are much lower, < 20 $\mu g\ m^{-3}$). The AER/AFWA dust module (Table 1) does not

543   produce sufficient dust in this case, even though WS10 is overpredicted and is proportional to the

544   dust emissions. The sea-salt emission module by Gong et al. (1997), however, seems to produce a

545   reasonable amount of sea-salt as shown by the similar concentrations between simulated and

546   observational data for PM_{10} near the coastlines. In addition, the MADE/VBS module in

547 WRF/Chem does not explicitly simulate the formation/volatilization of coarse inorganic species.

548 The coarse inorganic species are available, however, in the emissions and are transported and

549 deposited in a manner that is similar to non-reactive tracers.

550 The model performs well for $PM_{2.5}$ over eastern U.S. (Figure 7c), where modeled

551 concentrations are close to the observations; however, over the western U.S. there are

552 underpredictions in $PM_{2.5}$, especially in central to southern California. Even though Table 2 shows

553 in general an overprediction of SO_4^{2-} against STN sites, the model underpredicts SO_4^{2-} in regions

554 of elevated SO_4^{2-} concentrations, in particular, where concentrations are above 10 μg m^{-3} in the

555 vicinity of significant point sources of SO_2 and SO_4^{2-} over eastern U.S. (Figure 7d). This is likely

556 due to the coarse resolution ($0.5^o \times 0.5^o$) of RCP emissions, which probably results in a general

557 overprediction of SO_2 emissions over a grid but cannot resolve point sources smaller than the grid

558 resolution. A similar pattern is found for NH_4^+ over eastern U.S. due to underpredictions of high

559 concentrations of SO_4^{2-} (Figure 8a). There are also large underpredictions in NH_4^+ over the western

560 U.S. The underpredictions in NH_4^+ are likely due to underpredictions of NH_3 emissions from RCP.

561 The NH_3 emissions from RCP are much lower than those of NEI emissions over western U.S., by

562 more than a factor of 5, especially over portions of California. Large underpredictions occur over

563 both eastern and western U.S. for NO_3^-, EC, and TC (Figures 8b, c, and d). The underpredictions

564 in NO_3^- are more likely influenced by the underpredictions of NH_4^+ rather than NO_x emissions.

565 NO_x emissions for NEI are higher than those of RCP for a number of point sources, however, in

566 general RCP has higher NO_x emissions. Other possible reasons for the underpredictions of NO_3^-

567 concentrations include both prediction and measurement errors associated ~~from~~with SO_4^{2-} and

568 TNH_4 that can greatly affect the performance of NO_3^-, inaccuracies in the assumptions used in the

569 thermodynamic model (e.g., the assumption that inorganic ions are internally mixed and the

570 equilibrium assumption might not be representative, especially for particles with larger diameters),

571 as well as inaccuracies in T2 and RH predictions (Yu et al., 2005). The statistics for IMPROVE

572 TC indicate overpredictions; however the statistics for STN TC indicate larger underpredictions

573 with an MB of -2.0 μg m^{-3}, which would explain the large underpredictions in PM$_{2.5}$ concentrations

574 over western U.S. The large underpredictions are in part impacted by uncertainties in emissions as

575 well as due to uncertainties in the precursor gas emissions for these species, especially for TC. The

576 RCP emissions of EC and POA are lower when compared to those of NEI. NEI emissions have a

577 higher spatial resolution, and thus more adequately represent the emissions from point sources

578 compared to RCP. The underpredictions of TC are also more likely due to underpredictions in EC

579 as compared to OC, as shown in underpredictions of EC by Figure 8c. As T2 is slightly

580 underpredicted, these could have resulted in underpredictions in the isoprene and terpene, which

581 are major gas precursors of biogenic SOA, resulting in lower SOA and OC concentrations. In

582 addition, the emissions of anthropogenic VOC species from RCP which are also of a lower spatial

583 resolution compared to their emissions in the NEI tend to also be lower than NEI levels especially

584 at point sources. The underpredictions for these particulate species, especially for water-soluble

585 species including NH$_4^+$ and NO$_3^-$ are also likely impacted by overpredictions in precipitation

586 (Figure 2d), which leads to an overprediction in their wet deposition rates and thus a reduction of

587 their ambient concentrations. The overpredictions in WS10 also help contribute to the deposition

588 of PM$_{2.5}$ and PM$_{2.5}$ species onto the ground (Sievering et al., 1987).

589 **3.3 Aerosol, Cloud, and Radiation Predictions**

590   There are uncertainties in the satellite retrievals of various aerosol-cloud-radiation

591 variables from the Clouds and the Earth's Radiant Energy System (CERES) and the Moderate

592 Resolution Imaging Spectroradiometer (MODIS). Loeb et al. (2009) reported that the major

593 uncertainties of the top of atmosphere radiative fluxes from CERES are derived from instrument

594 calibration (with a net error of 4.2 W m^{-2}), and the assumed value of 1 W m^{-2} for total solar

595 irradiance. However, there is good correlation (R > 0.8) between the model and CERES for the

596 radiation variables SWDOWN, GSW, and GLW, which are all measured at the surface (Table 2).

597 Modeled OLR at the top of the atmosphere also has relatively good correlation (R ~ 0.6).

598 SWDOWN and GLW are both slightly overpredicted due to influences from biases in PM

599 concentrations and clouds, but GSW and OLR are slightly underpredicted.

600   The overpredictions of the surface radiation variables are also impacted by the

601 underpredictions in AOD and COT. AOD is underpredicted with an NMB of -24.0%, and COT is

602 underpredicted with an NMB of -44.3%. These underpredictions indicate that less radiation is

603 attenuated (i.e., absorbed or scattered) or reflected while traversing through the atmospheric

604 column and clouds, thus allowing more radiation to reach the ground. Using the CESM model, He

605 et al. (2015) also showed underpredictions in AOD and COT over CONUS against MODIS

606 satellite retrievals. Figure 9 compares the spatial distributions of the 10-year average predictions

607 of AOD (a and b) against the satellite retrieval data from MODIS. The simulated AODs show

608 relatively large values over eastern U.S., due to the relatively higher PM concentrations in this

609 region of the U.S. The MODIS AOD, however, shows slightly elevated ~~AOD~~ values over eastern

610 U.S., but the magnitudes are not as high as the simulated AOD over eastern U.S. MODIS-derived

611 AOD is also higher over ~~the~~ western U.S. compared to ~~the~~ eastern U.S~~, and are also not as high as~~

612 ~~the MODIS-derived AOD over the western U.S~~., and this trend is not ~~seen~~found in the simulated

613 AOD. The differences between the MODIS AOD and the simulated AOD are likely due to the

614 differences in the algorithms used to retrieve AOD based on MODIS measurements and calculate

615 AOD in WRF/Chem. For MODIS, AOD is calculated by matching the spectral reflectance

616　observations with a lookup table based on a set of aerosol parameters including the aerosol size

617　distributions from a variety of aerosol models, which differ based on seasons and locations (Levy

618　et al., 2007). There are also different algorithms for dark land, bright land, and over oceans (Levy

619　et al., 2013). The MODIS data are aggregated into a global 1° gridded (Level-3) dataset with

620　monthly　　　　　　　　(MOD08_M3)　　　　　　　temporal　　　　　　　　resolution

621　(https://www.earthsystemcog.org/site_media/projects/obs4mips/TechNote_MODIS_L3_C5_Aer

622　osols.pdf). The inaccuracies for the calculation of AOD in WRF/Chem include biases in aerosol

623　size distribution, aerosol composition, aerosol water content, and reflectances. They can also arise

624　from parameterizations in the calculations including the assumption of an internally-mixed aerosol

625　composition. Therefore, caution should also be taken when comparing simulated AOD with the

626　satellite-derived AOD products. Toth et al. (2013) compared Aqua MODIS AOD products over

627　the mid to high latitude Southern Ocean where a band of enhanced AOD is observed, to cloud and

628　aerosol products produced by the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP)

629　project; and AOD data from the Aerosol Robotic Network (AERONET) and the Maritime Aerosol

630　Network (MAN). They concluded that the band of enhanced AOD is not detected in the CALIOP,

631　AERONET, or MAN products. The enhanced AOD band is attributed to stratocumulus and low

632　broken cumulus cloud contamination, as well as the misidentification of relatively warm cloud

633　tops compared with surrounding open seas. ~~Figures 9 a and b show the spatial distributions of the~~

634　~~10-year average predictions of AOD compared against the satellite retrieval data from MODIS.~~

635　~~The simulated AODs show relatively large values over eastern U.S., due to the relatively higher~~

636　~~PM concentrations in this region of the U.S. The MODIS AOD, however, does not show a similar~~

637　~~spatial pattern.~~

638   Figure 9 also shows spatial distributions of the 10-year average predictions of CDNC (c

639   and d), CWP (e and f), and COT (g and h), compared against the satellite retrieval data from

640   MODIS. The cloud variables CDNC, CWP, and COT tend to be underpredicted for most of the

641   regions over the U.S. However, CWP is largely overpredicted over the Atlantic ocean. This is also

642   likely due to the ~~infrequent monthly reinitialization of the WRF/Chem simulations in this study,~~

643   ~~which results in a~~ build-up of moisture over the Atlantic ocean, also influencing precipitation as

644   mentioned previously. CDNC is overpredicted over some regions in eastern U.S., but there are

645   also relatively large areas of underpredictions over both the land and ocean. This leads to an

646   average domain-wide underprediction for CDNC (Table 2). This is likely due to the differences in

647   deriving CDNC in the model and in the satellite retrievals. CDNC in the model is calculated based

648   on the activation parameterization by Abdul Razzak and Ghan (2000) based on the aerosol size

649   distribution, aerosol composition, and the updraft velocity. The MODIS-derived CDNC from

650   Bennartz (2007) is calculated based on cloud effective radius and COT, which would explain the

651   differences in spatial patterns between model and observed data. As indicated by Bennartz (2007),

652   the errors in CDNC can be up to 260%, especially for regions with low CF (< 0.1). The model and

653   MODIS spatial patterns are similar for CWP and COT over land, although the model values are

654   underpredicted. King et al. (2013) reported that the MODIS retrieval of cloud effective radius

655   when compared to in-situ observations is overestimated by 13% on average. Combined with

656   overestimations in COT, this leads to overestimation of liquid water path. In addition, there can

657   also be differences in satellite-derived cloud products from different satellites. For example, Shan

658   et al. (2011) showed that the derived CLDFRA from MODIS and another satellite, the Polarization

659   and Directionality of Earth Reflectances (POLDER) can differ with a global average of 10%.

660   Figure 10 shows similar spatial plots for modeled versus CERES derived SWDOWN,

661   OLR, SWCF, and LWCF. We note that modeled SWCF is calculated based on the differences

662   between the net cloudy sky and net clear sky shortwave radiation at the top of atmosphere, which

663   in turn are dependent on cloud properties including the CLDFRA, COT, cloud asymmetry

664   parameter, and cloud albedo. It is possible that due to the overprediction of CLDFRA, the

665   magnitudes of the simulated SWCF are greater than those from CERES (Figures 10c and 10g),

666   even though the other cloud variables are underpredicted. LWCF is calculated based on the

667   differences in clear-sky OLR and cloudy-sky OLR, which in turn are dependent on CLDFRA,

668   COT, and absorbance and radiance due to atmospheric gases. The underprediction of total-sky

669   OLR (Table 2 and Figures 10b and 10f) leads to an overprediction in LWCF. SWCF is largely

670   overpredicted over eastern U.S. and especially over the Atlantic ocean (Figures 10c and 10g).

671   LWCF is also overpredicted by the model in similar locations as SWCF, such as in southeastern

672   U.S., and over the ocean in the eastern portion of the domain (Figures 10d and 10h). This is further

673   confirmed by the underpredictions in SWDOWN over the Atlantic ocean and in general over the

674   eastern portion of the domain, as increased clouds (as a consequence of overpredicted AOD, CWP

675   and COT) and SWCF lead to less SWDOWN reaching the ground (Figures 10a and 10e) which

676   also eventually leads to a reduction in the OLR also over the eastern portion of the domain. The

677   larger negative SWCF and positive LWCF in the model compared to CERES, however, lead to an

678   overall good agreement with CERES for the net cloud forcing (SWCF + LWCF; not shown). The

679   mean bias for SWCF against CERES of 7.8 W m^{-2} and that for LWCF against CERES of 6.9 W

680   m^{-2} are comparable to the results from the CMIP5 models of -10 to 10 W m^{-2} over CONUS region

681   (Figure 9.5 in Flato et al., 2013). The evaluation of 10-year averaged predictions of aerosol-cloud-

682   radiation variables is similar to the results from the WRF/Chem simulations in 2006 and 2010 by

683 Yahya et al. (2014 and 2015). For example WRF/Chem generally performs well for cloud fraction

684 but AOD, CDNC, CWP and COT are underpredicted in both studies, which possibly indicate

685 consistent biases for every year contributing to climatological biases.

686 **4. Summary and Conclusions**

687   Overall, the model slightly underpredicts T2 with a mean bias of ~-0.3 °C, which is

688 consistent or better than other studies based on chemical transport models and regional climate

689 models. The underpredictions in T2 correlate to the overpredictions in RH2. WS10 biases are

690 likely due to issues with unresolved topography or due to inaccuracies in the selection of

691 representative grid points. There are seasonal biases in precipitation, where overpredictions tend

692 to occur largely over the summer months; however, precipitation is overpredicted every year

693 between 2001 and 2010 likely due ~~mainly~~ to uncertainties in WRF ~~microphysics and~~ cumulus and

694 microphysics parameterization~~s schemes~~ ~~as well as the accumulation of moisture due to the~~

695 ~~monthly reinitialization.~~ ~~T~~in particular, the use of a different cumulus parameterization scheme,

696 e.g., based on the MSKF available in WRF/Chem version 3.7 or newer ~~version~~ has been shown in

697 the sensitivity study to ~~possibly~~ significantly reduce precipitation biases. Other factors

698 contributing to the precipitation bias include the use of bias-corrected CESM_NCSU data (instead

699 of NCEP reanalysis data), and the use of an reinitialization frequency of 1-month. ~~More frequent~~

700 ~~reinitializations would help to reduce the biases in moisture, precipitation, and related cloud and~~

701 ~~radiation variables, however, a balance would have to be achieved between running continuous~~

702 ~~climate simulations and the frequency of reinitializations.~~ A satisfactory model performance for

703 meteorological variables is important and necessary when simulating future years, as data

704 evaluation is not possible. Meteorological variables such as temperature, humidity, wind speed

705 and direction, PBL height, and radiation have a strong impact on chemical predictions, and thus

31

706 are critical to the satisfactory model performance when predicting chemical variables such as O_3

707 and $PM_{2.5}$. Biases in O_3 and $PM_{2.5}$ concentrations can be attributed to biases in any of the

708 meteorological and chemical variables. The model performs generally well for radiation variables,

709 as well as for the main chemical species such as O_3 and $PM_{2.5}$, which indicates that the processed

710 RCP 8.5 emissions are reasonably accurate to produce acceptable results for the concentrations of

711 chemical species.

712 Modeled O_3 mixing ratios at the CASTNET sites are slightly underpredicted, but are

713 slightly overpredicted at AIRS-AQS sites, in part due to the fact that the CASTNET sites are

714 classified as rural, while the AIRS-AQS sites are classified as both urban and rural. O_3 mixing

715 ratios at the AIRS-AQS sites tend to be overpredicted during the colder fall and winter seasons,

716 and annually, O_3 mixing ratios are overpredicted every year from 2001 to 2010. O_3 mixing ratios

717 at the CASTNET sites are underpredicted for all climatological months, while the largest

718 underpredictions are observed from January to May. However, on a decadal time scale,

719 WRF/Chem adequately represents the different O_3 probability distributions at the AIRS-AQS and

720 CASTNET sites. This study also showed that peak O_3 mixing ratios are observed over April and

721 May rather than June to August, which is consistent with Cooper et al. (2014) who attributed this

722 to emission reductions and opposite trends in O_3 mixing ratios over eastern and western U.S. over

723 the last 20 years. Modeled $PM_{2.5}$ concentrations tend to be overpredicted at the IMPROVE sites

724 but underpredicted at the STN sites. $PM_{2.5}$ at the IMPROVE sites tend to be underpredicted in

725 spring and summer but overpredicted in fall and winter, while $PM_{2.5}$ concentrations against STN

726 are persistently underpredicted for all climatological months. The IMPROVE and STN sites are

727 classified as rural and urban, respectively. Due to the relatively coarse horizontal resolution of the

728 model (36×36 km), the model is unable to capture the locally higher $PM_{2.5}$ concentrations at the

729　STN sites. In general, however, the model performs relatively well for total $PM_{2.5}$ concentrations

730　at the IMPROVE and STN sites with NMBs of within ±25%, although larger biases exist for $PM_{2.5}$

731　species. Model performance for PM_{10} should be improved, as PM_{10} also has important impacts on

732　climate through influencing the radiative budget both directly and indirectly due to its larger size

733　and higher concentrations. The choice of observational networks for model evaluation are

734　therefore important as both networks can show positive and negative biases depending on the type

735　and location of the sites (e.g., O_3 against AIRS-AQS and CASTNET, and $PM_{2.5}$ against STN and

736　IMPROVE). The major uncertainties lie in the predictions of cloud-aerosol variables. As

737　demonstrated in this study, large biases and error in simulating cloud variables even in the most

738　advanced models such as WRF/Chem, indicating a need for future improvement in relevant model

739　treatments such as cloud dynamics and thermodynamics, as well as aerosol-cloud interactions. In

740　addition, there are large uncertainties in satellite retrievals of cloud variables for evaluation. In this

741　study, most of the cloud-aerosol variables including AOD, COT, CWP, and CDNC are on average

742　underpredicted across the domain; however, the overpredictions of cloud variables including COT

743　and CWP over the Atlantic ocean and eastern U.S. lead to underpredictions in radiation and

744　overpredictions in cloud forcing, which are important parameters when simulating future climate

745　change.

746　　　In summary, the model is able to predict O_3 mixing ratios and $PM_{2.5}$ concentrations

747　relatively well with regards to decadal scale air quality and climate applications. The model is able

748　to predict meteorological variables satisfactorily and with results comparable to RCM and GCM

749　applications from literatures. Possible reasons behind the chemical and meteorological biases

750　identified through this work should be taken into account when simulating longer climatological

751　periods and/or future years. Aerosol-cloud-radiation variables are important for climate

752 simulations, the performance of these variables are not as good as that of the chemical and

753 meteorological variables. They contain consistent biases in single-year evaluations of WRF/Chem.

754 However, magnitudes of biases for SWCF and LWCF are comparable to those from literature,

755 which suggests that model improvements should be made in terms of bias correction of

756 downscaled ICs/BCs as well as aerosol-cloud-radiation parameterizations in the model. In

757 addition, having consistent physical and chemical mechanisms between the GCM and RCMs could

758 help to reduce uncertainties in the results (Ma et al., 2014). Although CESM and WRF/Chem use

759 similar chemistry and aerosol treatments in this work, they use somewhat different physics

760 schemes which may contribute to such uncertainties. The development of scale-aware

761 parameterizations that can be applied at both global and regional scales would help reduce

762 uncertainties associated with the use of different schemes for global simulations and downscaled

763 regional simulations.

770 _____

771 **Code and Data Availability**

772 The WRF/Chem v3.6.1 code used in this paper will be available upon request. However,

773 we highly encourage users to download the latest available version of the WRF/Chem code from

774 NOAA's web site at http://www2.mmm.ucar.edu/wrf/users/download/get_source.html. The

775 updates in our in-house version of WRF/Chem v3.6.1 has been implemented into WRF/Chem

776 v3.7 and WRF/Chem v3.7.1 for scientific community release. The WRF/Chem v3.7 and

777 WRF/Chem v3.7.1 codes are now publicly available at

778 http://www2.mmm.ucar.edu/wrf/users/download/get_source.html. These latest versions of the

779 source codes contain all major changes in the standard version of WRF/Chem v3.6.1 used in for

780 this study. In addition, they have been rigorously tested for compatibility and compiling issues

781 on various platforms. The inputs including the meteorological files, meteorological initial and

782 boundary conditions, chemical initial and boundary conditions, model set-up and configuration,

783 and the namelist set-up, and instructions on how to run the simulations for a 1-day test case, as

784 well as a sample output for 1-day test can be provided upon request.

785

786 **References**

787 Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation, 2. Multiple aerosol

788   types, J. Geophys. Res., 105(5), 6837-6844, 2000.

789 Aitken, A.C., DeCarlo, P.F., Kroll, J.H., Worsnop, D.R., Huffman, J.A., Docherty, K.S., Ulbrich,

790   I.M., Mohr, C., Kimmel, J.R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M.,

791   Ziemann, P.J., Canagaratna, M.R., Onasch, T.B., Alfarra, M.R., Prevot, A.S.H., Dommen, J.,

792   Duplissy, J., Metzger, A., Baltensperger, U. and Jimenez, J.L.: O/C and OM/OC ratios of

793   primary, secondary and ambient organic aerosols with high-resolution time of flight aerosol

794   mass spectrometry, Environ. Sci. Technol., 42, 4478 – 4485, 2008.

795 Alapaty, K., Herwehe, J., Nolte, C.G., Bullock, R.O., Otte, T.L., Mallard, M.S., Dudhia, J. and

796   Kain, J.S.: Introducing subgrid-scale cloud feedbacks to radiation in WRF, the 13th WRF Users

797   Workshop, Boulder, CO, June 26 to 29, 2012.

798   Ahmadov, R., McKeen, S.A., Robinson, A.L., Bareini, R., Middlebrook, A.M., De Gouw, J.A.,

799   Meagher, J., Hsie, E.-Y., Edgerton, E., Shaw, S. and Trainer, M.: A volatility basis set model

800   for summertime secondary organic aerosols over the eastern United States in 2006, J.

801   Geophys. Res. 117, D06301, doi:10.1029/2011JD016831, 2012.

802   Beniston, M., Stephenson, D.B., Christensen, O.B., Ferro, C.A.T., Frei, C., Goyette, S.,

803   Halsnaes, K., Holt, T., Jylha, K., Koffi, B., Palutikof, J., Scholl, R., Semmler, T. and Woth,

804   K.: Future extreme events in European climate: an exploration of regional climate model

805   projections, Clim. Change, 81, 71 – 95, doi: 10.1007/s10584-006-9226-z, 2007.

806   Bennartz, R.: Global assessment of marine boundary layer cloud droplet number concentration

807   from satellite, J. Geophys. Res-Atmos, 112(D2), D02201, doi:10.1029/2006JD007547, 2007.

808   Brunner, D., Savage, N., Jorba, O., Eder, B., Giordano, L., Badia, A., Balzarini, A., Baro, R.,

809   Bianconi, R., Chemel, C., Curci, G., Forkel, R., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A.,

810   Hozak, L., Im, U., Knote, C., Makar, P., Manders-Groot, A., van Meijgaard, E., Neal, L.,

811   Perez, J.L., Pirovano, G., San Jose, R., Schroder, W., Sokhi, R.S., Syrakov, D., Torian, A.,

812   Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Hogrefe, C. and

813   Galmarini, S.: Comparative analysis of meteorological performance of coupled chemistry-

814   meteorology models in the context of AQMEII phase 2, Atmos. Environ., in press,

815   doi:10.1016/j.atmosenv.2014.12.032, 2014.

816   Caldwell, P., H.-N.S. Chin, D.C. Bader, and G. Bala (2009), Evaluation of a WRF dynamical

817   downscaling simulation over California, Clim. Change., 95, 499-521.

818   Campbell, P. C**.,** Zhang, Y., Yahya, K., Wang, K., Hogrefe, C., Pouliot, G., Knote, C., Hodzic,

819   A., San Jose, R., Perez, J., Jimenez-Guerrero, P., Baro, R. and Makar, P.: A Multi-Model

820   Assessment for the 2006 and 2010 Simulations under the Air Quality Model Evaluation

821   International Initiative (AQMEII) Phase 2 over North America: Part I, Indicators of the

822   Sensitivity of O_3 and $PM_{2.5}$ Formation Regimes, Atmos. Environ., in press,

823   doi:10.1016/j.atmosenv.2014.12.026, 2014.

824   Chen, F. and Dudhia, J.: Coupling an advanced land-surface/hydrology model with the Penn

825   State/NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Mon.

826   Wea. Rev., 129, 569-585, 2001.

827   Clough, S.A., Shephard, M.W., Mlawer, J.E., Delamere, J.S., Iacono, M.J., Cady-Pereira, K.,

828   Boukabara, S. and Brown, P.D.: Atmospheric radiative transfer modeling: a summary of the

829   AER codes, J. Quant. Spectrosc. Radiat. Transfer, 91(2), 233 – 244, doi:

830   10.1016/j.qsrt.2004.05.058, 2005.

831   Cooper, O.R., Parrish, D.D., Ziemke, J., Balashov, N.V., Cupeiro, M., Galbally, I.E., Gilge, S.,

832   Horowitz, L., Jensen, N.R., Lamarque, J.-F., Naik, V., Oltmans, S.J., Schwab, J., Shindell,

833   D.T., Thompson, A.M., Thouret, V., Wang, Y. and Zbinden, R.M.: Global distribution and

834   trends of tropospheric ozone: An observation-based review, Elem. Sci. Anth., 2, 000029,

835   doi:10.12952/journal.elementa.000029, 2014.

836   Dasari, H.P., Salgado, R., Perdigao, J. and Challa, V.S.: A regional climate simulation study

837   using WRF-ARW model over Europe and evaluation for extreme temperature weather

838   events, Intl., J. of Atmos. Sci., ID 704079, doi:10.1155/2014/704079, 2014.

839   Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. and Tarpley,

840   J.D.: Implementation of NOAH land surface model advances in the National Centers for

841   Environmental Prediction operational mesoscale model, J. Geophys. Res., 108, D22, 8851,

842   doi:10.1029/2002JD003296, 2003.

843   EPA.: Our Nation's Air – Status and Trends through 2010, Particle Pollution, Report by the U.S.

844   EPA, 4pp, http://www.epa.gov/airtrends/2011, 2011, last accessed July 6th, 2015.

845   Fan, F., Bradley, R.S. and Rawlins, M.A.: Climate change in the northeastern U.S.: regional

846   climate validation and climate change projections, Clim. Dyn., 43, 145 – 161,

847   doi:10.1007/s00382-014-2198-1, 2014.

848   Feser, F., Rockel, B., Von Storch, H., Winterfeldt, J. and Zahn, M.: Regional climate models add

849   value to global model data, Bull. Amer. Meteor. Soc., 92, 1181 – 1192, 2011.

850   Flato et al.: Evaluation of Climate Models, In: Climate Change 2013: The Physical Science

851   Basis. Contribution of Working Group I to the Fifth Assessment Report of the

852   Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M.

853   Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)],

854   Cambridge University Press, Cambridge, United Kingdom and New York, NY, U.S.A.,

855   2013.

856   Gao, Y., Fu, J.S., Drake, J.B., Liu, Y. and Lamarque, J.F.: Projected changes of extreme weather

857   events in the eastern United States based on a high resolution climate modeling system,

858   Environ. Res. Lett., 7, 044025, 2012.

859   Gao, Y., Fu, J.S., Drake, J.B., Lamarque, J.-F. and Liu, Y.: The impact of emission and climate

860   change on ozone in the United States under representative concentration pathways (RCPs),

861   Atmos. Chem. Phys., 2013, 9607 – 9621, 2013.

862   Glotfelty, T., He, J. and Zhang, Y.: Updated organic aerosol treatments in CESM/CAM5:

863   development and initial application, in preparation, 2015.

864   Gong, S., Barrie, L.A. and Blanchet, J.P.: Modeling sea salt aerosols in the atmosphere: 1. Model

865   development, J. Geophys. Res., 102, 3805-3818, doi:10.1029/96JD02953, 1997.

866 Grell, G.A., Knoche, R., Peckham, S.E. and McKeen, S.A.: Online versus offline air quality

867   modeling on cloud-resolving time scales, Geophys. Res. Lett., 31 (16),

868   doi:10.1029/2004GL020175, 2004.

869 Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C. and Eder,

870   B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957-

871   6975, 2005.

872 Grell, G.A. and Freitas, S.R.: A scale and aerosol aware stochastic convective parameterization

873   for weather and air quality modeling, Atmos. Chem. Phys., 14, 5233-5250, doi:10.5914/acp-

874   14-5233-2014, 2014.

875 Guenther, A., Kart, T., Harley, P., Wiedinmyer, C., Palmer, P.I. and Geron, C.: Estimates of

876   global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and

877   Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210, 2006.

878 He, J., Zhang, Y., Glotfelty, T., He, R., Bennartz, R., Rausch, J. and Sartelet, K.: Decadal

879   simulation and comprehensive evaluation of CESM/CAM5.1 with advanced chemistry,

880   aerosol microphysics and aerosol-cloud interactions, J. Adv. Model. Earth Syst., 7, 110 –

881   141, doi:10.1002/2014MS000360, 2015.

882 Hong, S.-Y., Noh, Y. and Dudhia, J.: A new vertical diffusion package with an explicit treatment

883   of entrainment processes, Mon. Wea. Rev., 134, 2318-2341, 2006.

884 Hong, S.-Y.: A new stable boundary-layer mixing scheme and its impact on the simulated East

885   Asian summer monsoon, Q.J.R. Meteorol. Soc., 136, 1481 – 1496, doi:0.1002/qj.665, 2010.

886 Hurrell, J.W., Holland, M.M., Gent, P.R., Ghan, S., Kay, J.E., Kushner, P.J., Lamarque, J.-F.,

887   Large, W.G., Lawrence, D., Lindsay, K., Lipscomb, W.H., Long, M.C., Mahowald, N.,

888   Marsh, D.R., Neale, R.B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W.D.,

889   Hack, J.J., Kiehl, J. and Marshall, S.: The Community Earth System Model: A framework for

890   collaborative research, Bull. Am. Meteorol. Soc., 94, 1339 – 1360, doi:10.1175/BAMS-D-

891   12-00121.1, 2013.

892   Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shepard, M.W., Clough, S.A. and Collins, W.D.:

893   Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative

894   transfer models, J. Geophys. Res., 113, D13103, doi:10.1029/2008JD009944, 2008.

895   IPCC : Climate change 2013: the physical science basis. In: Stocker, T.F., Qin, D., Plattner, G.-

896   K., Tignor, M.M.B., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M.

897   (Eds.), Contribution of Working Group I to the Fifth Assessment Report of the

898   Intergovernmental Panel on Climate Change, Summary for Policymakers, 2013.

899   Jacob, D., Barring, L., Christensen, O.B., Christensen, J.H., de Castro, M., Deque, M., Giorgi, F.,

900   Hagemann, S., Hirschi, M., Jones, R., Kjellstrom, E., Lenderink, G., Rockel, B., Sanchez, E.,

901   Schar, C., Seneviratne, S.I., Somot, S., van Ulden, A. and van den Hurk, B.: An inter-

902   comparison of regional climate models for Europe: model performance in present-day

903   climate, Clim. Change, 81, 31 – 52, 2007.

904   Jimenez, P.A. and Dudhia, J.: Improving the representation of resolved and unresolved

905   topographic effects on surface wind in the WRF model, J. Appl. Meteor. Climatol., 51, 300 –

906   316, 2012.

907   Jones, R.G., Noguer, M., Hassell, D.C., Hudson, D., Wilson, S.S., Jenkins G.J. and Mitchell,

908   J.F.B.: Generating high resolution climate change scenarios using PRECIS, Met Office

909   Hadley Centre, Exeter, UK, 40 pp., April 2004, 2004.

910   Jones, S. and Creighton, G.: AFWA dust emission scheme for WRF/Chem-GOCART, 2011

911   WRF workshop, June 20-24, Boulder, CO, USA, 2011.

912   Karamchandani, P., Zhang, Y., Chen, S.-Y., and Balmori-Bronson, R.: Development of an

913   extended chemical mechanism for global-through-urban applications, Atmos. Poll. Res., 3, 1

914   – 24, doi:10.5094/apr.2011.047.

915   Kim, J., Waliser, D.E., Mattmann, C.A., Mearns, L.O., Goodale, C.E., Hart, A.F., Crichton,

916   D.J., McGinnis, S., Lee, H., Loikith, P.C. and Boustani, M.: Evaluation of the surface

917   climatology over the conterminous United States in the North American Regional Climate

918   Change Assessment Program Hindcast Experiment using a regional climate model evaluation

919   system, J. Climate, 26, 5698 – 5715, 2013.

920   King, N.J., Bower, K.N., Crosier, J. and Crawford, I.: Evaluating MODIS cloud retrievals with in

921   situ observations from VOCALS-REx, Atmos. Chem. Phys., 13, 191 – 209, 2013.

922   Legates, D.R. and McCabe Jr., G.J.: Evaluating the use of "goodness-of-fit" measures in

923   hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233 – 241,

924   doi:10.1029/1998WR900018, 1999.

925   Levy, R.C., Remer, L.A. and Dubovik, O.: Global aerosol optical properties and application to

926   Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land, J. Geophys.

927   Res., 112(D13), doi:10.1029/2006JD007815, 2007.

928   Levy, R.C., Mattoo, S., Muchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., Hsu, N.C.: The

929   Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989 –

930   3034, 2013.

931   Leung, R.L., Qian, Y. and Bian, X.: Hydroclimate of the Western United States based on

932   Observations and Regional Climate Simulation of 1981 – 2000, Part I: Seasonal Statistics, J.

933   Clim., 16(12), 1892 – 1911, 2003.

934 Liu, X., Easter, R.C., Ghan, S.J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A.,

935   Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A.M.L., Hess, P.,

936   Mahowald, N., Collins, W., Iacono, M.J., Bretherton, C.S., Flanner, M.G., and Mitchell, D.:

937   Toward a minimal representation of aerosols in climate models: description and evaluation in

938   the Community Atmosphere Model CAM5, Geosci. Mod. Dev., 5, 709 – 739,

939   doi:10.5194/gmd-5-709-2012, 2012.

940 Lo, J.C.F., Yang, Z.L. and Pielke Sr., R.A.: Assessment of three dynamical climate downscaling

941   methods using the Weather Research and Forecasting (WRF) model, J. Geophys. Res., 113,

942   D09112, doi:10.1029/2007JD009216, 2008.

943 Loeb, N.G., Wielicki, B.A., Doelling, D.R., Smith, L., Keyes, D.F., Kato, S., Manalo-Smith, N.

944   and Wong, T.: Toward Optimal Closure of the earth's top-of-atmosphere radiation budget, J.

945   Climate, 22, 748 – 766, 2009.

946 Ma, P.-L., Rasch, P.J., Fast, J.D., Easter, R.C., Gustafson Jr., W.I., Liu, X., Ghan, S.J. and Singh,

947   B.: Assessing the CAM5 physics suite in the WRF-Chem model: implementation, resolution

948   sensitivity, and a first evaluation for a regional case study, Geosci. Model Dec., 7, 755 – 778,

949   2014.

950 Mass, C.: Improved subgrid drag or hyper PBL/vertical resolution? Dealing with the stable PBL

951   problems in WRF, presented at the 13th WRF Users' Workshop, June 26 – 29, Boulder, CO,

952   2012.

953 Molders, N., Bruyere, C.L., Gende, S. and Pirhala, M.A.: Assessment of the 2006-2012

954   Climatological Fields and Mesoscale Features from Regional Downscaling of CESM Data by

955   WRF/Chem over Southeast Alaska, Atmos. Clim. Sci., 4, 589 – 613, 2014.

956 Morrison, H., Thompson, G. and Tatarskii, V.: Impact of cloud microphysics on the development

957   of trailing stratiform precipitation in a simulated squall line: Comparison of One- and Two-

958   Moment Schemes, Mon. Wea. Rev., 137, 991-1007, 2009.

959 Moss, R. H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P.,

960   Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic,

961   N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P. and Wilbanks, T.J.:

962   The next generation of scenarios for climate change research and assessment, Nature, 463,

963   747 – 756, doi: 10.1038/nature0882, 2010.

964 Nasrollahi, N., AghaKouchak, A., Li, J., Gao, X., Hsu, K. and Sorooshian, S.: Assessing the

965   Impacts of Different WRF Precipitation Physics in Hurricane Simulations, Wea. Forecasting,

966   27, 1003 – 1016, 2012.

967 Neale R.B., Jadwiga, H.R., Conley, A.J., Park, S., Lauritzen, P.H., Gettelman, A., Williamson,

968   D.L., Rasch, P., Vavrus, S.J., Taylor, M.A., Collins, W.D., Zhang, M. and Lin, S.-J.:

969   Description of the NCAR Community Atmosphere Model (CAM 5.0), NCAR Tech. Note

970   NCAR/TN-486+STR, Natl. Cent. for. Atmos. Res., Boulder, CO, available at

971   http://www.cesm.ucar.edu/models/cesm1.0/cam/docs/description/cam5_desc.pdf, 2010, last

972   accessed July 6th, 2015.

973 Otte, T.L., Nolte, C.G., Otte, M.J. and Bowden, J.H.: Does Nudging squelch the extremes in

974   regional climate modeling? J. Clim., 25, 7046 – 7066, doi:10.1175/JCLI-D-12-00048.1,

975   2012.

976 Penrod, A., Zhang, Y., Wang, K., Wu, S-Y. and Leung, R.L.: Impacts of future climate and

977   emission changes on U.S. air quality, Atmos. Environ., 89, 533 – 547, 2014.

978   Petikainen, J.-P., O'Donnell, D., Teichmann, C., Karstens, U., Pfeifer, S., Kazil, J., Podzun, R.,

979   Fiedler, S., Kokkola, H., Birmili, W., O'Dowd, C., Baltensperger, U., Weingartner, E.,

980   Gehrig, R., Spindler, G., Kulmala, M., Feichter, J., Jacob, D. and Laaksonen, A.: The

981   regional aerosol-climate model REMO-HAM, Geosci. Mod. Dev., 5, 1323 – 1339, 2012.

982   Pleim, J.E. and Gilliam, R.: An indirect data assimilation scheme for deep soil temperature in the

983   Pleim-Xiu Land Surface Model, J. Appl. Meteor. Climatol., 48, 1362 – 1376, 2009.

984   Pouliot, G., van der Gon, H.A.C.D., Kuenen, J., Zhang, J., Moran, M. and Makar, P.: Analysis of

985   the Emission Inventories and Model-Ready Emission Datasets of Europe and North America

986   for Phase 2 of the AQMEII Project, Atmos. Environ., in press,

987   doi:10.1016/j.atmosenv.2014.10.061, 2014.

988   Rawlins, M.A., Bradley, R.S. and Diaz, H.F.: Assessment of regional climate model simulation

989   estimates over the northeast United States, J. Geophys. Res., 117, D23112,

990   doi:10.1029/2012JD018137, 2012.

991   Refslund, J., Dellwik, E., Hahmann, A.N., Barlage, M.J. and Boegh, E.: Development of satellite

992   green vegetation fraction time series for use in mesoscale modeling: application to the

993   European heat wave 2006, Theor. Appl. Climatol., 117, 377-392, doi:10.1007/s00704-013-

994   1004-z, 2014.

995   Sarwar, G., Luecken, D.J. and Yarwood, G.: Developing and implementing an updated chlorine

996   chemistry into the Community Multiscale Air Quality Model, presented at the 28th

997   NATO/CCMS International Technical Meeting, Lepzig, Germany, May 15 – 19, 2006.

998   Sarwar, G., Luecken, D. and Yarwood, G.: Chapter 2.9: Developing and implementing an

999   updated chlorine chemistry into the community multiscale air quality model, Developments

1000   in Environmental Science, Volume 6, C. Borrego and E. Renner (Eds.), Elsevier Ltd,

1001   DOI:10.1016/S1474-8177(07)06029-9, 168 pp., 2007.

1002   Sarwar, G., Fahey, K., Napelenok, S., Roselle, S. and Mathur, R.: Examining the impact of

1003   CMAQ model updates on aerosol sulfate predictions, the 10th Annual CMAS Models-3

1004   User's Conference, October, Chapel Hill, NC, 2011.

1005   Shan, Z., Parol, F., Riedi, J., Cornet, C. and Thieuleux, F.: Examination of POLDER/PARASOL

1006   and MODIS/Aqua cloud fractions and properties representativeness, J. Climate, 24, 4435 –

1007   4450, 2011.

1008   Sievering, H.: Small-particle dry deposition under high wind speed conditions: Eddy flux

1009   measurements at the boulder atmospheric observatory, Atmos. Environ., 21 (10), 2179 –

1010   2185, 1987.

1011   Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M.A., Mitchell, K., Ek, M., Gayno, G.,

1012   Wegiel, J. and Cuenca, R.H.: Implementation and verification of the unified NOAH land

1013   surface model in the WRF model. 20th conference on weather analysis and forecasting/16th

1014   conference on numerical weather prediction, pp. 11 – 15, 2004.

1015   Toth, T.D., Zhang, J., Campbell, J.R., Reid, J.S., Shi, Y., Johnson, R.S., Smirnov, A., Vaughan,

1016   M.A. and Winker, D.M.: Investigating enhanced Aqua MODIS aerosol optical depth

1017   retrievals over the mid-to-high latitude Southern Oceans through intercomparison with co-

1018   located CALIOP, MAN and AERONET data sets, J. Geophys. Res: Atmos, 18, 1- 15, 2013.

1019   van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C.,

1020   Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith,

1021   S.J. and Rose, S.K.: The representative concentration pathways: an overview, Climate

1022   Change, 109, 5 – 31, doi: 10.1007/s10584-011-0148-z, 2011.

1023   Wang, K., Zhang, Y., Yahya, K., Wu, S.-Y. and Grell, G.:, Implementation and initial

1024   application of new chemistry-aerosol options in WRF/Chem for simulating secondary

1025   organic aerosols and aerosol indirect effects for regional air quality, Atmos. Environ., in

1026   press, doi: 10.1016/j.atmosenv.2014.12.007, 2014a.

1027   Wang, K., Yahya, K., Zhang, Y., Hogrefe, C., Pouliot, G., Knote, C., Hodzic, A., San Jose, R.,

1028   Perez, J.L., Guerrero, P.J., Baro, R. and Makar, P.: Evaluation of Column Variable

1029   Predictions Using Satellite Data over the Continental United States: A Multi-Model

1030   Assessment for the 2006 and 2010 Simulations under the Air Quality Model Evaluation

1031   International Initiative (AQMEII) Phase 2, Atmos. Environ., in press,

1032   doi:10.1016/j.atmosenv.2014.07.044, 2014b.

1033   Warrach-Sagi, K., Schwitalla, T., Wulfmeyer, V. and Bauer, H.-S.: Evaluation of a climate

1034   simulation in Europe based on the WRF-NOAH model system: precipitation in Germany,

1035   Clim. Dyn., 41, 755 – 774, doi:10.1007/s00382-013-1727-7, 2013.

1036   Willmott, C. J.: On the validation of models, Phys. Geog., 2, 184 – 194, 1981.

1037   Xing, J., Mathur, R., Pleim, J., Hogrefe, C., Gan, C.-M., Wong, D.C., Wei, C., Gilliam, R. and

1038   Pouliot, G.: Observations and modeling of air quality trends over 1990-2010 across the

1039   Northern Hemisphere: China, the United States and Europe, Atmos. Chem. Phys., 15, 2723 –

1040   2747, doi:10.5194/acp-15-2723-2015.

1041   Xu, Z. and Yang, Z.-L.: An improved dynamical downscaling method with GCM Bias

1042   Corrections and Its Validation with 30 years of climate simulations, J. Clim., 25, 6271 –

1043   6286, 2012.

1044 Yahya, K., Wang, K., Gudoshava, M., Glotfelty, T. and Zhang, Y.: Application of WRF/Chem

1045   over North America under the AQMEII Phase 2. Part I. Comprehensive Evaluation of 2006

1046   Simulation, Atmospheric Environment, in press, doi:10.1016/j.atmosenv.2014.08.063, 2014.

1047 Yahya, K., He, J., and Zhang, Y.: Multi-Year Applications of WRF/Chem over Continental U.S.:

1048   Model Evaluation, Variation Trend, and Impacts of Boundary Conditions over CONUS, J.

1049   Geophy. Res., in review, 2015a.

1050 Yahya, K., Wang, K., Zhang, Y. and Kleindienst, T.E.: Application of WRF/Chem over North

1051   America under the AQMEII Phase 2. Part II. Comprehensive Evaluation of 2010 Simulation

1052   and Responses of Air Quality and Meteorology-Chemistry Interactions to Changes in

1053   Emissions and Meteorology from 2006 to 2010, Geosci. Model Dev., in press, 2015b.

1054 Yarwood, G., Rao, S., Yocke, M. and Whitten, G.Z.: Final Report – Updates to the Carbon Bond

1055   Chemical Mechanism: CB05, Rep. RT-04-00675, 246 pp., Yocke and Co., Novato, Calif.,

1056   2005.

1057 Yu, S., Dennis, R., Roselle, S., Nenes, A., Walker, J., Eder, B., Schere, K., Swall, J., and

1058   Robarge, W.: An assessment of the ability of 3-D air quality models with current

1059   thermodynamic equilibrium models to predict aerosol NO3-, J. Geophys. Res., 110, D07S13,

1060   doi:10.1029/2004JD004718, 2005.

1061 Yu, S., Eder, B., Dennis, R., Chu, S.-H., and Schwartz, S.: New unbiased symmetric metrics for

1062   evaluation of air quality models, Atmos. Sci. Lett., 7, 26 – 34, 2006.

1063 Yu, S., Mathur, R., Pleim, J., Wong, D., Gilliam, R., Alapaty, K., Zhao, C., and Liu, X.: Aerosol

1064   indirect effect on the grid-scale clouds in the two-way coupled WRF-CMAQ: model

1065   description, development, evaluation and regional analysis, Atmos. Chem. Phys., 14, 11247 –

1066   11285, doi:10.5194/acp-14-1-2014, 2014.

1067   Zhang, Y., Liu, P., Pun, B., and Seigneur, C.: A comprehensive performance evaluation of

1068   MM5-CMAQ for summer 1999 Southern Oxidants Study Episode, Part-I. Evaluation

1069   Protocols, Databases and Meteorological Predictions, Atmos. Environ., 40, 4825 – 4838,

1070   2006.

1071   Zhang, Y., Wen, X.-Y. and Jang, C.J.: Simulating chemistry-aerosol-cloud-radiation-climate

1072   feedbacks over the CONUS using the online-coupled Weather Research Forecasting Model

1073   with chemistry (WRF/Chem), Atmos. Environ., 44, 3568 – 3582, 2010.

1074   Zhang, Y., Y.-C. Chen, G. Sarwar, and K. Schere, 2012a, Impact of Gas-Phase Mechanisms on

1075   Weather Research Forecasting Model with Chemistry (WRF/Chem) Predictions: Mechanism

1076   Implementation and Comparative Evaluation, ~~Journal of Geophysical Research~~J. Geophys.

1077   Res., 117, D1, doi:10.1029/2011JD015775.

1078   Zhang, Y., P. Karamchandani, T. Glotfelty, D. G. Streets, G. Grell, A. Nenes, F.-Q. Yu, and R.

1079   Bennartz, 2012b, Development and Initial Application of the Global-Through-Urban

1080   Weather Research and Forecasting Model with Chemistry (GU-WRF/Chem), J. Geophys.

1081   Res.~~Journal of Geophysical Research~~, 117, D20206, doi:10.1029/2012JD017966.

1082

1083

Table 1. Model configurations and set-up

Model Attribute	Configuration	Reference
Domain and Resolutions	36km × 36km, 148 × 112 horizontal resolution over continental U.S., with 34 layers vertically from surface to 100 hPa	-
Simulation Period	January 2001 to December 2010	-
Chemical and Meteorological ICs/BCs	Downscaled from the modified Community Earth System Model/Community Atmosphere Model (CESM/CAM5) v1.2.2; Meteorological ICs/BCs bias-corrected with National Center for Environmental Protection's Final (FNL) Operational Global Analysis data	He et al. (2014) Glotfelty et al. (2015)
Biogenic Emissions	Model of Emissions of Gases and Aerosols from Nature (MEGAN2)	Guenther et al. (2006)
Dust Emissions	Atmospheric and Environmental Research Inc. and Air Force Weather Agency (AER/AFWA)	Jones and Creighton (2011)
Sea-Salt Emissions	Gong et al. parameterization	Gong et al. (1997)
Radiation	Rapid and accurate Radiative Transfer Model for GCM (RRTMG) SW and LW	Clough et al. (2005) Iacono et al. (2008)
Boundary Layer	Yonsei University (YSU)	Hong et al. (2006) Hong (2010)
Land Surface	National Center for Environmental Prediction, Oregon State University, Air Force and Hydrologic Research Lab (NOAH)	Chen and Dudhia (2001) Ek at al. (2003) Tewari et al. (2004)
Microphysics	Morrison double moment scheme	Morrison et al. (2009)
Cumulus Parameterization	Grell 3D Ensemble	Grell and Freitas (2014)
Gas-phase chemistry	Modified CB05 with updated chlorine chemistry	Yarwood et al. (2005) Sarwar et al. (2006) Sarwar et al. (2007)
Photolysis	Fast Troposphere Ultraviolet Visible (FTUV)	Tie et al. (2003)
Aqueous-phase chemistry	AQ chemistry module (AQCHEM) for both resolved and convective clouds	Based on AQCHEM in CMAQv4.7 of (Sarwar et al. 2011)
Aerosol module	MADE/VBS	Ahmadov et al. (2012)
Aerosol Activation	Abdul-Razzak and Ghan	Abdul-Razzak and Ghan (2000)

Table 2. The 10-year (2001 – 2010) average performance statistics for the simulated meteorological, aerosol, cloud, radiation variables, and chemical species against surface observational networks and satellite retrieval products.

Database and Variable	Mean Obs	Mean Sim	R	MB	NMB (%)	NME (%)
NCDC T2 (ºC)	12.5	12.2	1.0	-0.3	-2.6	7.9
NCDC RH2 (%)	68.4	70.8	0.8	2.4	3.5	6.8
NCDC WS10 (m s^{-1})	3.54	3.84	0.3	0.3	8.6	28.4
NCDC WD10 (deg)	151.4	180.0	0.2	28.6	18.9	22.0
NADP Precip (mm day^{-1})	18.0	26.3	0.5	8.3	45.9	65.1
CERES SWDOWN (W m^{-2})	184.1	184.6	0.8	0.5	0.3	8.4
CERES GSW (W m^{-2})	157.5	151.8	0.8	-5.7	-3.6	9.6
CERES GLW (W m^{-2})	323.3	325.7	1.0	2.4	0.7	1.8
CERES OLR (W m^{-2})	240.0	224.8	0.6	-15.0	-6.3	6.3
MODIS AOD	0.14	0.10	0.1	-0.03	-24.0	38.5
MODIS CLDFRA	58.3	62.0	0.7	3.7	6.4	11.9
MODIS-derived CDNC (cm^{-3})	169.8	130.0	0.4	-39.9	-23.5	38.0
MODIS CWP (g m^{-2})	179.5	170.0	0.3	-9.6	-5.3	61.2
MODIS COT	16.5	9.2	0.2	-7.3	-44.3	54.0
CERES SWCF (W m^{-2})	-41.8	-49.6	0.5	7.8	18.6	31.4
CERES LWCF (W m^{-2})	24.8	31.8	0.6	6.9	28.0	34.7
AQS Hourly O$_3$ (ppb)	29.3	32.1	0.6	2.8	9.7	22.4
AQS Max 1-hr O$_3$ (ppb)	48.9	49.7	0.6	0.8	1.7	7.9
AQS Max 8-hr O$_3$ (ppb)	43.7	45.9	0.6	2.2	5.0	9.3
CASTNET Hourly O$_3$ (ppb)	35.0	31.9	0.7	-3.1	-8.8	19.8
CASTNET Max-1hr O$_3$ (ppb)	47.4	38.5	0.4	-8.9	-18.8	31.4
CASTNET Max 8-hr O$_3$ (ppb)	43.3	37.9	0.5	-5.4	-12.5	29.6
AQS 24-hr PM$_{10}$ (µg m^{-3})	22.5	11.0	0.1	-11.5	-51.2	57.1
IMPROVE PM$_{2.5}$ (µg m^{-3})	5.33	6.57	0.4	1.2	23.3	53.4
STN PM$_{2.5}$ (µg m^{-3})	12.0	10.7	0.2	-1.3	-10.8	38.3
IMPROVE SO$_4^{2-}$ (µg m^{-3})	1.45	1.86	0.8	0.4	28.0	41.8
STN SO$_4^{2-}$ (µg m^{-3})	3.10	3.74	0.7	0.6	20.7	36.8
IMPROVE[1] NO$_3^-$ (µg m^{-3})	0.54	0.44	0.7	-0.1	-17.9	64.6
STN NO$_3^-$ (µg m^{-3})	1.62	0.70	0.4	-0.9	-56.9	65.3
IMPROVE NH$_4^+$ (µg m^{-3})	1.02	0.72	0.4	-0.3	-29.6	45.5
STN NH$_4^+$ (µg m^{-3})	1.34	1.05	0.5	-0.3	-21.5	38.7
IMPROVE EC (µg m^{-3})	0.23	0.16	0.6	-0.1	-30.7	48.3
STN EC (µg m^{-3})	0.65	0.38	0.2	-0.3	-42.0	52.8
IMPROVE OC (µg m^{-3})	1.10	1.88	0.2	0.8	71.7	134.6
IMPROVE TC (µg m^{-3})	1.33	2.05	0.2	0.7	53.9	116.3
STN TC (µg m^{-3})	4.42	2.42	0.1	-2.0	-45.3	69.7

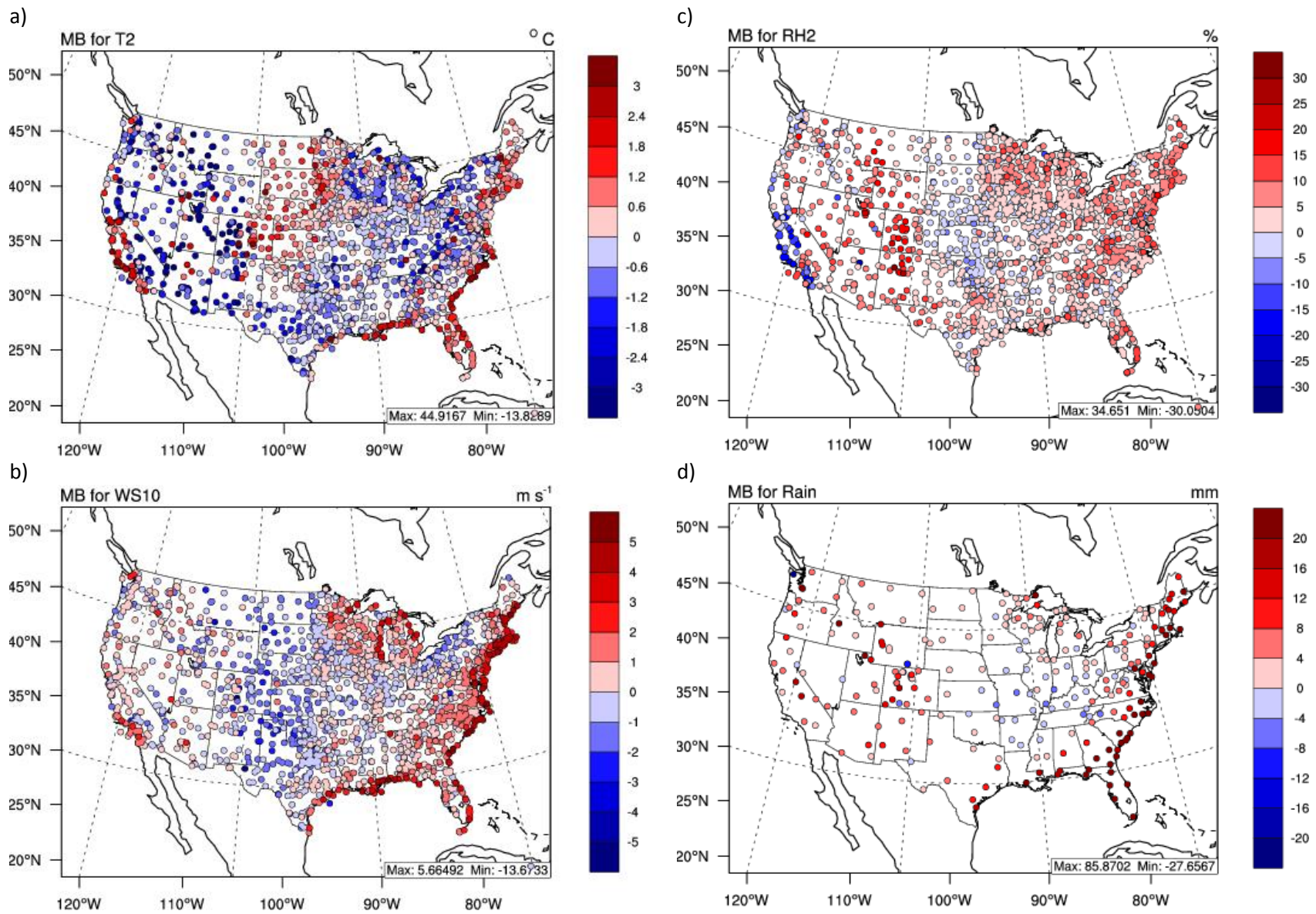[1] NH$_4^+$ IMPROVE data only available up to 2005.

Figure 1. Spatial distributions of MBs for: a) 2-m temperature (T2), b) 2-m relative humidity (RH2), c) 10-m wind speed (WS10) from NCDC, and d) weekly precipitation from NADP. Each marker represents the MB of each variable at each observational site.
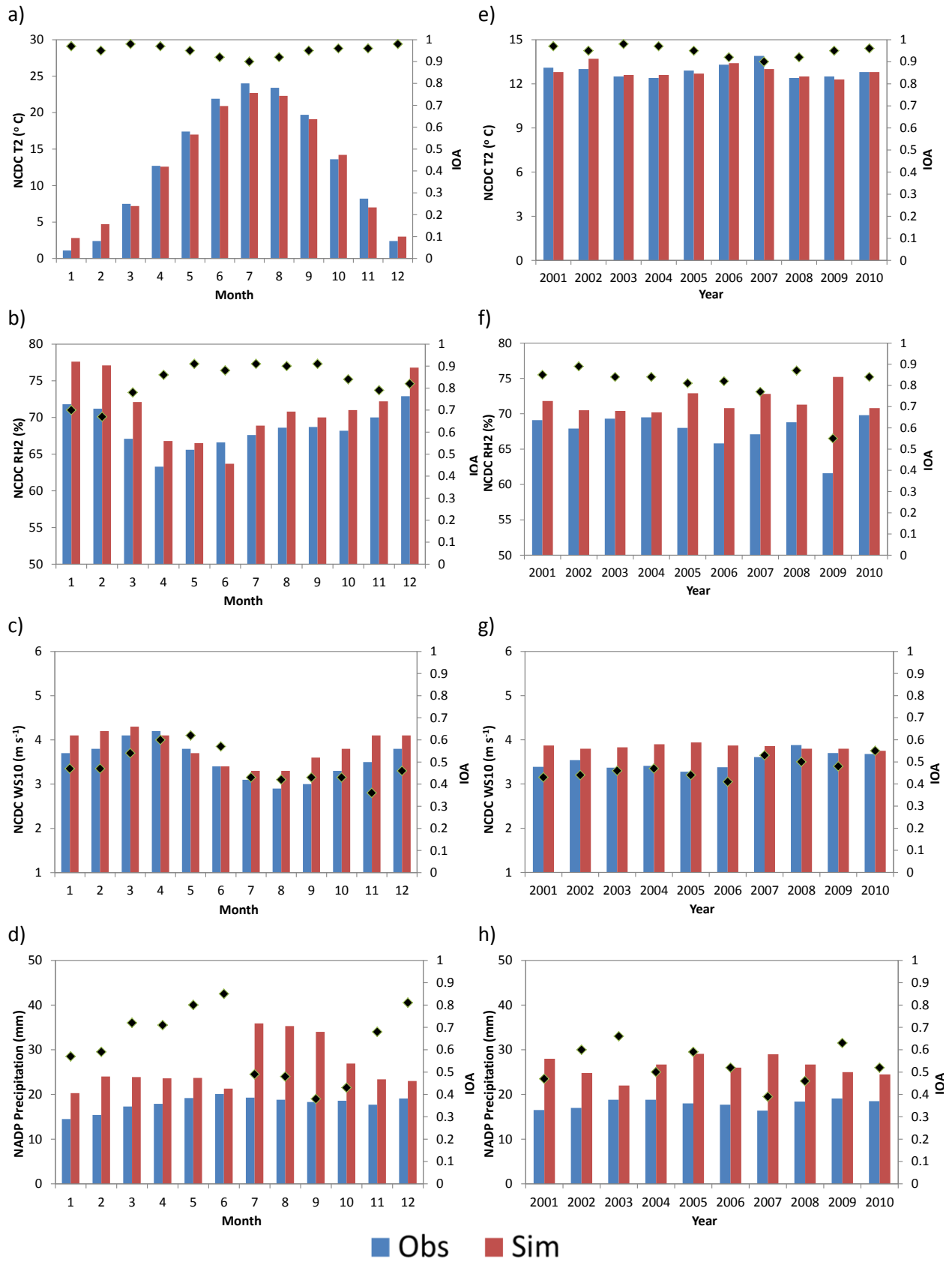
Figure 2. Time series of 10-year averaged monthly-mean observations (blue) versus simulations (red) for: a) T2, b) RH2, and c) WS10 against NCDC data, and d) precipitation against NADP data, and annual averages for e) T2, f) RH2, and g) WS10 against NCDC data, and h) precipitation against NADP. IOA statistics (black diamonds) are also provided on the secondary y-axes in panels a) – h).
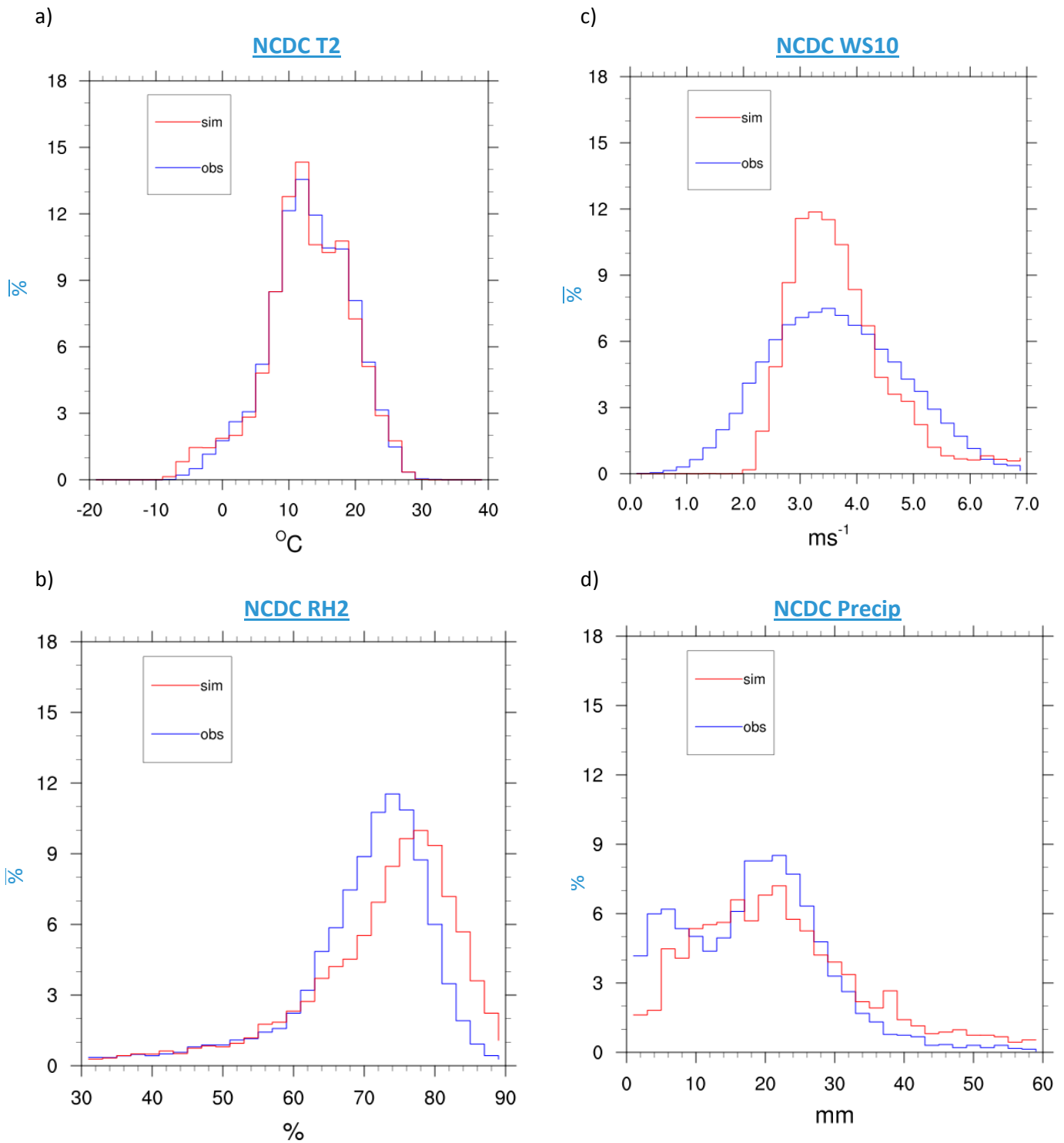
4

Figure 3. Probability distribution functions (PDFs)s of a) T2, b) RH2, c) WS10 against NCDC, and d) precipitation against NADP for 2001 to 2010 over 30 bins in the respective ranges of these variables. ~~The values for Y axis are in %.~~
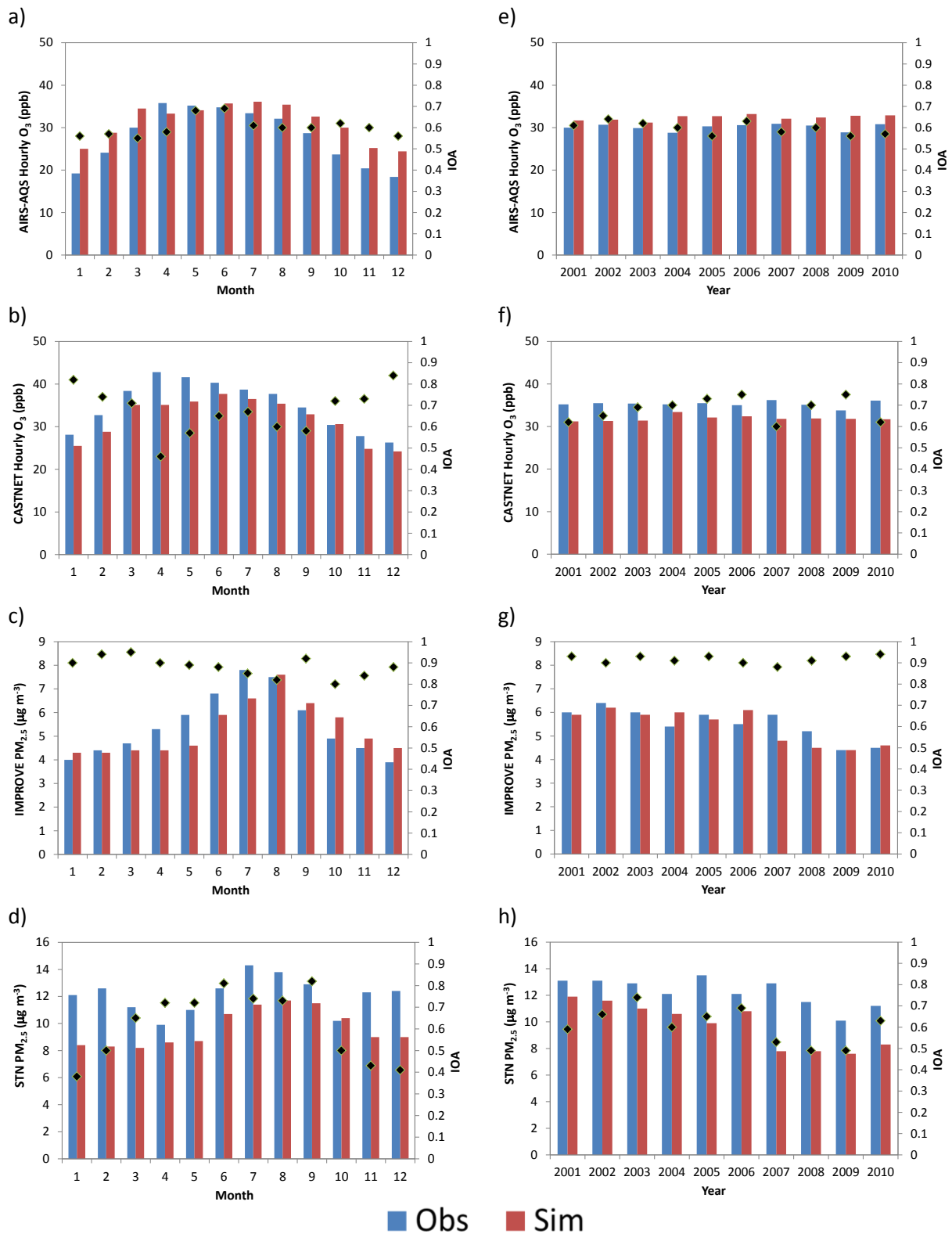
Figure 4. Time series of 10-year averaged monthly-mean observations (blue) versus simulations (red) for: a) O_3 against AIRS-AQS data, b) O_3 against CASTNET data, c) $PM_{2.5}$ against IMPROVE, and d) $PM_{2.5}$ against STN, and annual averages for e) O_3 against AIRS-AQS data, f) O_3 against CASTNET data, g) $PM_{2.5}$ against IMPROVE, and h) $PM_{2.5}$ against STN. IOA statistics (black diamonds) are also provided on the secondary y-axes in panels a) – h).
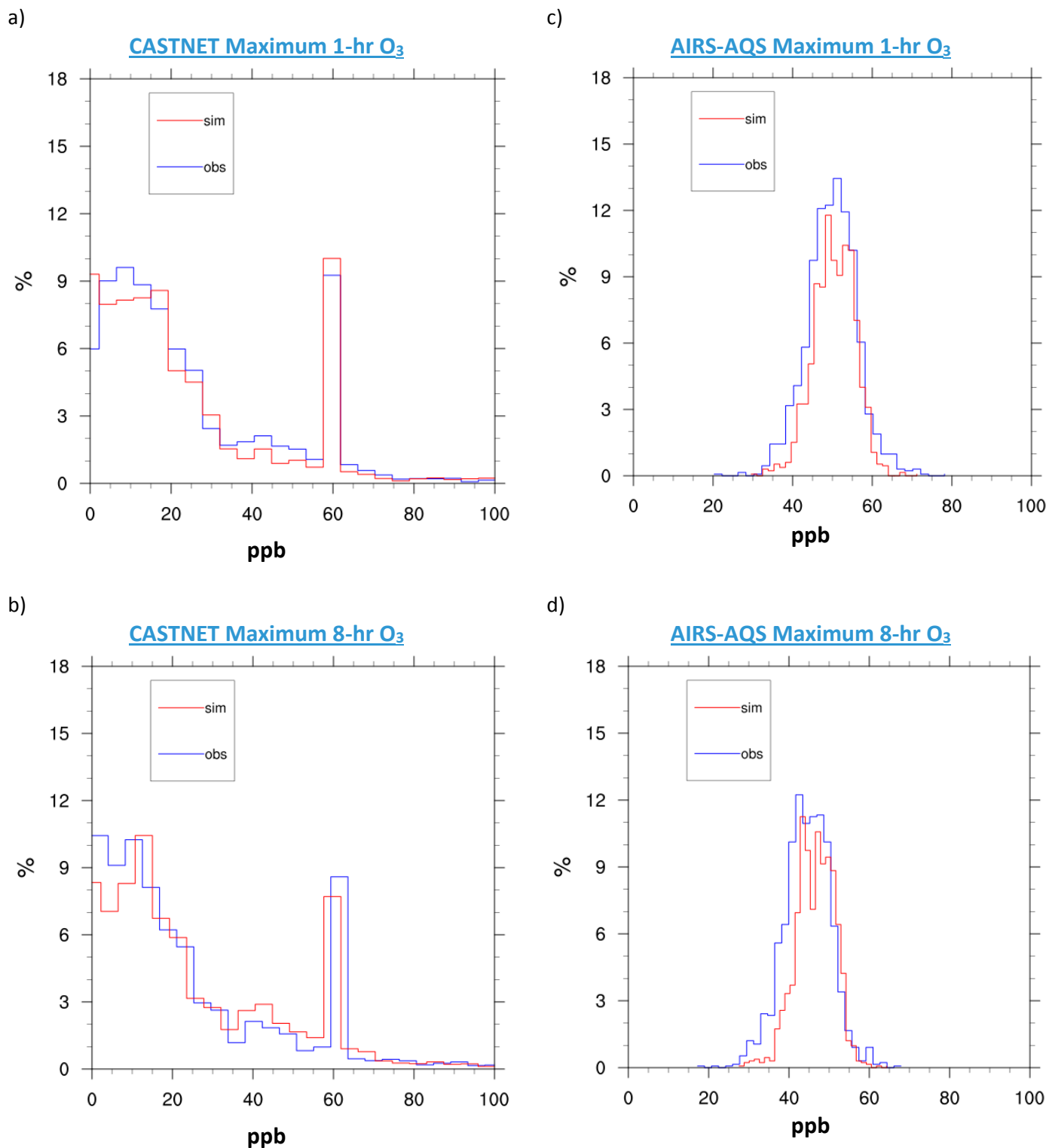
a)

CASTNET Maximum 1-hr O$_3$

c)

AIRS-AQS Maximum 1-hr O$_3$

b)

CASTNET Maximum 8-hr O$_3$

d)

AIRS-AQS Maximum 8-hr O$_3$

Figure 5. Probability ~~distributions~~ distribution functions (PDFs) of a) maximum 1-hr O$_3$ against CASTNET, b) maximum 8-hr O$_3$ against CASTNET, c) maximum 1-hr O$_3$ against AIRS-AQS, and d) maximum 8-hr O$_3$ against AIRS-AQS for 2001 to 2010~~.~~ over 30 bins in the respective ranges for all these variables.
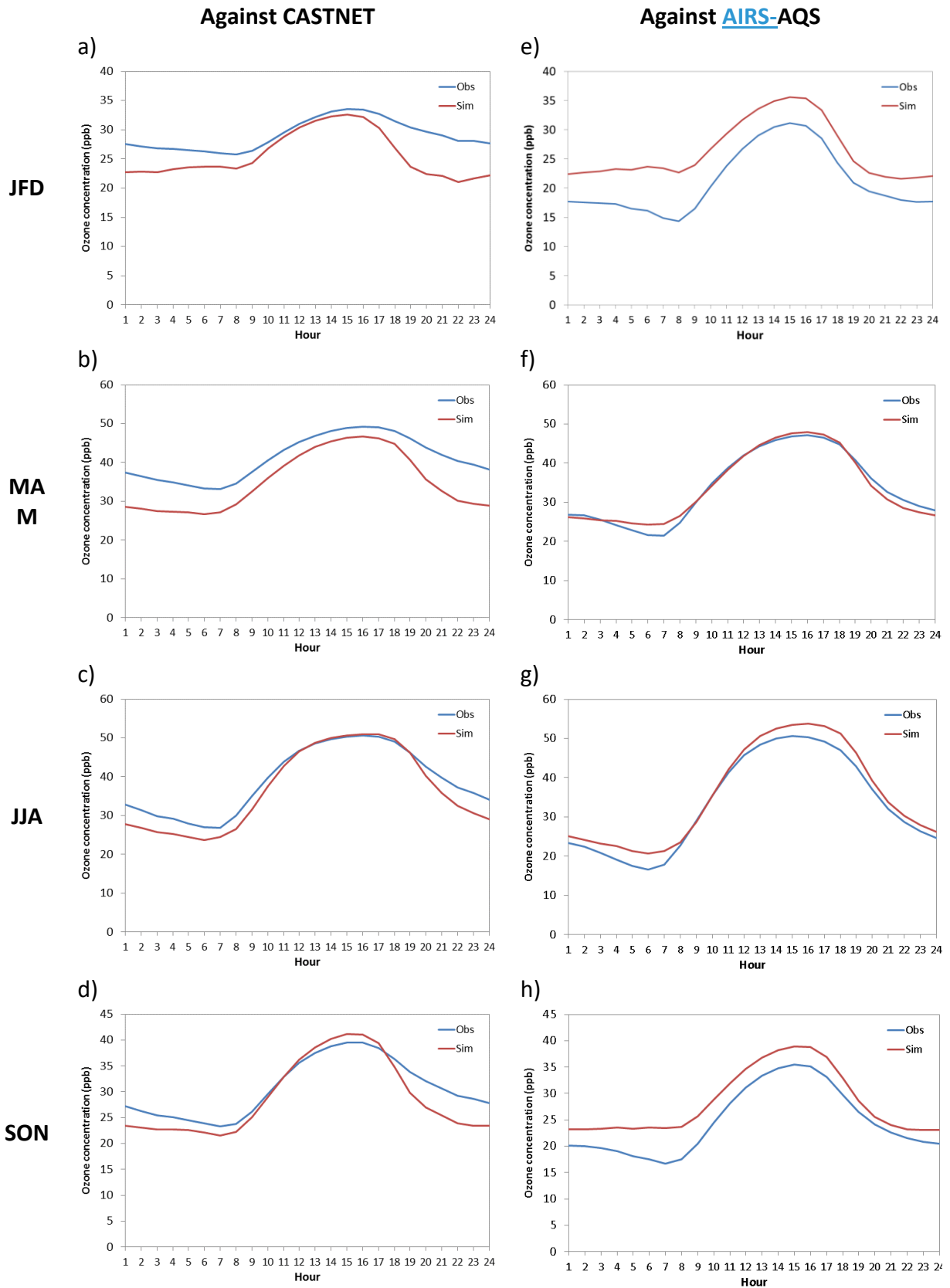
Figure 6. Diurnal variation of observed vs. simulated hourly O$_3$ concentrations against CASTNET (left column from a) to d)) and AIRS-AQS (right column from e) to h)) for all climatological seasons. The x-axes refer to hours in local standard time.
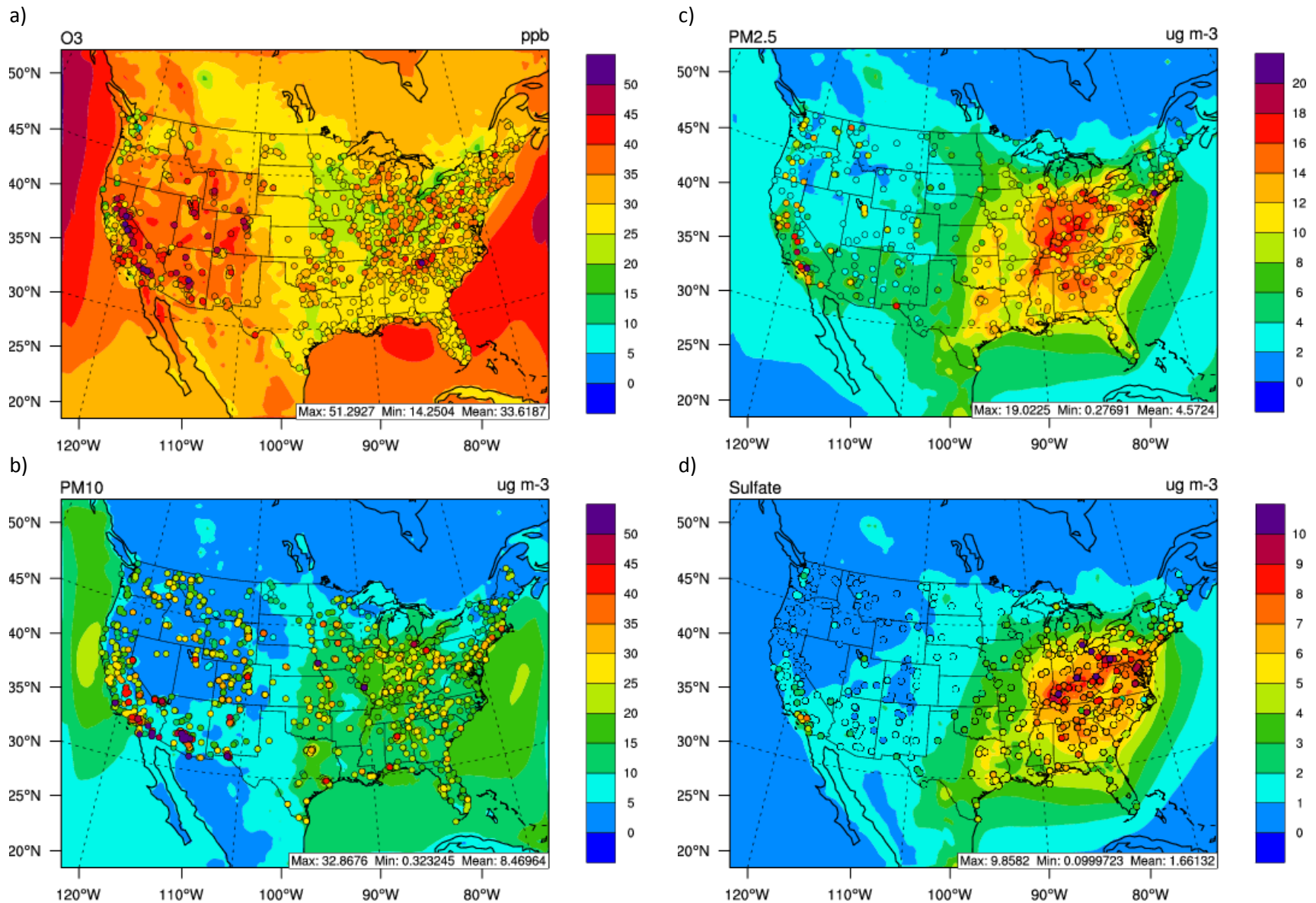
Figure 7. Spatial distributions of 10-year averaged hourly observed vs. simulated a) O_3 for CASTNET and AIRS-AQS, b) PM_{10} from AIRS-AQS, c) $PM_{2.5}$, and d) $PM_{2.5}$ sulfate from STN and IMPROVE. The background plots represent the simulated data while observations are represented by the markers.
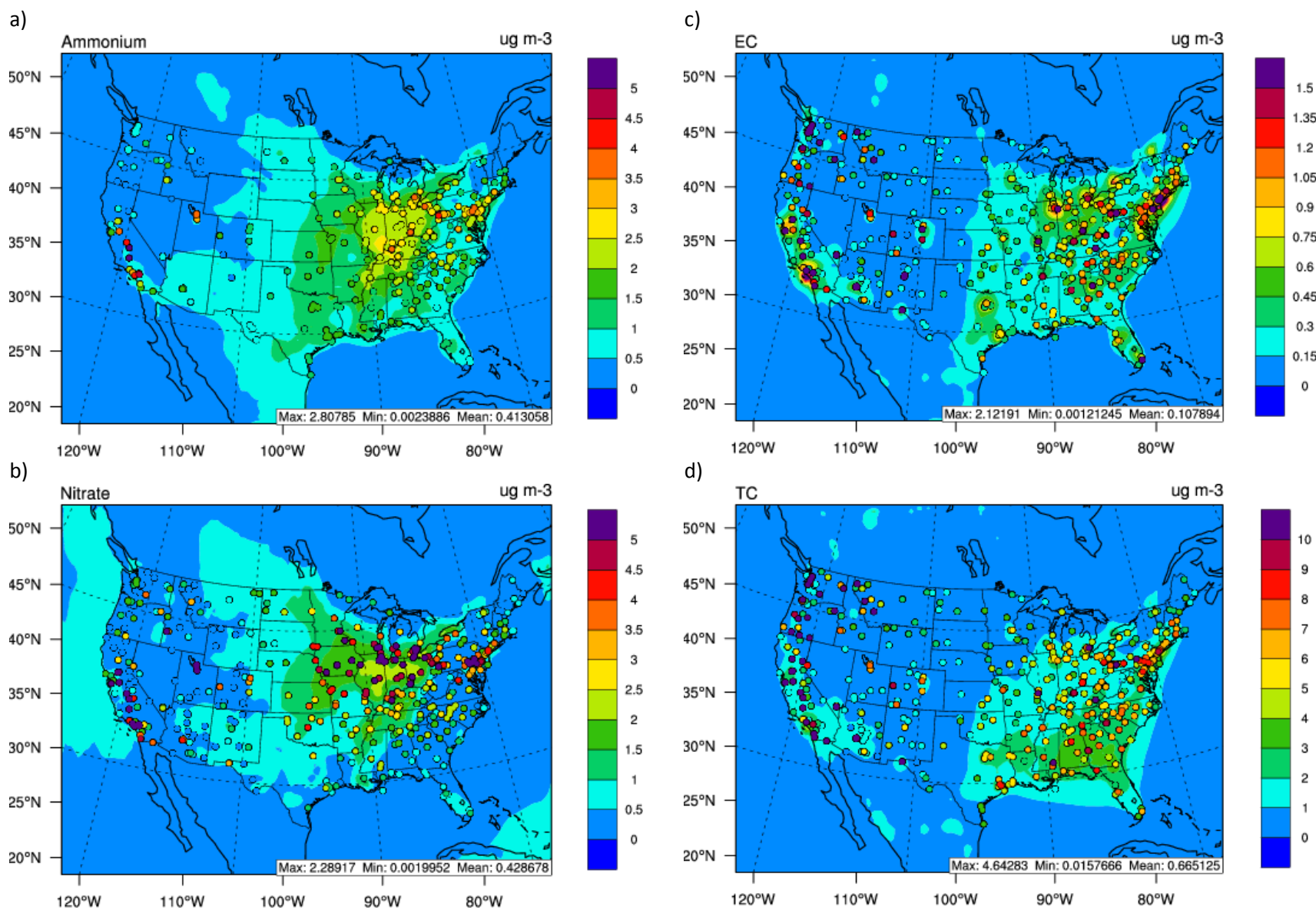
Figure 8. Spatial distributions of 10-year averaged hourly observed vs. simulated a) Ammonium, b) Nitrate, c) EC, and d) TC from STN and IMPROVE. The background plots represent the simulated data while observations are represented by the markers.
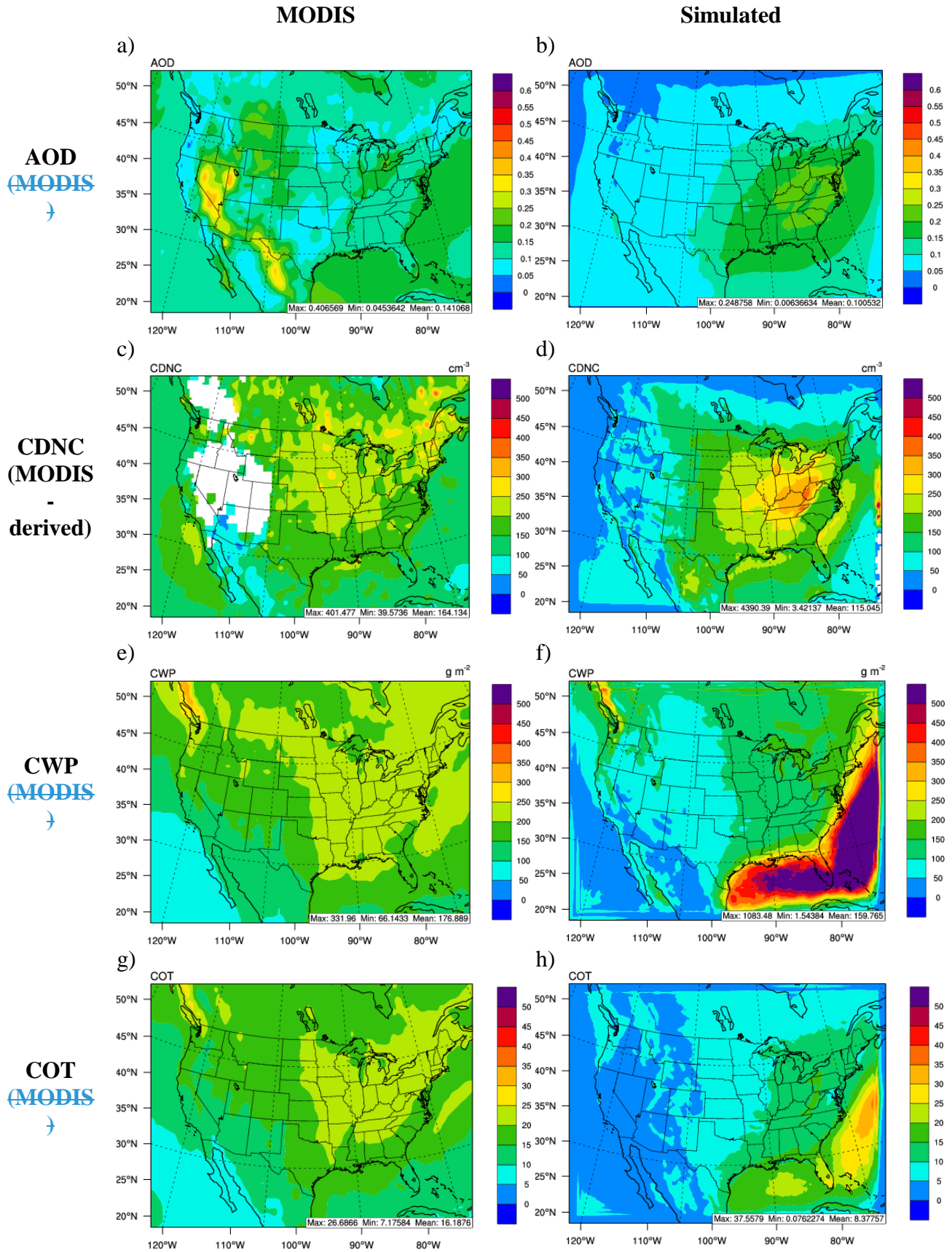
Figure 9. 10-year averaged MODIS (left) vs. simulated (right) AOD (a and b), CDNC (c and d), CWP (e and f), and COT (f and g).
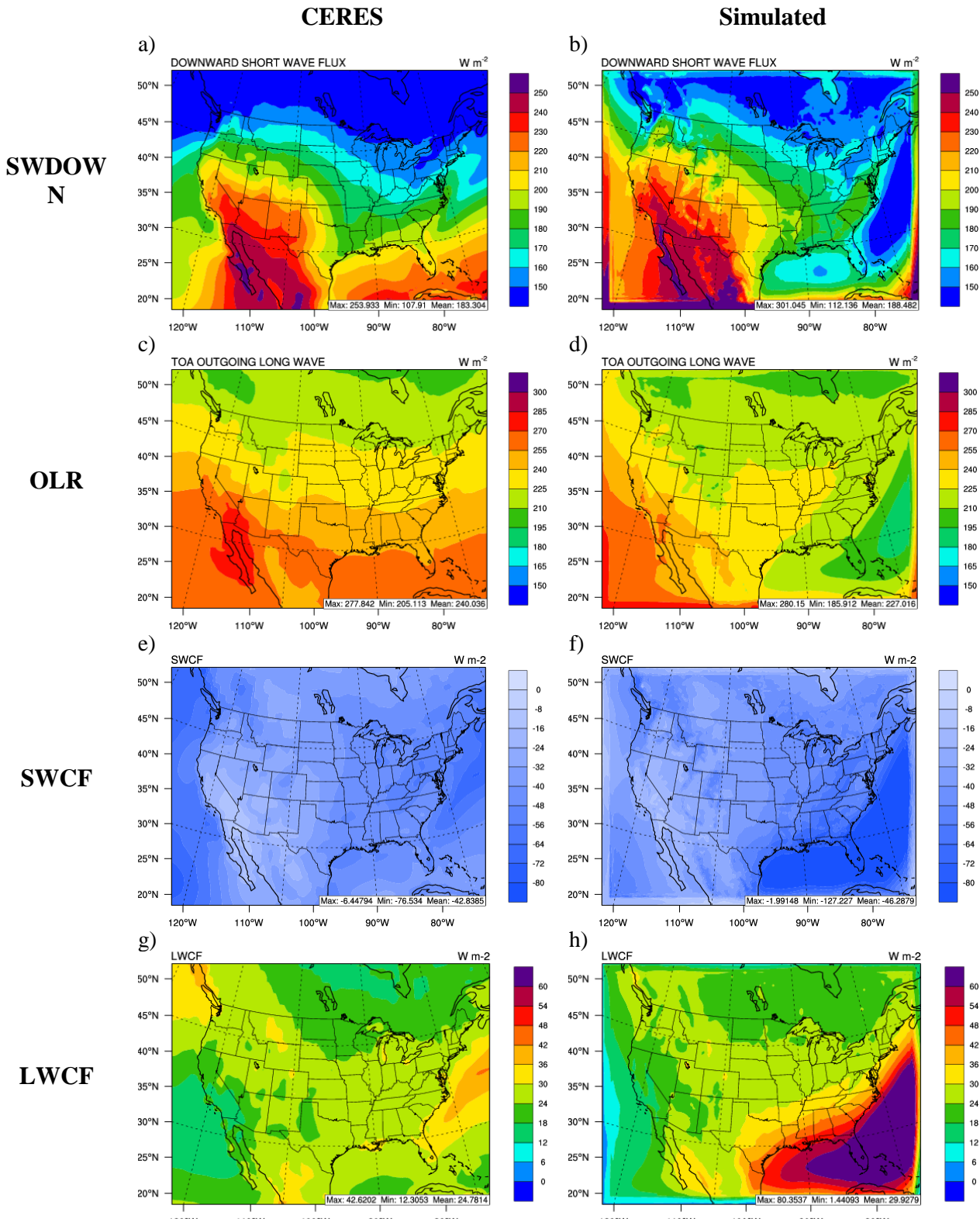
Figure 10. 10-year averaged CERES (left) vs. simulated (right) SWDOWN (a and b), OLR (c and d), SWCF (e and f), and LWCF (f and g).

Supplementary Material

Decadal Evaluation of Regional Climate, Air Quality, and Their Interactions using WRF/Chem

Khairunnisa Yahya, Kai Wang, Patrick Campbell, Timothy Glotfelty, Jian He and Yang Zhang*

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC 27695, U.S.A.

Email: *yang_zhang@ncsu.edu

1. List of Acronyms

Table S1. List of Acronyms used in the paper

Acronym	Full Name
AER/AFWA	The Atmospheric and Environmental Research Inc. and Air Force Weather Agency scheme
AERONET	The Aerosol Robotic Network
AIRS-AQS	the Aerometric Information Retrieval System– Air Quality System
AOD	Aerosol optical depth
BCs	Boundary Conditions
CAM5	The Community Atmosphere Model version 5
CASTNET	The Clean Air Status and Trends Network
CALIOP	The Cloud-Aerosol Lidar with Orthogonal Polarization
CB05	The Carbon Bond 2005
CCN	Cloud condensation nuclei
CDNC	Cloud droplet number concentration
CERES	The Clouds and the Earth's Radiant Energy System
CESM	The Community Earth System Model
CESM_NCSU	CESM/CAM5 developed at the North Carolina State University
CLDFRA	Cloud fraction
CMAQ	The Community Multiscale Air Quality Model
CMIP5	The Coupled Model Intercomparison Project Phase 5
CONUS	Continental U.S.
COT	Cloud optical thickness
CRU	Climatic Research Unit
CWP	Cloud water path
EC	Elemental carbon
GCMs	General circulation models
GCTMs	Global chemical transport models
GLW	Longwave radiation
GPCP	Global Precipitation Climatology Project

GSW	Net shortwave radiation
ICs	Initial Conditions
IMPROVE	The Interagency Monitoring of Protected Visual Environments
IOA	Index of Agreement
IPCC	The Intergovernmental Panel on Climate Change
JFD	January, February and December
JJA	June, July, and August
LSM	Land Surface Model
LST	local standard time
LWCF	Longwave cloud forcing
MADE/VBS	The Modal for Aerosol Dynamics in Europe / Volatility Basis Set
MAM	March, April, and May
MAN	The Maritime Aerosol Network
MB	Mean bias
MEGAN2	The Model of Emissions of Gases and Aerosols from Nature version 2
MODIS	The Moderate Resolution Imaging Spectroradiometer
MSKF	The Multi-Scale Kain-Fritsch cumulus scheme
NADP	The National Atmospheric Deposition Network
NARR	The North American Regional Reanalyses
NCDC	The National Climatic Data Center
NCEP	The National Centers for Environmental Prediction
NCEP FNL	The NCEP Final Reanalyses
NEI	The National Emission Inventory
NH_4^+	Ammonium
NMB	Normalized mean bias
NME	Normalized mean error
NO_3^-	Nitrate
NO	Nitric oxide
NO_2	Nitrogen dioxide
NO_x	Nitrogen oxide
NOAH	The National Center for Environmental Prediction, Oregon State University, Air Force, and Hydrologic Research Lab
O_3	Ozone
OA	Organic aerosol
OC	Organic carbon
OMI	The Ozone Monitoring Instrument
$PM_{2.5}$ and PM_{10}	Particulate matter with diameter less than and equal to 2.5 and 10 μm
POA	Primary organic aerosol
PRECIS	Providing Regional Climates for Impacts Studies
PRISM	The Parameter-elevation Regressions on Independent Slopes Model
R	Correlation coefficient
RCMs	Regional climate models
RCP	The Representative Concentration Pathway
RH2	Relative humidity at 2-m
RRTMG	The Rapid and accurate Radiative Transfer Model for GCM

SEARCH	The Southeastern Aerosol Research and Characterization
SMOKE	The Sparse Matrix Operator Kernel Emissions model
SOA	Secondary organic aerosol
SO_2	Sulfur dioxide
SO_4^{2-}	Sulfate
SON	September, October, and November
STN	The Speciated Trends Network
SWCF	Shortwave cloud forcing
SWDOWN	Downward shortwave radiation
T2	Temperature at 2-m
TC	Total carbon, = EC + OC
WD10	Wind direction at 10-m
WRF	Weather Research and Forecasting model
WRF/Chem	The Weather Research and Forecasting model with Chemistry
WS10	Wind speed at 10-m

2. Mapping of RCP Emissions to CB05 species

Table S2 summarizes the mapping of species from RCP emissions to CB05 species for input into the model. The explanation for the mapping process can be found in the main text.

Table S2. CB05 emissions species for WRF/Chem, their associated full names, their availability in regards to the RCP emissions dataset, and the lumped RCP group species.

CB05 Species WRF/Chem	Species Long name	RCP Species Available	RCP Group
E_ALD2	Acetaldehyde	Group	Other Alkanals
E_ALDX	Higher Aldehydes	Group	Hexanes and Higher Alkanes
E_BENZENE	Benzene	Yes	
E_CH4	Methane	Yes	
E_CL2	Chlorine	No	
E_CO	Carbon Monoxide	Yes	
E_ECI, E_ECJ, E_ECC	Elemental Carbon - Nuclei, Accumulation, Coarse Modes	No, Group, No	Black Carbon
E_ETH	Ethene	Yes	
E_ETHA	Ethane	Yes	
E_ETOH	Ethanol	Group	Alcohols
E_FORM	Formaldehyde	Yes	
E_HCL	Hydrogen Chloride	No	
E_HONO	Nitrous Acid	No	
E_IOLE	Internal Olefin Carbon Bond	Group	Other Alkenes and Alkynes
E_ISOP	Isoprene	No	
E_MEOH	Methanol	Group	Alcohols
E_NH3	Ammonia	Yes	
E_NH4I, E_NH4J	Ammonium – Nuclei, Accumulation Modes	No, No	
E_NO	Nitrogen Oxides	Yes	
E_NO2	Nitrogen Dioxide	No	
E_NO3I, E_NO3J, E_NO3C	Nitrate – Nuclei, Accumulation, Coarse Modes	No, No, No	
E_OLE	Terminal Olefin Carbon Bond	Group	Other Alkenes and Alkynes
E_ORGI, E_ORGJ, E_ORGC	Organics – Nuclei, Accumulation, Coarse Modes	No, Group, No	Organic Carbon
E_PAR	Paraffin Carbon Bond	No	
E_PM10	Unspeciated PM_{10}	No	
E_PM25	Unspeciated $PM_{2.5}$	No	
E_PM25I, E_PM25J	Unspeciated $PM_{2.5}$ – Nuclei, Accumulation Modes	No, No	
E_PSULF	Sulfuric Acid	No	
E_SO2	Sulfur Dioxide	Yes	
E_SO4I, E_SO4J, E_SO4C	Sulfate – Nuclei, Accumulation, Coarse Modes	No, No, No	
E_TERP	Terpene	No	
E_TOL	Toluene	Yes	
E_XYL	Xylene	Yes	

3. Observational Datasets for Model Evaluation and Operational Evaluation

Table S3 summarizes the observational databases and the variables evaluated in this work. For evaluation of chemical concentrations and meteorological variables, the surface networks include the National Climatic Data Center (NCDC) Quality Controlled Local Climatological Data (QCLCD), Clean Air Status and Trends Network (CASTNET), the Aerometric Information Retrieval System (AIRS) – Air Quality System (AQS), the Interagency Monitoring of Protected Visual Environments (IMPROVE), the Speciated Trends Network (STN), the Southeastern Aerosol Research and Characterization (SEARCH), and the National Atmospheric Deposition Network (NADP). Several aerosol-cloud-radiation variables are also evaluated against satellite retrievals including the Clouds and the Earth's Radiant Energy System (CERES) and the Moderate Resolution Imaging Spectroradiometer (MODIS).

NCDC QCLCD data contains data over 700 U.S. locations from July 1996 to December 2004, and over 1600 locations from 2005 onwards (http://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/quality-controlled-local-climatological-data-qclcd). CASTNET observations have been collected in a range of rural environments, from desert to agricultural locations, and from flat to complex terrains (http://java.epa.gov/castnet/epa_jsp/sites.jsp). It contains measurement data for meteorological variables and chemical concentrations. AIRS-AQS is the U.S. EPA's repository for ambient air quality data from over 5000 active monitors (http://www.epa.gov/ttn/airs/airsaqs/). While IMPROVE observations have been collected in protected visual environments, i.e., in National Parks and Wilderness Areas (http://vista.cira.colostate.edu/improve/), STN sites are located in a range of locations from urban to rural areas (http://www.epa.gov/ttnamti1/specgen.html). Both networks contain data for $PM_{2.5}$ and major $PM_{2.5}$ species. NADP contains precipitation data from rain gauges.

Table S3. Observational datasets and variables evaluated in this study.

Gases and PM Species			
Observational Database	**Variables Evaluated**	**Sampling Frequency**	**Number of Sites**
CASTNET	Max 1-hr and 8-hr O_3	Daily for O_3	~90
AIRS–AQS	O_3	Hourly	~1150
IMPROVE	$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, EC, OC	24-hour data. Data availability once every 3 days	~160
STN	$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, EC, TC	24-hour data. Data availability once every 3 days	~200
Meteorology			
Observational Database	*Variables evaluated*	*Temporal Resolution*	*Spatial Resolution*
NCDC QCLCD	T2, RH, WS10,WD10	Hourly	~700 before 2005 ~1600 after 2005
NADP	Precipitation	Weekly	255
Radiation and other Aerosol/Cloud variables			
Observational Database/ Satellite	*Variables evaluated*	*Temporal Resolution*	*Number of sites/ Spatial Resolution*
CERES	SWDOWN	Monthly	$1^o \times 1^o$
MODIS	AOD, CF, COT, CWP, QVAPOR, CCN	Monthly	$1^o \times 1^o$
MODIS derived based on Bennartz (2007)	CDNC	Monthly	$1^o \times 1^o$

4. Sensitivity simulations to determine precipitation and cloud bias over the Atlantic Ocean

A number of sensitivity simulations were conducted for the month of July 2005 to determine the cause of the precipitation bias, especially over the Atlantic Ocean. The sensitivity simulations consist of (i) **Base**, which is the set-up for the main simulations in this study consisting of monthly reinitialization frequency with CESM_NCSU ICs/BCs with the Grell 3D cumulus parameterization scheme; (ii) **Sen1**, which is similar to the Base case except with a 5-day reinitialization period; (iii) **Sen2**, which is similar to Base except using NCEP for the meteorological ICs/BCs; and (iv) **Sen3**, which is similar to Base except using WRF/Chem v3.7 with the MSKF cumulus parameterization, instead of Grell 3D. An additional sensitivity simulations using WRF/Chem v3.7 with both MSKF and Grell 3D and their comparison with Figure S1 showed that the differences between Sen3 and Base are mainly caused by the use of different cumulus parameterizations; other model updates between WRF/Chem v3.7 and WRF/Chem v3.6.1 only have minor contributions to such differences. A summary of the set-up of the sensitivity simulations can be found in Table S4.

The sensitivity simulations are evaluated against GPCP and PRISM data and the statistics are summarized in Tables S5 and S6, respectively. GPCP has data over the land and ocean while PRISM only has data over land. The results show that the R value for the **Base** case is the highest against both GPCP and PRISM, even though the NMB is the highest. While using more frequent reinitialization with 5-day (Sen1) reduces both the NMB and NME with slight to moderate improvements, it also reduces the R value. Using NCEP data as ICs/BCs (Sen2) also slightly-to-moderately improve the NMB and NME, indicating that using CESM_NCSU ICs/BCs contributes to the biases in precipitation. However, NCEP data are not available for future climate simulations.

Lastly, using CESM_NCSU IC/BCs with the new Multi-Scale Kain Fritsch (MSKF) scheme (Sen3) drastically reduce NMB and NME, but the correlation becomes much worse.

Table S4. Summary of set-up of sensitivity simulations

No.	Sensitivity Simulation	Reinitialization Frequency	IC/BCs	Cumulus Parameterization Scheme
1.	Base	Monthly	CESM_NCSU	Grell 3D
2.	Sen1	5-day	CESM_NCSU	Grell 3D
3.	Sen2	Monthly	NCEP	Grell 3D
4	Sen3	Monthly	CESM_NCSU	MSKF

Table S5. Statistics for sensitivity simulations against GPCP

Sensitivity Simulation	Mean Obs (mm)	Mean Sim (mm)	R	NMB (%)	NME (%)
Base	2.4	5.3	0.5	121.1	150.2
Sen1	2.4	4.2	0.4	74.1	140.9
Sen2	2.4	4.5	0.5	85.1	122.4
Sen3	2.4	2.9	0.1	18.9	109.2

Table S6. Statistics for sensitivity simulations against PRISM

Sensitivity Simulation	Mean Obs (mm)	Mean Sim (mm)	R	NMB (%)	NME (%)
Base	2.3	4.0	0.7	77.8	96.5
Sen1	2.3	2.5	0.3	11.5	102.8
Sen2	2.3	3.6	0.5	60.9	105.0
Sen3	2.3	2.2	-0.2	-2.1	111.9

Figure S1 compares the spatial plots of the simulated precipitation with daily average observational precipitation data from GPCP and PRISM for July 2005. The high precipitation over the Atlantic ocean shown in all sensitivity simulations particularly in Sen1 and Sen2 does not exist in the GPCP observational data. The 5-day reinitialization case (Sen1) does not help to reduce the high precipitation over the ocean. Using NCEP data (Sen2) helps to reduce the precipitation over the ocean slightly. Using the MSKF scheme (Sen3) completely reduces the precipitation over the

ocean, however it does not capture well precipitation over the southeastern U.S. The comparison of Sen3 and Base illustrates a very high sensitivity of the simulated precipitation to different cumulus parameterizations, which warrants future study.
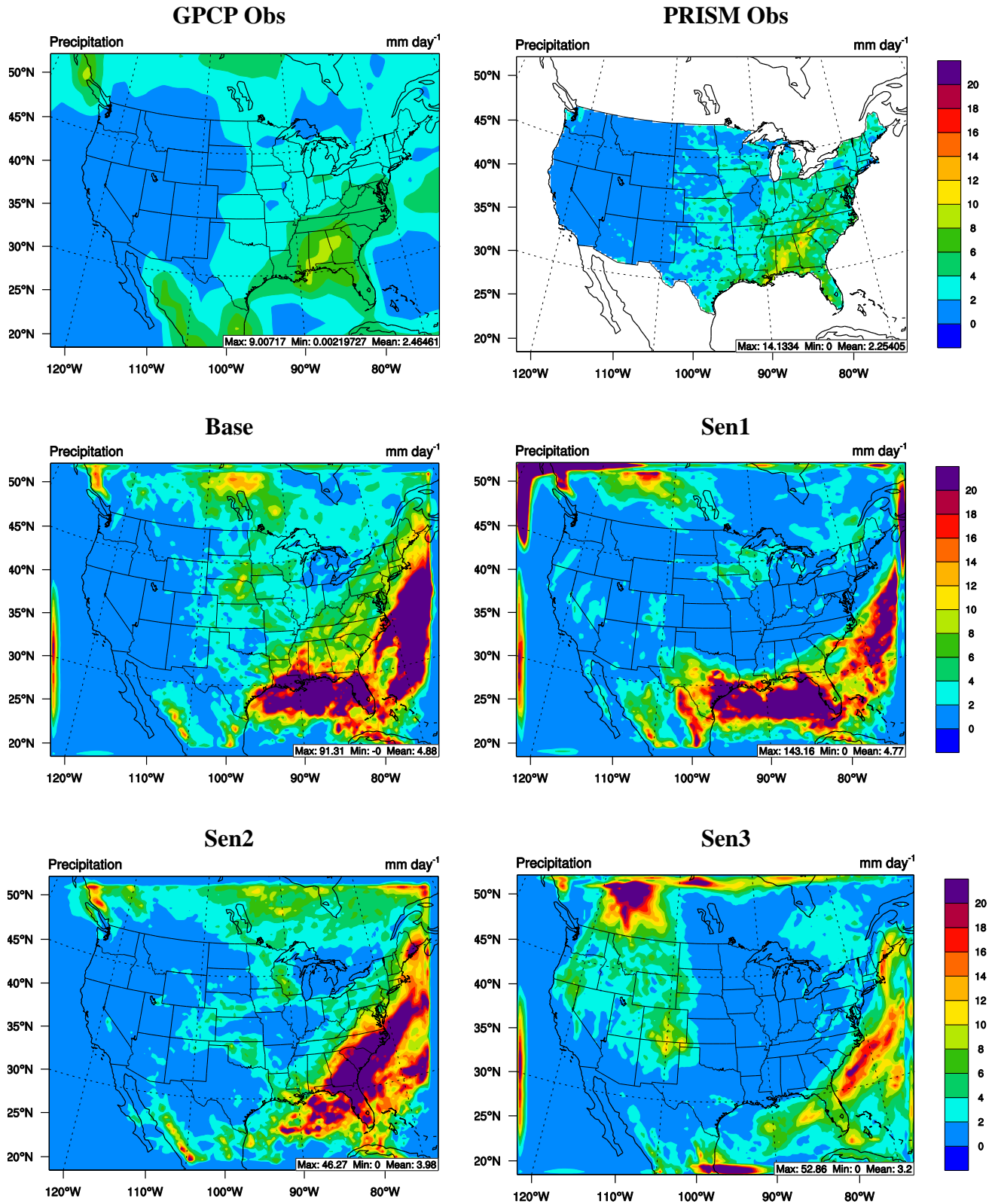
Figure S1. Spatial plots of average daily precipitation for GPCP and PRISM and sensitivity simulation cases for July 2005.

References

Bennartz, R. (2007), Global assessment of marine boundary layer cloud droplet number concentration from satellite, J. Geophys. Res., 112, D02201, doi:10.1029/2006JD007547.