



1       **Evaluation of the NAQFC Driven by the NOAA Global Forecast**  
2       **System Version 16: Comparison with the WRF-CMAQ Downscaling**  
3       **Method During the Summer 2019 FIREX-AQ Campaign**

4       Youhua Tang<sup>1,2</sup>, Patrick C. Campbell<sup>1,2</sup>, Pius Lee<sup>1</sup>, Rick Saylor<sup>1</sup>, Fanglin Yang<sup>3</sup>, Barry Baker<sup>1</sup>,  
5       Daniel Tong<sup>1,2</sup>, Ariel Stein<sup>1</sup>, Jianping Huang<sup>3,4</sup>, Ho-Chun Huang<sup>3,4</sup>, Li Pan<sup>3,4</sup>, Jeff McQueen<sup>3</sup>,  
6       Ivanka Stajner<sup>3</sup>, Jose Tirado-Delgado<sup>5,6</sup>, Youngsun Jung<sup>5</sup>, Melissa Yang<sup>7</sup>, Ilann Bourgeois<sup>8,9</sup>,  
7       Jeff Peischl<sup>8,9</sup>, Tom Ryerson<sup>9</sup>, Donald Blake<sup>10</sup>, Joshua Schwarz<sup>9</sup>,  
8       Jose-Luis Jimenez<sup>8</sup>, James Crawford<sup>11</sup>, Glenn Diskin<sup>7</sup>, Richard Moore<sup>7</sup>, Johnathan Hair<sup>7</sup>, Greg  
9       Huey<sup>11</sup>, Andrew Rollins<sup>9</sup>, Jack Dibb<sup>12</sup>, Xiaoyang Zhang<sup>13</sup>

10      1. NOAA Air Resources Laboratory, College Park, MD, USA.

11      2. Center for Spatial Information Science and Systems, George Mason University, Fairfax, VA,  
12      USA.

13      3. NOAA National Centers for Environmental Prediction, College Park, MD, USA

14      4. I.M. Systems Group Inc., Rockville, MD, USA

15      5. Office of Science and Technology Integration, NOAA National Weather Service, Silver  
16      Spring, MD, USA

17      6. Eastern Research Group Inc, USA

18      7. NASA Langley Research Center, Hampton, VA, USA

19      8. Cooperative Institute for Research in Environmental Sciences, University of Colorado  
20      Boulder, Boulder, CO, USA

21      9. NOAA Chemical Sciences Laboratory, Boulder, CO, USA

22      10. Department of Chemistry, University of California at Irvine, Irvine, CA, USA

23      11. School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA,  
24      USA

25      12. Earth Systems Research Center, University of New Hampshire, Durham, NH, USA

26      13. Department of Geography & Geospatial Sciences, South Dakota State  
27      University, Brookings, SD, USA

28

29      **Correspondence:** Youhua Tang (youhua.tang@noaa.gov)

30      **Abstract**

31      The latest operational National Air Quality Forecasting Capability (NAQFC) has been advanced  
32      to use the Community Multi-scale Air Quality (CMAQ) model version 5.3.1 with CB6 (carbon  
33      bond version 6)-Aero7 (version 7 of the aerosol module) chemical mechanism and is driven by  
34      the Finite Volume Cubed-Sphere (FV3)-Global Forecast System, version 16 (GFSv16). This has  
35      been accomplished by development of the meteorological preprocessor, NOAA-EPA  
36      Atmosphere-Chemistry Coupler (NACC), which is adapted from the existing Meteorology-  
37      Chemistry Interface Processor (MCIP). Differing from the typically used Weather Research and  
38      Forecasting (WRF)/CMAQ system in the air quality research community, the interpolation-based  
39      NACC can use various meteorological output to drive CMAQ (e.g., FV3-GFSv16) even though



1 they are in different grids. Here we compare and evaluate GFSv16-CMAQ vs. WRFv4.0.3-  
2 CMAQ using observations over the contiguous United States (CONUS) in summer 2019. During  
3 this period, the Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ)  
4 field campaign was performed and we compare the two models with airborne measurements  
5 mainly from the NASA DC-8 aircraft. The GFS-CMAQ and WRF-CMAQ systems have overall  
6 similar performance with some differences for certain events, species and regions. The GFSv16  
7 meteorology tends to have stronger planetary boundary layer height diurnal variability (higher  
8 during daytime, and lower at night) than WRF over the U.S. Pacific coast, and it also predicted  
9 lower nighttime 10-m winds. In summer 2019, GFS-CMAQ system showed better surface O<sub>3</sub>  
10 than WRF-CMAQ at night over the CONUS domain; however, their PM<sub>2.5</sub> predictions showed  
11 mixed verification results: GFS-CMAQ yielded better mean bias but poorer correlations over the  
12 Pacific coast. These results indicate that using global GFSv16 meteorology with NACC to  
13 directly drive CMAQ via the interpolation is feasible and yields reasonable results compared to  
14 the commonly-used WRF downscaling approach.  
15

## 16 **1. Introduction**

17 Traditionally, mesoscale meteorological models such as the Weather Research and Forecasting  
18 Model (WRF) (Powers et al., 2017) are used as the meteorological drivers for air quality models  
19 (AQMs) on the same (“native”) model grid, such as Community Multiscale Air Quality Model  
20 (CMAQ) (Byun & Schere, 2006). The NOAA National Weather Service’s (NWS) National Air  
21 Quality Forecasting Capability (NAQFC) has historically used a different approach, in which the  
22 hourly meteorological outputs from prior operational models, such as North American Mesoscale  
23 Model (NAM), need to be interpolated to the AQM grid to drive its air quality prediction. Prior  
24 to this work, a “PREMAQ” coupler (Otte et al, 2004) combined both meteorological processing  
25 and Sparse Matrix Operator Kernel Emissions (SMOKE) (Houyoux et al., 2000) processes, such  
26 as point source plume rise effects. However, since the release of CMAQ version 5, the  
27 meteorology-dependent plume rise, sea salt, and dust emission processes are included as inline  
28 modules in CMAQ, and thus the corresponding emission processes are no longer needed in  
29 PREMAQ. Furthermore, PREMAQ has no built-in interpolator, and thus relied on external  
30 interpolators to remap the non-native-grid meteorological inputs, such as NAM, to the targeted  
31 CMAQ domain, though it did perform vertical layer collapsing/interpolation to reduce layers.  
32 The interpolation approach allows for more flexibility in using different meteorological drivers  
33 (i.e., besides just WRF) for CMAQ; however, there is potential to raise mass-consistency issues  
34 between models. It should be noted that the mass-consistency issues may also exist using native-  
35 grid couplers (Byun, 1999a, 1999b), which can stem from the mass-inconsistent meteorological  
36 inputs or due to the temporal interpolation of the input data. The well-developed offline AQMs,  
37 such as CMAQ, have already considered such mass-consistency treatments using different  
38 meteorological inputs (Byun et al., 1999c).  
39

40 To upgrade the NAQFC system with the latest CMAQ chemistry and NOAA operational  
41 meteorology, we developed an updated interpolation-based meteorological coupler, the NOAA-  
42 EPA Atmosphere-Chemistry Coupler (NACC) (Campbell et al., 2022) adapted from the U.S.



1 EPA’s Meteorology-Chemistry Interface Processor (MCIP) version 5 (Otte and Pleim, 2010;  
2 <https://github.com/USEPA/CMAQ>). The NACC system replaced PREMAQ, and effectively  
3 couples the Finite-Volume Cubed-Sphere (FV3) Dynamical Core - Global Forecast System  
4 version 16 (GFSv16) (Yang et al., 2020; Harris et al, 2021) to CMAQ v5.3.1 (hereafter referred  
5 to as GFS-CMAQ). Campbell et al. (2022) described the development and application of the  
6 GFS-CMAQ system using NACC (in their work referred to as “NACC-CMAQ”) and a  
7 comprehensive comparison between the current (GFS-CMAQ since July 20, 2021) and previous  
8 (NAM-CMAQv5.0.2) operational NAQFC model performances.

9  
10 In this study, we analyze the impacts of the meteorological model drivers, and compare GFS-  
11 CMAQ using NACC interpolation to the commonly-used downscaled, native-grid WRF-CMAQ  
12 application and its impact on air quality predictions. Yu et al. (2012a, 2012b) had previously  
13 compared the CMAQ performance driven by WRF-NMM and WRF-ARW during the 2006  
14 TexAQS/GoMACCS field campaign, and found that the NMM-CMAQ and ARW-CMAQ  
15 showed overall similar performance with some differences for certain events, chemical species,  
16 and regions. Similarly, this study focuses on the comparison of GFS-CMAQ versus WRF-  
17 CMAQ (see Section 2, Methodology), and verifies the model performance against the aircraft  
18 observations from the Fire Influence on Regional to Global Environments and Air Quality  
19 (FIREX-AQ) field experiment during summer 2019 (Section 3). Surface verification is also  
20 performed using AIRNow data for August 2019 (Section 4), serving as a benchmark for the new  
21 NAQFC versus the traditional WRF-CMAQ used in the air quality modeling community.

## 23 **2. Methodology**

24 Here we compare the two CMAQ (version 5.3.1) runs driven by the interpolated GFSv16  
25 meteorology (GFS-CMAQ) and WRF downscaled meteorology (WRF-CMAQ). All other  
26 settings, such as emission and lateral boundary conditions are the same. The meteorology-  
27 related physics is discussed in the following sections to address the models’ performance  
28 discrepancies. Both the GFS-CMAQ and WRF-CMAQ simulations are run from a period  
29 covering 12 July – 31 August, 2019, each using the last 10 days in July as the model spin-ups  
30 that are not included in the analyses.

31

### 32 **2.1 GFS Meteorological Inputs**

33 The GFSv16 is the current operational global forecast system in NOAA/NCEP using FV3  
34 dynamical core. Its detailed configuration can be found in Campbell et al. (2022) and Yang et al.  
35 (2020). Compared to the previous version (v15), GFSv16 updated many physical schemes (Table  
36 1) and added the parameterization for subgrid scale nonstationary gravity-wave drag. To use the  
37 GFS’s meteorology to drive CMAQ, a meteorological coupler, NACC, is developed (Campbell  
38 et al., 2022). Differing from the original MCIP which was developed to process WRF/ARW  
39 meteorology for CMAQ, the NACC coupler interpolates non-native-grid meteorology to a user-



1 defined grid and has parallel processing capability, which drastically reduces its run time for  
2 operational forecasts (Campbell et al., 2022). Currently, NACC's horizontal interpolation  
3 employs two methods: bilinear and nearest-neighbor. In this study, we use the nearest-neighbor  
4 method to categorical (discontinuous) variables that include land use types, vegetation fraction,  
5 terrain elevation, Monin–Obukhov length, friction velocity, and soil temperatures, while the  
6 bilinear interpolation is used for mainly smoothly varying (continuous) meteorological variables  
7 that include wind fields, temperature, pressure, and specific humidity. The CMAQ model is  
8 defined in the Arakawa C-grid (Arakawa and Lamb, 1977), and thus the GFSv16 horizontal wind  
9 components (U, V) need to be interpolated to the perpendicular cell faces instead of the cell  
10 center (Otte and Pleim, 2010) after rotation to the defined map projection. The scalar variables  
11 are defined in the target grid cell center, and thus their GFSv16 interpolations are more  
12 straightforward. The NACC coupler can either use the native layers or collapse (i.e., interpolate)  
13 to a set number of user-defined vertical layers for CMAQ use. The GFSv16 has 127 vertical  
14 layers with global coverage in 13 km horizontal resolution, where the targeted domain is a 12×12  
15 km Contiguous United States (CONUS) with 35 vertical layers (Campbell et al, 2022). Here we  
16 use 24-hour GFSv16 forecasts starting at 12 UTC each day.

17  
18 Most variables needed by CMAQ are directly interpolated from the GFSv16 outputs. The NACC  
19 processor has options to calculate diagnostic variables, such as planetary boundary layer (PBL)  
20 height, if they are needed. In this study, we use the interpolated GFSv16's PBL height instead of  
21 the diagnostic one. It also has an option to import the externally provided land-surface variables.  
22 Here we import updated 2018–2020 climatological averaged leaf area index (LAI) and NOAA  
23 near-real-time (NRT) greenness vegetation fraction (GVF) from satellite-based Visible Infrared  
24 Imaging Radiometer Suite (VIIRS) retrievals (Campbell et al., 2022). The updated satellite-based  
25 LAI and GVF impact CMAQ's biogenic emissions and dry deposition processes, which were  
26 described in detail in Campbell et al. (2022).

27

## 28 **2.2 WRF Meteorology**

29 To compare with GFSv16 meteorology processed by NACC, a corresponding WRF version 4.0.3  
30 (Skamarock et al, 2021) simulation is run covering the NAQFC's native grid, which is a 12 km  
31 horizontal resolution, Lambert conformal map projection over CONUS. Table 1 shows the WRF  
32 configuration, which is commonly employed in CONUS meteorological and air quality studies in  
33 the community, versus the current NOAA/NWS operational version of GFSv16. In contrast to  
34 GFSv16, which is a global model that uses the NOAA/NCEP's Global Data Assimilation System  
35 (GDAS) ([https://www.emc.ncep.noaa.gov/data\\_assimilation/data.html](https://www.emc.ncep.noaa.gov/data_assimilation/data.html)) for its initial conditions  
36 and runs on its own global dynamics and physics without any other constraints, the regional  
37 WRF simulation uses downscaled GFSv16 for its initial conditions. Furthermore, WRF also uses  
38 downscaled lateral boundary conditions taken from GFSv16 every 6 hours. Here WRF runs  
39 continuously after spin-up and we have enabled the four-dimensional data assimilation (FDDA)



1 for the u- and v-component winds, temperature, and humidity (Table 1) every 6 hours, thus  
2 nudging towards GFSv16.  
3  
4 WRF and GFSv16 have similar settings for the land surface model, surface layer and radiation  
5 schemes; however, their microphysics and PBL schemes are different (Table 1). Compared to the  
6 35-layer WRF with a 100 hPa domain top, GFSv16 has a much higher domain top (0.2 hPa) and  
7 127 vertical layers, which are collapsed by NACC to 35 sigma layers up to 14 km for CMAQ.  
8 We use NACC (inherited from MCIP version 5.0) to process WRF hourly meteorology, while  
9 maintaining the vertical layer structure. Thus, in contrast to GFS-CMAQ, the WRF-CMAQ  
10 system uses the native grid without interpolation.

11

### 12 **2.3 CMAQ Configuration**

13 Here CMAQ version 5.3.1 (Appel et al., 2021) is used with the Carbon Bond 6 version r3  
14 (CB6r3; Yarwood et al., 2010, 2014; Luecken, et al., 2019) chemical mechanism and Aero7  
15 treatment of secondary organic aerosols (CB6r3\_AE7\_AQ). CMAQ 5.3.1 includes a series of  
16 scientific updates from the previous version (Appel et al., 2021), including the updated air-  
17 surface exchange and deposition modules, which showed significant impact on ozone prediction  
18 compared to the previous NAQFC (Campbell et al., 2022). We also include the bi-directional  
19 NH<sub>3</sub> (BIDI-NH<sub>3</sub>) exchange model for NH<sub>3</sub> surface fluxes. An updated Biogenic Emissions  
20 Landuse Dataset v5 (BELD5) is used in this study to drive the inline Biogenic Emissions  
21 Inventory System (BEIS) version 3.61. The anthropogenic emissions are provided by the  
22 National Emissions Inventory Collaborative (NEIC) with base year 2016 version 1 (NEIC 2019).  
23 We replace the U.S. EPA default CMAQ dust emissions model with a novel inline windblown  
24 dust model known as “FENGSHA” (Fu et al., 2014; Huang et al., 2015; Dong et al., 2016).  
25 Campbell et al. (2022) include the details of the CMAQ 5.3.1 configuration for this study.

26

27 We have updated the wildfire emissions system in CMAQv5.3.1 based on the Blended Global  
28 Biomass Burning Emissions Product (GBBEPx) (Zhang and Kondragunta, 2006; Zhang et al.,  
29 2011). The GBBEPx uses satellite-detected fire radiative power (FRP) to estimate wildfire  
30 smoke emissions for a number of species: CO (carbon monoxide), NO<sub>x</sub> (nitrogen oxides), SO<sub>2</sub>  
31 (sulfur dioxide), elemental carbon, total organic aerosols, and PM<sub>2.5</sub>. The satellite FRP is  
32 estimated from satellite brightness temperature anomaly, and the GBBEPx processor assumes  
33 that the wildfire emissions are proportional to the FRP over certain land use type in certain  
34 regions. The GBBEPx emissions are based on polar orbiting satellites: MODIS (Aqua and Terra  
35 satellites) and VIIRS (Suomi-NPP and NOAA-20 satellites) instruments, which are updated once  
36 per day. A wildfire emission preprocessor converts the GBBEPx emissions to CMAQ-ready  
37 input files using emission speciation and diurnal profiles (high during daytime and low at night)  
38 (adopted from U.S. EPA-based profiles) (Baker et al., 2016), and a daily scaling factor. Here we  
39 classify the wildfire into either a long-lasting fire (longer than 24 hours) or short-term fire  
40 (shorter than 24 hours) based on land use types and regions. Only the fires west of 110°W that



1 have a model grid cell total forest fraction  $> 0.4$  are assumed to be long-lasting fires, which incur  
2 daily scaling factors of 1, 0.25, 0.25 for days 1, 2 and 3, respectively. All other short-term  
3 GBBEPx fires are assumed to have smoke emissions for 24 hours (i.e., day 1 only). CMAQ  
4 treats wildfire emissions as point sources that undergo inline plume rise to distribute the smoke  
5 vertically. The default CMAQ plume rise used here is based on Briggs (1965), which is driven  
6 by fire heat flux (converted from FRP with a ratio of 1) and fixed burning area (assumed to be  
7 10% of the  $0.1^\circ \times 0.1^\circ$  grid cell).

8

### 9 **3. Model Evaluations over the U.S. for August 2019**

10 To first gain a general picture and compare the overall GFS-CMAQ and WRF-CMAQ model  
11 performances, in this section we evaluate near-surface meteorological and air quality predictions  
12 during the FIREX-AQ August 2019 period against NOAA's METeorological Aerodrome Report  
13 (METAR; [https://madis.ncep.noaa.gov/madis\\_metar.shtml](https://madis.ncep.noaa.gov/madis_metar.shtml)) and the U.S. EPA's AirNow  
14 (<https://www.airnow.gov/>) observation networks.

15

#### 16 **3.1 Domain-Wide Meteorology against the METAR Network**

17 Figure 1 shows the mean bias (MB) of GFS and WRF predicted surface meteorological variables  
18 compared to METAR data during August, 2019. Both meteorological models have a cool bias  
19 over the Western and Northeastern United States, and a warm bias over the western Rocky  
20 Mountain region and Southeastern United States (Figure 1a, 1b). Similar temperature predictions  
21 are expected since WRF uses the FDDA method nudging toward GFS data. However, GFS tends  
22 to be cooler than WRF over the Rocky Mountains and in the central and northeastern USA due  
23 to their different dynamics and physics. The GFSv16 cold bias in the lower troposphere is  
24 impacted by excessive evaporative cooling from rainfall (personal communication with  
25 NOAA/NCEP). Campbell et al. (2022) had detailed discussions about GFSv16 biases.

26

27 Both GFSv16 and WRF models have similar and rather significant dry biases for specific  
28 humidity (SH) predictions across CONUS (Figure 1c, 1d). Qian et al (2020) investigated this  
29 common dry bias in many models, and found that neglecting an irrigation contribution could  
30 cause this dry bias. GFS has widespread dry biases (Campbell et al. 2022) and WRF has similar  
31 dry biases, too as it is nudged toward GFS. There are some noticeable differences for certain  
32 regions. For instance, WRF has less dry bias over Southern Texas than GFS.

33

34 Both models underestimate the mean 10-m wind speeds compared to METAR stations over the  
35 western U.S. WRF has stronger underpredictions over the Rocky Mountains and overpredictions  
36 over northeastern U.S., while GFS has stronger underpredictions over the Appalachian  
37 Mountains and overpredictions over Texas and Oklahoma. GFSv16's operational verification  
38 also ([https://www.emc.ncep.noaa.gov/gmb/emc.glopara/vsdb/v16rt2/g2o/g2o\\_00Z/index.html](https://www.emc.ncep.noaa.gov/gmb/emc.glopara/vsdb/v16rt2/g2o/g2o_00Z/index.html))  
39 shows that it tends to underpredict the 10-m wind over the western U.S. during both daytime and



1 nighttime, but shows overprediction over the eastern U.S. Besides the difference of physical  
2 schemes, etc. (Table 1), other possible reasons causing this surface wind difference could be  
3 effect of the gravity-wave drag (GFSv16 includes it, but the WRF run here does not), and  
4 vertical resolution (GFS's 127 layers versus WRF's 35 layers). Some studies (Skamarock et al,  
5 2019) revealed the necessity of fine vertical resolution for atmospheric simulations, especially  
6 within the PBL, near tropospheric top, and during convective events. Insufficient vertical  
7 resolution could also cause plume dilution on chemical transport modeling (Zhuang et al., 2018).  
8 The gravity-wave drag is also known to produce synoptic scale body forces on the atmospheric  
9 flow over irregularities at the earth's surface such as mountains and valleys, and uneven  
10 distribution of diabatic heat sources associated with convective systems (Kim et al., 2003). Its  
11 parameterization is needed for large-scale models.

12  
13 There is strong regional variability in the monthly mean PBL height differences between GFS  
14 and WRF during daytime (represented by 18 UTC) and nighttime (represented by 06 UTC)  
15 (Figure 2). During daytime, GFS has a higher PBL height compared to WRF over the U.S.  
16 Pacific coast, northern Rocky Mountains, northeastern and southeastern U.S., but it becomes  
17 lower over the central U.S. (e.g., Texas, Oklahoma, and Kansas). At night, however, most of  
18 these regional differences between GFS and WRF are reversed. This diurnal difference is  
19 mainly driven by the different PBL schemes employed in GFS (Han and Bretherton, 2019) and  
20 WRF (i.e., YSU) and the associated other physical suites, including the land surface data. The  
21 GFS's PBL height has a strong diurnal variation over these regions, including the western and  
22 northeastern U.S. in the summer, including a sharp rise and collapse after sunrise and sunset,  
23 respectively (Campbell et al., 2022). The strong PBL diurnal variation has significant effects on  
24 the air quality predictions in GFS-CMAQ.

### 25 26 **3.2 Evaluation of Regional Meteorology and Air Quality against the AirNow Network**

27 The U.S EPA AirNow network provides hourly observations of near-surface ozone, fine  
28 particulate matter (PM<sub>2.5</sub>), and meteorology. Campbell et al. (2022) showed detailed verification  
29 of GFS-CMAQ with the surface AIRNow data. Here we focus on the difference between the  
30 interpolation-based GFSv16 versus WRF downscaling and the impacts on meteorological and  
31 chemical model performances. Figure 3 shows a comparison of these two models over two  
32 specific regions, the U.S. West (CA, OR and WA) and Northeast states (CT, DE, MA, MD, ME,  
33 NH, NJ, NY, PA, RI, VT and District of Columbia) (Figure S1), where the two models had  
34 relatively large differences for some meteorological variables. GFS and WRF had very similar 2-  
35 m temperatures over the Pacific coast states: Washington, Oregon and California, and both of  
36 them had similar cool bias (around 1K), R and RMSE (Figure 3a). However, these two models  
37 had significant differences for 10-m wind speed prediction over the Pacific coast (Figure 3c),  
38 where WRF overpredicted the wind speed, especially at night and in later August. Most AIRNow  
39 stations are located near urban or suburban areas, which generally have weaker 10-m wind speed  
40 than those at the METAR aviation weather stations near airports. For this reason, although



1 Figure 1e and 1f shows that GFS and WRF underpredict monthly-mean wind speed over the  
2 METAR stations in the West, they still tend to overpredict AIRNow wind (Figure 3c), especially  
3 for the WRF 10-m wind speed at night. Considering that the model grid cells represent  $12 \times 12$   
4  $\text{km}^2$  averages, the true model-observation comparisons likely fall somewhere between the  
5 urban/suburban AIRNow stations and METAR stations, depending on the land use fractions of  
6 each grid. Obviously the observation representation characteristics could affect the verification  
7 results. Compared to AIRNow stations, GFSv16 has overall better scores for surface wind speed  
8 predictions over the U.S. West, where the WRF's larger surface wind speed overprediction is  
9 associated with its PBL height predictions (Figure 3e, 3f). During the nighttime, GFS has a lower  
10 PBL height (10–50% lower than WRF) and weaker vertical mixing, which brings less  
11 momentum from the upper layers to the surface, which led to lower nighttime wind and better  
12 agreements with the AIRNow wind-speed observation.

13

14 Over the northeast, the mean bias (MB) of GFS temperature was about  $-1\text{K}$ , while the WRF has  
15 a smaller, slightly positive MB of about  $0.22\text{K}$  (Figure 3b). However, the GFS's temperature  
16 prediction has a better correlation coefficient,  $R$ , and RMSE, implying that it better captures  
17 some events, such as the 28–29 of August. Both models overpredict 10-m wind speeds in the  
18 northeast, but the GFS model yields better results due to a slightly lower PBLH at night (Figure  
19 3f) than WRF that had significant overpredictions, especially during 25–29, August (Figure 3d)  
20 when the tropical storm Erin approached this region. Especially on 28 of August, when the storm  
21 was centered near the east coast of North Carolina, the WRF run significantly underpredicts 2-m  
22 temperature (Figure 3b) and overpredict 10-m wind speed (Figure 3d), implying that the some  
23 WRF settings lead to relatively large surface prediction bias for the storm weather, such as its  
24 relatively coarse vertical resolution compared to the 127-layer GFS model.

25

26 Figures 4a and 4b show the ozone predictions of the two models over these two regions, and  
27 GFS-CMAQ yields predominantly lower  $\text{O}_3$  than WRF-CMAQ, especially at night. Over the  
28 northwest, the lower ozone in GFS-CMAQ is associated with their PBL height difference. First,  
29 with a certain dry deposition velocity between the models, it is easier to deplete ozone given the  
30 smaller volume of a shallower PBL. Second, the thinner PBL results in higher  $\text{NO}_x$   
31 concentrations and ozone titration rates near  $\text{NO}_x$  source regions, and consequently lower ozone  
32 there at night. Last, the lower PBL leads to weaker vertical mixing and downward transport of  
33 ozone from the residual-layer at night (Caputi et al, 2019). All these factors contributed to the  
34 lower nighttime ozone of GFS-CMAQ compared to WRF-CMAQ. Since GFS-CMAQ already  
35 underpredicts ozone due to combined meteorological factors, such as the temperature  
36 underprediction (Figure 4a), the GFS-CMAQ's further ozone reduction (possibly due to its lower  
37 PBLH at night) exacerbates its low bias. However, over the Northeast, the similar impacts help  
38 the GFS-CMAQ yield much better MB due to its better agreement with the observed nighttime  
39 low ozone over the Northeast. Over the entire CONUS domain, the situation is similar: for an  
40 average August 2019, the GFS-CMAQ has a lower ozone MB ( $1.1 \text{ ppb}$ ) compared to WRF-



1 CMAQ (4.7 ppb). Figure 5 shows that both models have similar daytime ozone prediction over  
2 CONUS. However, GFS-CMAQ better captures low nighttime ozone over the U.S. East than  
3 WRF-CMAQ (Figure 5c, 5d).  
4  
5 GFS-CMAQ has substantially higher PM<sub>2.5</sub> mean concentrations over the U.S. West, but lower  
6 over the U.S. Northeast compared to WRF-CMAQ (Figures 4c, 4d). These model differences are  
7 also related to their interpolated GFSv16 versus downscaled WRF meteorological drivers.  
8 Because both models use the same emissions under relatively clean background conditions in the  
9 west (i.e., prevailing westerly flow from the Pacific Ocean), the PBL and wind speed differences  
10 have significant impacts on their near-surface pollutant concentrations, especially at night. Both  
11 models show strong PM<sub>2.5</sub> diurnal variation (high at night and low during daytime), driven by  
12 the meteorological diurnal variation (e.g. PBL), which overcomes the emission diurnal variation  
13 (usually high during daytime and low at night). Compared to WRF-CMAQ, GFS-CMAQ has  
14 lower nighttime PBL height and weaker wind speed at night, which leads to weaker vertical  
15 mixing and venting, and increases the pollutant concentrations near the surface and yields higher  
16 surface PM<sub>2.5</sub> over the U.S. West (Figure 4c). Its higher surface PM<sub>2.5</sub> could also result in  
17 stronger local dry deposition. In contrast to local vertical mixing and venting effects on PM<sub>2.5</sub>  
18 discussed above, there are strong (and potentially counterbalancing) impacts of model PBL and  
19 horizontal wind speed differences on downstream PM<sub>2.5</sub> concentrations at night. WRF-CMAQ's  
20 deeper PBL and stronger wind speeds at night (Figures 3c–3f) tends to transport aerosols and  
21 their precursors more efficiently downstream via the dominant advection pathway. Figure 6  
22 shows that these monthly mean background PM<sub>2.5</sub> differences appear in East of Rocky  
23 Mountain (WRF-CMAQ is about 2  $\mu\text{g}/\text{m}^3$  higher) during both daytime and nighttime. This effect  
24 is very prominent in the Northeast region. Although both models predicted similar PM<sub>2.5</sub>  
25 magnitude over the U.S. Northeast, GFS-CMAQ yields the overall PM<sub>2.5</sub> underprediction, and  
26 its monthly-mean PM<sub>2.5</sub> is 2.6  $\mu\text{g}/\text{m}^3$  lower than the WRF-CMAQ prediction (Figure 4d).  
27 Especially during 01–09 August, WRF-CMAQ had about 4  $\mu\text{g}/\text{m}^3$  higher surface PM<sub>2.5</sub>  
28 background than that of GFS-CMAQ. In this case, the WRF-CMAQ model has a better  
29 agreement with observations (Figure 4d). It is possible that the GFS-CMAQ's nighttime PBL  
30 heights (wind speeds) are too shallow (weak) in this case, which does not allow enough transport  
31 of pollutants to the downstream (Eastern USA). Overall, GFS-CMAQ and WRF-CMAQ have  
32 mixed performances for PM<sub>2.5</sub> predictions during the August 2019 period: GFS-CMAQ has  
33 better PM<sub>2.5</sub> prediction over the U.S West, and WRF-CMAQ yields better results over east of  
34 Rocky mountain (Figure 6).

#### 35 **4 Model Comparisons against the FIREX-AQ Aircraft Data**

36 From late July to early September, 2019, the joint NOAA-NASA FIREX-AQ field campaign  
37 (<https://csl.noaa.gov/projects/firex-aq/>) employed a suite of satellites, aircraft, vehicles and  
38 ground site platforms aimed to observe, analyze, and characterize air pollutants emitted from  
39 wildfire sources over the CONUS (Ye et al., 2021). The FIREX-AQ airborne measurements



1 provide a three-dimensional dataset from various meteorological, gas, and aerosol instruments  
2 that is used to verify the GFS-CMAQ and WRF-CMAQ model performance, while elucidating  
3 reasons for any model differences. Here the focus of the FIREX-AQ model comparison and  
4 verification is against observations taken primarily from the NASA DC-8 aircraft, which include  
5 meteorological variables, gaseous and aerosol concentrations, and aerosol optical properties. The  
6 majority of the FIREX-AQ flights were over the western United States, and sampled within  
7 environments that both were *and* were not (see section 4.1) influenced by wildfire emissions  
8 ([https://daac.ornl.gov/MASTER/guides/MASTER\\_FIREX\\_AQ\\_JulySept\\_2019.html](https://daac.ornl.gov/MASTER/guides/MASTER_FIREX_AQ_JulySept_2019.html)). During a  
9 cluster of major wildfire events (see Section 4.2), the DC-8 sampled both near-source and aged  
10 smoke plumes between 02–08 August, 2019 (i.e., the Williams Flats, Snow Creek, and Horsefly  
11 Fires) across the states of Idaho, Washington, and Montana.

12

#### 13 **4.1 Comparison of the July 22 non-wildfire event over the central California Valley**

14 On 22 July, the DC-8 aircraft flew from California to Boise, Idaho, while maintaining a  
15 relatively low-altitude (<1 km) above sea level (ASL) over the California Central Valley (Figure  
16 7). This flight was not impacted by any major wildfire event, and was mainly controlled by  
17 anthropogenic emissions and local meteorological conditions. Figure 7 shows that the GFSv16  
18 and WRF models had similar meteorological temperature and humidity predictions, and that both  
19 models have dry and warm biases over the Central Valley at lower altitudes (Figures 7d–7e)  
20 (Yun et al., 2020). GFS’s horizontal wind speeds tended to have stronger variability than WRF  
21 (Figure 7b), especially in high altitudes. For wind direction, WRF showed a better prediction  
22 than GFS around 20 and 24 UTC (Figure 7c).

23

24 Both GFS-CMAQ and WRF-CMAQ underestimate the vertical wind (W) variability by at least  
25 one order of magnitude, and WRF-CMAQ has weaker W variability than that of GFS-CMAQ,  
26 especially in high altitudes (Figure 7f). The model vertical velocities are not from the GFS or  
27 WRF model, but rather they are re-diagnosed in CMAQ to conserve mass (Otte and Pleim,  
28 2010), and thus represent the whole layer’s vertical movement across the 12 km by 12 km grid  
29 cell. With its flight speed around 80 to 240 m/s, the DC-8 aircraft’s one-minute average  
30 sampling frequency results in an approximate 4.8 to 14 km horizontal scale, respectively, which  
31 is comparable to the 12 km CMAQ model resolution. The aircraft observations, however, include  
32 turbulence effects during its one-minute averages, which may not be temporally resolved by  
33 CMAQ at this resolution. Thus, both the GFS-CMAQ and WRF-CMAQ model vertical  
34 velocities are much lower and have almost no correlation with the aircraft observations.

35

36 Although both GFS-CMAQ and WRF-CMAQ have reasonable comparisons for most  
37 meteorological variables, including the horizontal winds, it continues to be a challenge to  
38 compare them with the observed vertical velocities. Thus to further elucidate the model vs.  
39 observation differences in vertical motions, Figure 8 shows a curtain plot of vertical velocities  
40 along the flight path from the two models. Since WRF-CMAQ remains in a native grid, its wind



1 fields tend to be more balanced and have lower variability compared to the interpolated GFS-  
2 CMAQ wind fields. The stronger variability in W for GFS-CMAQ represents CMAQ's effort to  
3 counteract mass inconsistency effects from the interpolated horizontal wind fields (Byun,  
4 1999b).

5  
6 GFS-CMAQ and WRF-CMAQ overall yield similar results for specific chemical species during  
7 this DC-8 flight (Figure 9). Both models underestimate CO, O<sub>3</sub> and ethane (C<sub>2</sub>H<sub>6</sub>) concentrations  
8 over the lower altitudes in the California Central Valley. Over the same flight segment, they had  
9 better NO<sub>x</sub> (NO + NO<sub>2</sub>) and ethene (C<sub>2</sub>H<sub>4</sub>) predictions, implying that the emissions of these two  
10 species have better accuracy than those of CO and ethane. Figure 9f shows that the two models  
11 also underestimate NO<sub>z</sub> (NO<sub>y</sub>-NO<sub>x</sub>), or the oxidized nitrogen species besides NO<sub>x</sub>, indicating  
12 that the photochemical ozone production may also be underestimated. NO<sub>z</sub> is a good indicator of  
13 the ozone photochemical formation (Sillman et al., 1997), where the O<sub>3</sub>/NO<sub>z</sub> ratio represents the  
14 ozone photochemical efficiency per NO<sub>x</sub> oxidation products. Thus, NO<sub>z</sub> and O<sub>3</sub> are typically  
15 highly correlated over regions with active photochemical production. The O<sub>3</sub> and NO<sub>z</sub>  
16 underestimations are likely due to the underestimation of CO and some hydrocarbons, such as  
17 ethane, as they are precursors of O<sub>3</sub>.

18  
19 The two models show slight differences in peak values of CO, ethene, and NO<sub>x</sub> around 23:30  
20 UTC, where the GFS-CMAQ predicted concentrations are slightly higher and closer to  
21 observations (Figure 9). These differences are due to their PBL predictions (both from the  
22 corresponding meteorological model outputs), where GFS-CMAQ has a lower PBL height and  
23 weaker emission vertical dilution compared to WRF-CMAQ (Figure 8). GFS-CMAQ tends to  
24 underpredict O<sub>3</sub> more (Figure 9b), however, due to its higher NO<sub>x</sub> titration. This implies that the  
25 effects of the transport and non-local transformation of O<sub>3</sub> could be stronger than that of local  
26 precursor emissions. WRF-CMAQ has higher NO<sub>z</sub> (Figure 9f), but lower NO<sub>x</sub> compared to  
27 GFS-CMAQ due to the time lag of O<sub>3</sub> and NO<sub>z</sub> photochemical formation. Consequently, the  
28 peak O<sub>3</sub> values may not be well correlated with the emitted precursors, such as NO<sub>x</sub> and volatile  
29 organic compounds (VOCs). Furthermore, the modeled peak C<sub>2</sub>H<sub>6</sub> and C<sub>2</sub>H<sub>4</sub> concentrations do  
30 not occur at the same time around 23:30 UTC, while observations indicate that these two species  
31 should be highly correlated in this region. This model mismatch implies that the VOC speciation  
32 factors for a certain area or emission sector need to be improved over Southern California.

#### 34 **4.2 Comparison of the 6 August wildfire events over the U.S. Northwest**

35 On 06 August, the DC-8 observed a cluster of three wildfires: the Williams Flats Fire (47.98 °N,  
36 118.624 °W, 80 km to the northwest of Spokane, Washington), Snow Creek Fire (47.703°N,  
37 113.4°W, 32 km northeast of Condon, Montana), and Horsefly Fire (46.963 °N, 112.441°W, 24  
38 km east of Lincoln, Montana). Figure 10a shows the flight path on that date, where the DC-8  
39 aircraft departed from Boise, ID, flew over the Williams Flats Fire region, then flew to Montana  
40 to sample the Snow Creek and Horsefly Fires (i.e., Montana Fires), and finally returned to the



1 Boise base. The aircraft flew below 8 km for most flight segments near the fire plumes. Figure  
2 S1 shows the corresponding GOES-16 satellite true color image, where these 06 August fires and  
3 associated smoke plumes are visible and can be distinguished from the cloud bands to the south  
4 that move northward later that day (Figure S2). The Williams Flats Fire was ignited by  
5 lightning, and was the largest fire event sampled during the FIREX-AQ campaign burning from  
6 about 02–08, August, 2019.  
7  
8 Both models significantly underpredicted CO (Figure 10c), submicron organic aerosol (Figure  
9 10e) and aerosol optical extinction coefficient (AOE) (Figure 10f), which suggests an issue with  
10 the GBBEPx gas and aerosol emissions. The models performed well for NO<sub>2</sub> during the  
11 Williams Flats and Montana Fires Fire below 6 km ASL, but there were prominent  
12 underestimations for the high-altitude flight segments (Figure 10d). This indicates that the  
13 background NO<sub>2</sub> was underestimated, or the models had insufficient inject height for fire plume  
14 rise (both based on Briggs, 1965). WRF-CMAQ predicted higher O<sub>3</sub> values than the GFS-  
15 CMAQ, which overall agreed better with observations for the Williams Flats Fire (Figure 10b).  
16 However, for the Montana Fires (~ 23–24 UTC), WRF-CMAQ has higher O<sub>3</sub> biases and GFS-  
17 CMAQ yields better results. The difference in O<sub>3</sub> is largely driven by the background  
18 concentration difference between the two models, where WRF-CMAQ tends to have higher  
19 domain-wide O<sub>3</sub> than GFS-CMAQ due to the meteorological effects discussed in Sections 3.  
20  
21 Figure S3 shows the spatial overlay comparison of vertically averaged GFS-CMAQ predictions  
22 at 21 UTC and the DC-8 flight observations for the altitude 1–3 km above ground level (AGL),  
23 on 6 August, 2019. The peak NO<sub>2</sub> observation around 118.5°W, 48°N indicates the general  
24 location of the Williams Flats fire. The GBBEPx emission and GFS-CMAQ prediction showed  
25 shifted peak-value locations driven by the westerly modeled winds. For this flight, the GBBEPx  
26 had stronger NO<sub>x</sub> fire emission over two Montana locations than that over Williams Flats. The  
27 model overpredicts the column averaged NO<sub>2</sub> concentrations, especially over the Montana fires,  
28 which can not be reflected by the point-by-point NO<sub>2</sub> comparison result in Figure 10d. For this  
29 flight, the mean GFS-CMAQ NO<sub>2</sub> along the flight path for 1–3km AGL is about 0.125 ppbv  
30 compared to the observed mean NO<sub>2</sub> of 0.169 ppbv, and the model indeed showed NO<sub>2</sub>  
31 underprediction along the flight path. However, in this case, the flight path missed some  
32 locations where modeled peak NO<sub>2</sub> values existed or the modeled transport misplaced the  
33 plumes, especially over the Montana fires leading to this inconsistency. For ozone comparison  
34 (Figure S3b), this inconsistency could also exist, though not as significant as for the high-  
35 gradient NO<sub>2</sub> concentrations. In the GFS-CMAQ prediction, the high ozone concentrations are  
36 almost co-located with high NO<sub>2</sub> concentration (Figure S3b), but the observation did not show  
37 this feature. Instead, some high-O<sub>3</sub> flight segments had relatively low NO<sub>2</sub>, such as those circled  
38 in the black rectangle box of Figure S3b. The observed NO<sub>x</sub> titration was not able to be  
39 produced by the 12 km models. Wang et al (2021) used a 100m horizontal resolution large eddy  
40 simulation and demonstrated the capability of using such techniques to capture some high-



1 resolution fire plume features and associated chemical behavior. While such high resolution  
2 techniques are not currently feasible for the operational NAQFC, they demonstrate the limitation  
3 of using regional scale (12×12 km) models to capture such fine scale features of plume behavior.  
4  
5 GFS-CMAQ has higher wildfire-related CO, NO<sub>2</sub>, OA and AOE values that are closer to  
6 observations than WRF-CMAQ for the Montana Fires between 23–24 UTC at flight altitudes of  
7 ~ 4–5 km (Figure 10c–10f). Since these two models use the same GBBEPx emissions and  
8 wildfire plume rise algorithm (Briggs, 1965), the differences should be due to other reasons. To  
9 help explain these model differences, Figure 11a and 11b show the Differential Absorption High  
10 Spectral Resolution Lidar (DIAL-HSRL) retrieved aerosol backscatter coefficients (ABC)  
11 aboard the DC-8 aircraft without and with cloud screen, respectively. It shows that the major fire  
12 plumes of the William Flats Fire were below 4 km (~ 19–22 UTC), but the Montana Fires (~23–  
13 24 UTC) extended from the surface up to 6 km, with some detached plumes reaching 10 km. The  
14 model predicted AOE have an overall similar pattern, with major plumes below 4 km for the  
15 Williams Flats Fire (Figures 11c and 11d). Over the Montana Fires, the GFS-CMAQ predicts  
16 slightly higher PBL, thus allowing for the fire plume to reach a higher height near the DC-8  
17 cruising altitude. In contrast, the WRF-CMAQ wildfire plumes are slightly lower than the  
18 aircraft flight path around 23–24 UTC, which leads to underprediction in the fire emitted species  
19 (Figure 11d).  
20  
21 An interesting feature in the DIAL observations is the detached plume from 8 km to 10 km  
22 altitude (Figure 11a), where some cirrus clouds existed, and the DIAL retrieval could not  
23 distinguish whether they are pure clouds or clouds mixed with elevated aerosols above 8km. The  
24 cloud screen product (Figure 11b) mainly showed the enhanced aerosols below 7km and some  
25 scattered signals near the high cloud edges. Cloud mixing with aerosols was usual for fire  
26 induced clouds, or pyrocumulonimbus (Peterson et al., 2021). Although in this event, the middle-  
27 size fires did not show evident of inducing the high-altitude clouds, the indicator of mixed clouds  
28 and aerosols in high altitudes still existed: both in-situ measured OA (Figures 10e) and AOE  
29 (Figure 10f, 11c, 11d) showed the enhanced aerosols around 25 UTC above 8km. This elevated  
30 plume was generally captured by the GFS-CMAQ simulation, while underestimating its strength  
31 (Figure 11c); however, this feature was completely missed in WRF-CMAQ (Figure 11d).  
32 Considering the altitude range of the detached plume, the major model disparities are likely due  
33 to model convection differences in the free troposphere. To further investigate this impact,  
34 Figures 11e and 11f show curtain plots of RH predicted by the two models. GFS-CMAQ yields  
35 higher RH at such altitudes (10 km) compared to WRF-CMAQ around 23–24 UTC, indicating  
36 that the GFS-CMAQ has stronger convection. The CMAQ model uses inputted meteorology to  
37 diagnose convection activity and drive its ACM2 convection scheme. This convective activity is  
38 apparent in GOES-16 satellite images (Figure S2), as more fractional clouds appeared ahead of  
39 the northward moving frontal band. Both the GFSv16 and WRF models used here *do not*  
40 consider the fire heat feedback effect, and thus their predicted convection and clouds are only



1 driven by the synoptic weather conditions. If such synoptic-to-mesoscale weather models  
2 consider wildfire heat feedback effects, their predictions may result in stronger convection and  
3 help correct underpredictions in PBL heights.

4

### 5 **4.3 Statistical Results of Model Performances for FIREX-AQ**

#### 6 4.3.1 *Meteorological Statistics*

7

8 During the FIREX-AQ field campaign, the DC-8 aircraft performed more than 20 flights over  
9 CONUS with detailed observations of various chemical compounds. Tables 2 and 3 show the  
10 statistical results of mean bias (MB), normalized mean bias (NMB), root mean square error  
11 (RMSE), correlation coefficient (R), and linear regression/slopes for the two models'  
12 performance over the western U.S. (west of 110°W) only at low altitudes (<3km ASL) for both  
13 non-fire and fire flight segments. Most of these flights departed from Boise, ID, except the 22  
14 July flight that flew from California to Idaho. As a result, they mainly flew over Idaho and its  
15 surrounding regions. The GFS tends to have slightly higher wind speed with positive MB, while  
16 WRF has a small negative wind speed bias. Most of the DC-8 flights are during the daytime, and  
17 the GFS has a higher daytime wind speed than WRF at low altitudes. The GFS and WRF have  
18 very similar temperature predictions. For the RH, the GFS predictions were slightly dryer than  
19 those of WRF, especially for non-fire events. The meteorological models do not consider  
20 wildfire heat effects, and thus may have (in part) led to slightly warm MB for the non-fire events  
21 (Table 2) and slightly cool MB for the fire events (Table 3). Because both the GFSv16 and WRF  
22 models have similar MB shifts from an average temperature overprediction (Table 2; non-fire  
23 events) to an underprediction (Table 3; wildfire events), we can estimate that the fire effects  
24 made roughly a 1–2 Kelvin temperature enhancement to the background along the DC-8 flight  
25 paths below 3 km. This estimate assumes that the model temperature biases are generally  
26 representative of the western U.S. (west of 110°W), and are independent of the averaged flight  
27 segments that have different locations and periods in Table 2 and Table 3. Correspondingly, the  
28 air masses are dryer in the sampled wildfire plumes, as shown by the large reduction in the RH  
29 underpredictions (i.e., negative MBs) from Table 2 to Table 3.

30

#### 31 4.3.2 *Chemical Statistics During Non-Fire Events*

32

33 For most chemical species, the two models also have similar performance, indicating that the  
34 emissions and chemistry are major driving forces. For non-fire events, both models overpredict  
35 NO<sub>x</sub>, HNO<sub>3</sub>, toluene, EC, and ammonium (NH<sub>4</sub><sup>+</sup>), but underestimate PAN, benzene, C<sub>2</sub>H<sub>2</sub>, SO<sub>2</sub>,  
36 and submicron sulfate and organic aerosols (OA) (Table 2). The SO<sub>2</sub> and submicron sulfate  
37 underprediction may be impacted by underestimated NEIC2016v1 SO<sub>2</sub> emissions over the  
38 western U.S. Since point sources, including power plant emissions, are the SO<sub>2</sub> sources, this  
39 comparison implies that the point sources for 2019 events have large uncertainties.

40



1 Although the models agree well with NO<sub>y</sub> observations, they disproportionately underestimate  
2 NO<sub>z</sub> as shown by the regression slopes and MBs. One of the important NO<sub>z</sub> species is PAN, and  
3 both models underestimate PAN during the non-fire events (Table 2). PAN's carbonyl  
4 precursors include acetaldehyde (CH<sub>3</sub>CHO) (44% of the global source), methylglyoxal (30%),  
5 acetone (7%), and a suite of other isoprene and terpene oxidation products (19%) (Fischer et al.,  
6 2014). CH<sub>3</sub>CHO and acetone are also underestimated (Table 2), and help explain PAN's  
7 underestimation. For the oxidized hydrocarbons, like aldehydes (HCHO, CH<sub>3</sub>CHO), their main  
8 atmospheric sources are the oxidation of highly reactive VOCs, including alkanes, alkenes, and  
9 aromatics, instead of direct emissions (Parrish et al., 2012). So, the underestimation of HCHO  
10 and CH<sub>3</sub>CHO are associated with the underestimation of their precursor hydrocarbons, including  
11 anthropogenic and biogenic VOCs. Our other comparison indicated that BEIS tends to  
12 underpredict biogenic emission over the U.S. West, e.g. isoprene in Table 2. In this comparison,  
13 most anthropogenic hydrocarbons are disproportionately underestimated, except toluene,  
14 implying the VOC speciation issue in the NEIC2016v1 anthropogenic emissions (Table 2).  
15 Previous work had discovered that a model overprediction in toluene was also related to the  
16 toluene speciation in the NEI emission inventory (Lu et al., 2020).

17  
18 Submicron ammonium (NH<sub>4</sub><sup>+</sup>) and the nitrate ion are also underestimated by both models during  
19 non-fire events (Table 2), suggesting there are NH<sub>3</sub> underestimates due to either insufficient NH<sub>3</sub>  
20 emissions or exaggerated NH<sub>3</sub> removal processes. There are, however, overpredictions in the  
21 intermediate species nitric acid (HNO<sub>3</sub>). It implies that the HNO<sub>3</sub> accumulates in the atmosphere  
22 because the modeled NO<sub>z</sub> pathways toward the nitrate ion and organic nitrate aerosol products  
23 are reduced due to their other precursor (NH<sub>3</sub> and VOCs) underestimation.

24  
25 There are underestimations in the VOC and CO concentrations, which contributes to the ozone  
26 underestimation during non-fire events (Table 2). These non-fire comparisons also highlight that  
27 both models have similar biases due to similar meteorology (Section 4.3.1), and the use of the  
28 same anthropogenic emissions (NEIC2016v1), BEIS biogenic emission and chemical  
29 models/mechanisms (i.e CMAQv5.3.1). The differences in the two models' bias, error, and  
30 correlation/slope are much smaller than their individual magnitudes.

### 31 32 4.3.3 Chemical Statistics During Fire Events

33 The WRF-CMAQ and GFS-CMAQ models significantly underestimate CO, VOC, HONO, and  
34 OA for fire events at low altitudes (< 3 km) over the western U.S. (Table 3). In conjunction with  
35 underestimated GBBEPx emissions during these wildfire events, other possible causes for the  
36 average statistical underprediction are the CMAQ model's 12 km horizontal resolution and the  
37 flight sampling coverage. Most of the fires that are averaged in the statistics, such as the Horsefly  
38 (5.5 km<sup>2</sup> burning area) and Snow Creek Fires (7.3 km<sup>2</sup> burning area), are at a much finer scale  
39 than the model grid. Only the largest Williams Flats Fire, with a total burning area of 180 km<sup>2</sup>  
40 (Ye et al., 2021), had a comparable horizontal scale to the model resolution.



1  
2 The DC-8 aircraft had many flight segments near wildfire sources during the fire events in Table  
3 3, and thus dilution of the emissions due to the relatively coarse model resolution may lead to  
4 underestimations in the predicted slope for most wildfire emitted pollutants, such as CO and OA  
5 (Table 3). The O<sub>3</sub> concentrations are also underestimated; however, the O<sub>3</sub> underpredictions are  
6 reduced from the non-fire (Table 2) to fire events (Table 3). Abundant amounts of wildfire  
7 emitted NO<sub>x</sub> can titrate ozone near the fire source region, and the models likely underestimate  
8 these titration effects due to the 12 km model resolution. Thus, the models cannot capture the  
9 strong spatial O<sub>3</sub> variability that is observed due to both reduction near source regions and  
10 enhancement in downstream areas. Again, for this fire event comparison, both models showed  
11 similar behavior and their differences were relatively smaller compared to the overall model  
12 biases.

## 14 **5. Summary and Discussion**

15 The operational NOAA/NWS National Air Quality Forecasting Capability (NAQFC) recently  
16 underwent a major upgrade on July 20, 2021. The advanced NAQFC includes the recent  
17 Community Multi-scale Air Quality (CMAQ) model version 5.3.1 with CB6 (carbon bond  
18 version 6)-Aero7 (version 7 of the aerosol module) chemical mechanism, and is driven by the  
19 latest operational Finite Volume Cubed-Sphere (FV3)-Global Forecast System, version 16  
20 (GFSv16) (Campbell et al., 2022). Here we analyze the impacts of the driving meteorological  
21 models on CMAQ model performance with the new GFSv16 interpolation-based meteorology  
22 versus the commonly-used native-grid Weather Research and Forecasting (WRF) model version  
23 4.0.3 meteorology. The meteorological and chemical analysis includes both 2D ground-based  
24 and 3D aircraft measurements during the summer 2019, which encompasses the joint NOAA-  
25 NASA Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ)  
26 campaign. As CMAQ has existing mass conservation via adjustments of the contravariant  
27 vertical velocity (Otte and Pleim, 2010), the NACC interpolated GFSv16 wind field can be well  
28 handled in CMAQ (i.e., GFS-CMAQ).

29  
30 The different NWS/NOAA operational GFS and commonly chosen WRF physics schemes  
31 employed in this study (Table 1) clearly have impacts on temperature, wind fields (both  
32 horizontal/advection and vertical/convection), PBL heights, and the corresponding CMAQ  
33 model predictions. During this study period over the U.S. West, both models showed moisture  
34 dry bias and temperature warm bias in low altitudes, which could be due to the issue mentioned  
35 by Qian et al (2020) and impacts from soil moisture deficits on surface fluxes in both models.  
36 Due to their different physics, GFS had stronger diurnal variation of PBL height, lower at night  
37 and higher during daytime over the U.S. West and Northeast. The differences in the GFS and  
38 WRF physics have a larger impact than the meteorology driver methodologies (i.e., interpolation  
39 vs. native) on the models' meteorological and air quality predictions, even despite using FDDA  
40 to nudge WRF simulation toward the GFSv16 data. While FDDA nudging was used here in



1 WRF to avoid growing errors across a continuous 1-month simulation, we note that if it is turned  
2 off, the differences between GFSv16 and WRF predictions would have been even greater. This  
3 would further substantiate the dominance of using different model physics compared to using  
4 different meteorological-driver methodologies (i.e., native vs. interpolation) and their impacts on  
5 CMAQ model predictions. Overall, the results of this study further corroborate the use of the  
6 GFSv16 data and NACC interpolation-based methods (Campbell et al., 2022) for regional  
7 CMAQ model applications in the scientific community.

8  
9 Over CONUS, GFS-CMAQ demonstrated lower mean surface ozone (by about 3 ppb) and  
10 PM<sub>2.5</sub> (by about 1 µg/m<sup>3</sup>) than WRF-CMAQ in August 2019 (section 3). In the western U.S.,  
11 the GFS has a stronger diurnal variability in PBL height and a better performance in nighttime  
12 10-m wind speeds compared to WRF. The nighttime difference between these two models tends  
13 to be more significant than the corresponding daytime difference. Their difference is also  
14 impacted by both vertical/convective (mainly daytime) and upstream advective transport  
15 differences in GFS-CMAQ and WRF-CMAQ, which somewhat confounds the impact of  
16 different meteorological physics on chemical predictions from region to region. This transport  
17 effect is more significant on PM<sub>2.5</sub> than that on O<sub>3</sub>, as O<sub>3</sub> has a shorter lifetime and is more  
18 sensitive to local emissions in summer. In this study, neither GFS-CMAQ nor WRF-CMAQ  
19 show overwhelming performance advantage over the other, similar to the NMM-CMAQ and  
20 ARW-CMAQ comparison in Yu et al. (2012a, 2012b).

21  
22 GFS-CMAQ and WRF-CMAQ demonstrated rather similar performance for major chemical  
23 variables during both FIREX-AQ non-fire (Table 2) and fire events (Table 3). In most FIREX-  
24 AQ events, both GFS-CMAQ and WRF-CMAQ showed similar biases, indicating that other  
25 factors, including emissions, model resolution and chemistry etc. could be more important for the  
26 model predictions compared to the meteorological differences. The aircraft data comparison  
27 reveals many common issues in both model systems. One critical issue is whether the flight  
28 sampling coverage is comparable to the 12 km model resolution, especially for high-gradient fire  
29 emission, e.g. the case of 06 August flight (Figure S3). The observation representation issue also  
30 exists in other places, such as near-surface meteorological comparison between AIRNow stations  
31 and METAR stations. Emission is the driving force for atmospheric composition concentrations.  
32 The comprehensive aircraft measurements help verify that the anthropogenic NEIC2016v1  
33 inventory is overall reasonable, except for SO<sub>2</sub>, NH<sub>3</sub> and certain hydrocarbons. The wildfire  
34 emission has bigger uncertainties, including the emission intensities, specification and plume  
35 rise, shown by the both models' results.

36  
37 The NACC interpolation method is advantageous as it enables using the original meteorological  
38 driver directly via interpolation, and avoids running another model such as WRF as a downscaler  
39 for regional CMAQ applications. It is also faster, and more consistent with the original  
40 meteorological driver. These aspects can simultaneously benefit real-time forecasting and



1 retrospective air quality applications in the scientific community. NACC can also adapt to  
2 quickly use any regional domain globally, and may also use other global meteorological data  
3 including reanalysis products. This helps mitigate the confounding factors of using different  
4 model configurations across the myriad of WRF physics options, while alleviating the difficulty  
5 in understanding their impacts on air quality predictions. The operational GFSv16 and associated  
6 reanalysis products are well vetted and evaluated across different global agencies and  
7 laboratories, and thus are well suited for regional CMAQ applications using NACC. In fact, there  
8 is an ongoing project at NOAA to migrate both the GFSv16 data and NACC software to the  
9 Amazon Web Services (AWS) Cloud platform to provide a streamlined product for the user to  
10 generate the model-ready meteorological data for any regional CMAQ application globally.

11  
12 Finally, we note that the current operational GFSv16 has enough meteorological variables to  
13 drive CMAQ with other supplied data (fractional landuse, LAI etc), and its C768 grid has  
14 horizontal resolution from 10.21 km to 14.44 km, which is close to the NAQFC's 12 km  
15 horizontal resolution. However, some commonly available global meteorological data, such as  
16 NCEP or ECMWF reanalysis data, may not have all meteorological variables needed by CMAQ,  
17 and have relatively coarse model resolutions. In this case, the WRF downscaling may become  
18 the only available method to drive a finer scale CMAQ model application. WRF data generated  
19 by different physics may be good for a finer scale CMAQ simulation; however NACC  
20 developments are underway to also process/interpolate higher resolution FV3-based Limited  
21 Area Models (LAMs) for direct application to CMAQ. All the physics schemes were developed  
22 according to certain regions and meteorological conditions. We again stress, however, that the  
23 downscaled WRF physics may significantly alter the original meteorological fields even with the  
24 FDDA nudging. As shown in this study, GFS and WRF had mixed performance for driving  
25 CMAQ.

#### 26 27 **Code and Data Availability.**

28 The FIREX-AQ field campaign data used in this study is in [https://www-air.larc.nasa.gov/cgi-](https://www-air.larc.nasa.gov/cgi-bin/ArcView/firexaq)  
29 [bin/ArcView/firexaq](https://www-air.larc.nasa.gov/cgi-bin/ArcView/firexaq) (last access, 16 May 2022). The NACC code used in this study is publicly  
30 available at <https://doi.org/10.5281/zenodo.5507489> and via GitHub at [https://github.com/noaa-](https://github.com/noaa-oar-arl/NACC.git)  
31 [oar-arl/NACC.git](https://github.com/noaa-oar-arl/NACC.git) (last access: 5 April 2022). The modified CMAQv5.3.1 for GFS-CMAQ is  
32 available at <https://doi.org/10.5281/zenodo.5507511> and via GitHub at [https://github.com/noaa-](https://github.com/noaa-oar-arl/NAQFC)  
33 [oar-arl/NAQFC](https://github.com/noaa-oar-arl/NAQFC) (last access: 5 April 2022).

#### 34 35 **Author contributions.**

36 YT contributed to the project conceptualization, model run, software, data analysis, visualization,  
37 investigation, and writing of the original draft. PCC contributed to software, the model run, data  
38 analysis, investigation and draft revision. DT and XZ contributed to wildfire emissions data. BB  
39 contributed to software and funding acquisition. FY, JH and HH provided the GFS model data.  
40 LP provided the global aerosol model for the lateral boundary condition. PL, RS, AS, JF, IS, JT-



1 D, YJ contributed to project supervision, project administration, and funding acquisition. MY,  
2 IB, JF, TR, DB, JS, J-LJ, JC, GD, RM, JH, GH, AR and JD contributed to the FIREX-AQ  
3 aircraft data.

4

#### 5 **Competing interests.**

6 The contact author has declared that neither they nor their co-authors have any competing  
7 interests.

8

#### 9 **Acknowledgements**

10 This research was funded by NOAA's National Air Quality Forecasting Capability (NAQFC) in  
11 the National Weather Service Office of Science and Technology Integration (NWS/OSTI).

12

#### 13 **References:**

- 14 Appel, K.W., Bash, J.O., Fahey, K.M., Foley, K.M., Gilliam, R.C., Hogrefe, C., Hutzell, W.T.,  
15 Kang, D., Mathur, R., Murphy, B.N. and Napelenok, S.L., The Community Multiscale Air  
16 Quality (CMAQ) model versions 5.3 and 5.3. 1: system updates and evaluation. *Geoscientific*  
17 *Model Development*, 14(5), pp.2867-2897. 2021. <https://doi.org/10.5194/gmd-14-2867-2021>
- 18 Arakawa, A. and Lamb, V.: Computational design of the basic dynamical processes of the  
19 UCLA general circulation model, *Methods in Computational Physics*, 17, 173–265, 1977.
- 20 Baker, K.R., Woody, M.C., Tonnesen, G.S., Hutzell, W., Pye, H.O.T., Beaver, M.R., Pouliot, G.  
21 and Pierce, T., Contribution of regional-scale fire events to ozone and PM<sub>2.5</sub> air quality  
22 estimated by photochemical modeling approaches. *Atmospheric Environment*, 140, 539-554.  
23 <https://doi.org/10.1016/j.atmosenv.2016.06.032>, 2016.
- 24 Briggs, G. A., A Plume Rise Model Compared with Observations, *Journal of the Air Pollution*  
25 *Control Association*, 15:9, 433-438, doi:10.1080/00022470.1965.10468404, 1965.
- 26 Byun, D. W., Dynamically Consistent Formulations in Meteorological and Air Quality Models  
27 for Multiscale Atmospheric Studies. Part I: Governing Equations in a Generalized Coordinate  
28 System, *Journal of the Atmospheric Science*, 56, 3789–3807. 1999a.  
29 [https://doi.org/10.1175/1520-0469\(1999\)056<3789:DCFIMA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<3789:DCFIMA>2.0.CO;2)
- 30 Byun, D.W., Dynamically consistent formulations in meteorological and air quality models for  
31 multi-scale atmospheric applications: Part II. Mass conservation issues. *Journal of the*  
32 *Atmospheric Science* 56, 3808–3820, 1999b. [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0469(1999)056<3808:DCFIMA>2.0.CO;2)  
33 [0469\(1999\)056<3808:DCFIMA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<3808:DCFIMA>2.0.CO;2).
- 34 Byun, D.W., Ching, J.K.S.. Science algorithms of the EPA models-3 Community Multiscale Air  
35 Quality (CMAQ) modeling system, EPA/600/R-99/030, U.S. EPA. 1999c
- 36 Byun, D. and Schere, K.L., Review of the governing equations, computational algorithms, and  
37 other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling  
38 system. *Appl. Mech. Rev.* 59: 51–77. <https://doi.org/10.1115/1.2128636>. 2006.



- 1 Campbell, P., Tang, Y., Lee, P., Baker, B., Tong, D., Saylor, R., Stein, A., Huang, J., Huang, H.-  
2 C., Strobach, E., McQueen, J., Pan, L., Stajner, I., Sims, J., Tirado-Delgado, J., Jung, Y.,  
3 Yang, F., Spero, T., and Gilliam, R.: Development and evaluation of an advanced National  
4 Air Quality Forecast Capability using the NOAA Global Forecast System version 16, *Geosci.*  
5 *Model Dev.* 15, 3281–3313, <https://doi.org/10.5194/gmd-15-3281-2022>, 2022.
- 6 Chen, F., and Dudhia, J. Coupling an advanced land surface-hydrology model with the Penn  
7 State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly*  
8 *Weather Review*, 129(4), 569–585. 2001. [https://doi.org/10.1175/1520-  
9 0493\(2001\)129<0569:CAALSH>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2)
- 10 Chen J.-H and S.-J Lin. The remarkable predictability of inter-annual variability of atlantic  
11 hurricanes during the past decade. *Geophysical Research Letters*, 38(L11804):6, 2011.  
12 <https://doi.org/10.1029/2011GL047629>
- 13 Chen J.-H. and S.-J. Lin. Seasonal predictions of tropical cyclones using a 25-km-resolution  
14 general circulation model. *J. Climate*, 26(2):380–398, 2013. [https://doi.org/10.1175/JCLI-D-  
15 12-00061.1](https://doi.org/10.1175/JCLI-D-12-00061.1)
- 16 Clough, S. A., Shephard, M. W., Mlawer, J. E., Delamere, J. S., Iacono, M. J., Cady-Pereira, K.,  
17 et al. (2005). Atmospheric radiative transfer modeling: A summary of the AER codes. *Journal*  
18 *of Quantitative Spectroscopy & Radiative Transfer*, 91(2), 233–244.  
19 <https://doi.org/10.1016/j.jqsrt.2004.05.058>
- 20 Caputi, D. J., Faloon, I., Trousdell, J., Smoot, J., Falk, N., and Conley, S.: Residual layer ozone,  
21 mixing, and the nocturnal jet in California's San Joaquin Valley, *Atmos. Chem. Phys.*, 19,  
22 4721–4740, <https://doi.org/10.5194/acp-19-4721-2019>, 2019
- 23 Dong, X., J. S. Fu, K. Huang, D. Tong, and G. Zhuang. Model development of dust emission  
24 and heterogeneous chemistry within the Community Multiscale Air Quality modeling system  
25 and its application over East Asia. *Atmos. Chem. Phys.*, 16, 8157–8180. 2016.  
26 <https://doi.org/10.5194/acp-16-8157-2016>.
- 27 Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003).  
28 Implementation of Noah land surface model advances in the National Centers for  
29 Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*,  
30 108(D22), 8851. <https://doi.org/10.1029/2002JD003296>
- 31 Fischer, E. V., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Millet, D. B., Mao, J., Paulot, F.,  
32 Singh, H. B., Roiger, A., Ries, L., Talbot, R. W., Dzepina, K., and Pandey Deolal, S.:  
33 Atmospheric peroxyacetyl nitrate (PAN): a global budget and source attribution, *Atmos.*  
34 *Chem. Phys.*, 14, 2679–2698, <https://doi.org/10.5194/acp-14-2679-2014>, 2014.
- 35 Fu, X., S. X. Wang, Z. Cheng, J. Xing, B. Zhao, J. D. Wang, and J. M. Hao, 2014: Source,  
36 transport and impacts of a heavy dust event in the Yangtze River Delta, China, in 2011.  
37 *Atmos. Chem. Phys.*, 14, 1239–1254, doi:10.5194/acp-14-1239-2014.
- 38 Grell, G. A., Dudhia, J., and Stauffer, D. R. A description of the fifth-generation Penn  
39 State/NCAR Mesoscale Model (MM5). NCAR tech. Note NCAR TN-398-1-STR, 117 pp.  
40 1994.



- 1 Han, J., & Bretherton, C. S., TKE-Based Moist Eddy-Diffusivity Mass-Flux (EDMF)  
2 Parameterization for Vertical Turbulent Mixing, *Weather and Forecasting*, 34(4), 869-886.  
3 2019. <https://doi.org/10.1175/WAF-D-18-0146.1>
- 4 Harris, L., Chen, X., Putman, W., Zhou, L. and Chen, J.H., A Scientific Description of the GFDL  
5 Finite-Volume Cubed-Sphere Dynamical Core. <https://doi.org/10.25923/6nhs-5897>. 2021
- 6 Hodyss, D. and N. Nichols, The error of representation: basic understanding, *Tellus A: Dynamic*  
7 *Meteorology and Oceanography*, 67:1, <https://doi.org/10.3402/tellusa.v67.24822>, 2015.
- 8 Hong, Song-You, Yign Noh, Jimmy Dudhia, A new vertical diffusion package with an explicit  
9 treatment of entrainment processes. *Mon. Wea. Rev.*, 134, 2318–2341. 2006.  
10 doi:10.1175/MWR3199.1.
- 11 Houyoux, M.R., Vukovich, J.M., Coats, C.J., Wheeler, N.J.M., Kasibhatla, P.S., Emission  
12 inventory development and processing for the seasonal model for regional air quality  
13 (SMRAQ) project. *J. Geophys. Res.*, 105(D7), 9079-9090,  
14 <https://doi.org/10.1029/1999JD900975>. 2000.
- 15 Huang, M., D. Tong, P. Lee, L. Pan, Y. Tang, I. Stajner, R. B. Pierce, J. McQueen, and J. Wang,  
16 Toward enhanced capability for detecting and predicting dust events in the western United  
17 States: The Arizona case study. *Atmos. Chem. Phys.*, 15, 12 595–12 610, doi:10.5194/acp-15-  
18 12595-2015. 2015.
- 19 Kain, J. S., The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, 43, 170–  
20 181. doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.
- 21 Kim, Y.J., Eckermann, S.D. and Chun, H.Y., An overview of the past, present and future of  
22 gravity-wave drag parameterization for numerical climate and weather prediction models.  
23 *Atmosphere-Ocean*, 41(1), 65-98., doi: 10.3137/ao.410105, 2003.
- 24 Krueger, S. K., Q. Fu, K. N. Liou, and H-N. S. Chin. Improvement of an ice-phase microphysics  
25 parameterization for use in numerical simulations of tropical convection. *Journal of Applied*  
26 *Meteorology*, 34:281–287, January 1995. <https://doi.org/10.1175/1520-0450-34.1.281>
- 27 Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W.  
28 D.. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative  
29 transfer models. *Journal of Geophysical Research*, 113, D13103. 2008.  
30 <https://doi.org/10.1029/2008JD009944>
- 31 Jimenez, P. A., Dudhia, J., Gonzalez-Rouco, J. F., Navarro, J., Montavez, J. P., & Garcia-  
32 Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Monthly*  
33 *Weather Review*, 140(3), 898–918. <https://doi.org/10.1175/MWR-D-11-00056.1>
- 34 Lin, Y.-L., R. D. Farley, and H. D. Orville. Bulk parameterization of the snow field in a cloud  
35 model. *J. Climate Appl. Meteor.*, 22:1065–1092, 1983. [https://doi.org/10.1175/1520-  
36 0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2)
- 37 Lord, S. J., H.E. Willoughby, and J.M. Piotrowicz. Role of a parameterized ice-phase  
38 microphysics in an axisymmetric, nonhydrostatic tropical cyclone model. *J. Atmos. Sci.*,  
39 41(19):2836–2848, October 1984. [https://doi.org/10.1175/1520-  
40 0469\(1984\)041<2836:ROAPIP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1984)041<2836:ROAPIP>2.0.CO;2)



- 1 Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C.,  
2 and Robinson, A. L.: Simulation of organic aerosol formation during the CalNex study:  
3 updated mobile emissions and secondary organic aerosol parameterization for intermediate-  
4 volatility organic compounds, *Atmos. Chem. Phys.*, 20, 4313–4332,  
5 <https://doi.org/10.5194/acp-20-4313-2020>, 2020
- 6 Luecken, D. J., Yarwood, G., and Hutzell, W. T.: Multipollutant modeling of ozone, reactive  
7 nitrogen and HAPs across the continental US with CMAQ-CB6, *Atmos. Environ.*, 201, 62–  
8 72, <https://doi.org/10.1016/j.atmosenv.2018.11.060>, 2019.
- 9 Mlawer, E. J., Taubman S. J., Brown P. D., Iacono M. J., and Clough S. A., Radiative transfer  
10 for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J.*  
11 *Geophys. Res.*, 102, 16663–16682. 1997. <https://doi.org/10.1029/97JD00237>.
- 12 Monin, A. S., & Obukhov, A. M. Basic laws of turbulent mixing in the surface layer of the  
13 atmosphere (in Russian). Contribution Geophysics Institute, Academy of Sciences USSR,  
14 151, 163–187. 1954.
- 15 Morrison, H., G. Thompson, V. Tatarskii, 2009: Impact of Cloud Microphysics on the  
16 Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of  
17 One- and Two-Moment Schemes. *Mon. Wea. Rev.*, 137, 991–1007.  
18 <https://doi.org/10.1175/2008MWR2556.1>
- 19 National Emissions Inventory Collaborative (2019). 2016v1 Emissions Modeling Platform.  
20 Retrieved from <http://views.cira.colostate.edu/wiki/wiki/10202>.
- 21 Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A.V., Kostelich, E.J., Corazza, M., Kalnay, E., Patil,  
22 D.J. and Yorke, J.A., 2004. A local ensemble Kalman filter for atmospheric data  
23 assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 56(5), pp.415-428.  
24 <https://doi.org/10.3402/tellusa.v56i5.14462>
- 25 Otte, T. L., J E. Pleim, and G Pouliot. PREMAQ: A new pre-processor to cmaq for air-quality  
26 forecasting. Presented at 2004 Models-3 Conference, Chapel Hill, NC, October 18-20, 2004.
- 27 Otte, T. L. and Pleim, J. E.: The Meteorology-Chemistry Interface Processor (MCIP) for the  
28 CMAQ modeling system: updates through MCIPv3.4.1, *Geosci. Model Dev.*, 3, 243–256,  
29 <https://doi.org/10.5194/gmd-3-243-2010>, 2010.
- 30 Pan, L., Kim, H., Lee, P., Saylor, R., Tang, Y., Tong, D., Baker, B., Kondragunta, S., Xu, C.,  
31 Ruminski, M. G., Chen, W., McQueen, J., and Stajner, I.: Evaluating a fire smoke simulation  
32 algorithm in the National Air Quality Forecast Capability (NAQFC) by using multiple  
33 observation data sets during the Southeast Nexus (SENEX) field campaign, *Geosci. Model*  
34 *Dev.*, 13, 2169–2184, <https://doi.org/10.5194/gmd-13-2169-2020>, 2020.
- 35 Parrish, D. D., Ryerson, T. B., Mellqvist, J., Johansson, J., Fried, A., Richter, D., Walega, J. G.,  
36 Washenfelder, R. A., de Gouw, J. A., Peischl, J., Aikin, K. C., McKeen, S. A., Frost, G. J.,  
37 Fehsenfeld, F. C., and Herndon, S. C.: Primary and secondary sources of formaldehyde in  
38 urban atmospheres: Houston Texas region, *Atmos. Chem. Phys.*, 12, 3273–3288,  
39 <https://doi.org/10.5194/acp-12-3273-2012>, 2012.



- 1 Peterson, D.A., Fromm, M.D., McRae, R.H., Campbell, J.R., Hyer, E.J., Taha, G., Camacho,  
2 C.P., Kablick, G.P., Schmidt, C.C. and DeLand, M.T., Australia's Black Summer  
3 pyrocumulonimbus super outbreak reveals potential for increasingly extreme stratospheric  
4 smoke events. *npj Climate and Atmospheric Science*, 4(1), 2021.  
5 <https://doi.org/10.1038/s41612-021-00192-9>
- 6 Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and  
7 Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial  
8 uncertainty, *SOIL*, 7, 217–240, 2021. <https://doi.org/10.5194/soil-7-217-2021>
- 9 Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L.,  
10 Gochis, D.J., Ahmadov, R., Peckham, S.E. and Grell, G.A., The weather research and  
11 forecasting model: Overview, system efforts, and future directions. *Bulletin of the American*  
12 *Meteorological Society*, 98(8), pp.1717-1737. 2017. [https://doi.org/10.1175/BAMS-D-15-](https://doi.org/10.1175/BAMS-D-15-00308.1)  
13 [00308.1](https://doi.org/10.1175/BAMS-D-15-00308.1)
- 14 Putman, W. M. and S.-J. Lin, Finite-volume transport on various cubed-sphere grids, *Journal of*  
15 *Computational Physics*, 227(1), 55-78, <https://doi.org/10.1016/j.jcp.2007.07.022>, 2007.
- 16 Qian, Y., Yang, Z., Feng, Z., Liu, Y., Gustafson, W.I., Berg, L.K., Huang, M., Yang, B. and Ma,  
17 H.Y., Neglecting irrigation contributes to the simulated summertime warm-and-dry bias in the  
18 central United States. *npj Clim Atmos Sci* 3, 31. [https://doi.org/10.1038/s41612-020-00135-](https://doi.org/10.1038/s41612-020-00135-w)  
19 [w](https://doi.org/10.1038/s41612-020-00135-w), 2020
- 20 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Huang, X.-Yu.. A  
21 Description of the Advanced Research WRF Model Version 4. NCAR Tech Note,  
22 NCAR/TN-556+STR. doi:10.5065/1dfh-6p97, 2021.
- 23 Skamarock, W.C., Snyder, C., Klemp, J.B. and Park, S.H., vertical resolution requirements in  
24 atmospheric simulation. *Monthly Weather Review*, 147(7), 2641-2656.  
25 <https://doi.org/10.1175/MWR-D-19-0043.1>, 2019.
- 26 Shi, Y., Hu, F., Xiao, Z., Fan, G. and Zhang, Z., Comparison of four different types of planetary  
27 boundary layer heights during a haze episode in Beijing. *Science of the total environment*,  
28 711, <https://doi.org/10.1016/j.scitotenv.2019.134928>, 2020.
- 29 Sillman, S., He, D., Cardelino, C., and Imhoff, R. E.: The use of photochemical indicators to  
30 evaluate ozone-NO<sub>x</sub> –hydrocarbon sensitivity: Case studies from Atlanta, New York, and Los  
31 Angeles, *J. Air Waste Manage. Assoc.*, 47, 1030–1040,  
32 doi:10.1080/10962247.1997.11877500, 1997.
- 33 Sofiev, M., Ermakova, T., and Vankevich, R.: Evaluation of the smoke-injection height from  
34 wild-land fires using remote-sensing data, *Atmos. Chem. Phys.*, 12, 1995–2006,  
35 <https://doi.org/10.5194/acp-12-1995-2012>, 2012.
- 36 Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., et al. Implementation  
37 and verification of the unified NOAA land surface model in the WRF model. Paper Presented  
38 at the 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical  
39 Weather Prediction, Seattle, WA, 2004.



- 1 Wang, S., Coggon, M.M., Gkatzelis, G.I., Warneke, C., Bourgeois, I., Ryerson, T., Peischl, J.,  
2 Veres, P.R., Neuman, J.A., Hair, J. and Shingler, T., Chemical Tomography in a Fresh  
3 Wildland Fire Plume: a Large Eddy Simulation (LES) Study. *Journal of Geophysical*  
4 *Research: Atmospheres*, 126(18), <https://doi.org/10.1029/2021JD035203>. 2021
- 5 Yamartino, R. J., Nonnegative, conserved scalar transport using grid-cell-centered, spectrally  
6 constrained Blackman cubics for applications on a variable-thickness mesh. *Mon. Wea. Rev.*  
7 121, 753-763, 1999. [https://doi.org/10.1175/1520-493\(1993\)121<0753:NCSTUG>2.0.CO;2](https://doi.org/10.1175/1520-493(1993)121<0753:NCSTUG>2.0.CO;2)
- 8 Yang, F., Tallapragada, V., Kain, J. S., Wei, H., Yang, R., Yudin, V. A., Moorthi, S., Han, J.,  
9 Hou, Y. T., Wang, J., Treadon, R., and Kleist, D. T. (2020). Model Upgrade Plan and Initial  
10 Results from a Prototype NCEP Global Forecast System Version 16. 2020 AMS Conference,  
11 Boston, MA. <https://ams.confex.com/ams/2020Annual/webprogram/Paper362797.html>
- 12 Yarwood, G., Jung, J., Whitten, G.Z., Heo, G., Mellberg, J. and Estes, M., Updates to the Carbon  
13 Bond mechanism for version 6 (CB6). In 9th Annual CMAS Conference, Chapel Hill, NC  
14 (pp. 11-13). 2010.
- 15 Yarwood, Y., Sakulyanontvittaya, T., Nopmongkol, O., and Koo, K.: Ozone depletion by  
16 bromine and iodine over the Gulf of Mexico, final report for the Texas Commission on  
17 Environmental Quality, available at:  
18 [https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/pm/582111](https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/pm/5821110365FY1412-20141109-enviro-bromine.pdf)  
19 [0365FY1412-20141109-enviro-bromine.pdf](https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/pm/5821110365FY1412-20141109-enviro-bromine.pdf) (last access: 3 May 2021), November 2014.
- 20 Ye, X., Arab, P., Ahmadov, R., James, E., Grell, G. A., Pierce, B., Kumar, A., Makar, P., Chen,  
21 J., Davignon, D., Carmichael, G. R., Ferrada, G., McQueen, J., Huang, J., Kumar, R.,  
22 Emmons, L., Herron-Thorpe, F. L., Parrington, M., Engelen, R., Peuch, V.-H., da Silva, A.,  
23 Soja, A., Gargulinski, E., Wiggins, E., Hair, J. W., Fenn, M., Shingler, T., Kondragunta, S.,  
24 Lyapustin, A., Wang, Y., Holben, B., Giles, D. M., and Saide, P. E.: Evaluation and  
25 intercomparison of wildfire smoke forecasts from multiple modeling systems for the 2019  
26 Williams Flats fire, *Atmos. Chem. Phys.*, 21, 14427–14469, [https://doi.org/10.5194/acp-21-](https://doi.org/10.5194/acp-21-14427-2021)  
27 [14427-2021](https://doi.org/10.5194/acp-21-14427-2021), 2021.
- 28 Yu, S., Mathur, R., Pleim, J., Pouliot, G., Wong, D., Eder, B., Schere, K., Gilliam, R. and Rao,  
29 S.T. Comparative evaluation of the impact of WRF/NMM and WRF/ARW meteorology on  
30 CMAQ simulations for PM<sub>2.5</sub> and its related precursors during the 2006  
31 TexAQS/GoMACCS study. *Atmospheric Chemistry and Physics*, 12(9), doi:10.5194/acp-12-  
32 4091-2012, 2012a.
- 33 Yu, S., Mathur, R., Pleim, J., Pouliot, G., Wong, D., Eder, B., Schere, K., Gilliam, R. and Rao,  
34 S.T., 2012. Comparative evaluation of the impact of WRF–NMM and WRF–ARW  
35 meteorology on CMAQ simulations for O<sub>3</sub> and related species during the 2006  
36 TexAQS/GoMACCS campaign. *Atmospheric Pollution Research*, 3(2), pp.149-162, doi:  
37 10.5094/APR.2012.015, 2012b.
- 38 Zhang, X., and Kondragunta, S., 2006, Estimating forest biomass in the USA using generalized  
39 allometric model and MODIS land data, *Geographical Research Letter*, 33, L09402,  
40 doi:10.1029/2006GL025879.



- 1 Zhang, X., Kondragunta, S., and Quayle, B. Estimation of biomass burned areas using multiple-
- 2 satellite-observed active fires. *IEEE Transactions on Geosciences and Remote Sensing*, 49:
- 3 4469-4482, 10.1109/TGRS.2011.2149535, 2011.
- 4 Zhuang, J., Jacob, D. J., and Eastham, S. D.: The importance of vertical resolution in the free
- 5 troposphere for modeling intercontinental plumes, *Atmos. Chem. Phys.*, 18, 6039–6055,
- 6 <https://doi.org/10.5194/acp-18-6039-2018>, 2018.
- 7



1 Table 1. The two meteorological datasets used in this study

Model Settings	FV3-GFSv16/NACC	WRF-ARW/MCIP
Domain	Global C768L127 (~ 13 km horizontal resolution in 6 cubic spherical tiles , 127 vertical layers up to 80km), interpolated to the 12km CONUS domain with 35-layers up to about 14km (60hPa)	12km CONUS 35 vertical layers up to 100hPa
Dynamic core	Finite Volume 3, non-hydrostatic (Putman and Lin, 2007)	WRF-ARW dynamic in hybrid vertical coordinate (Skamarock et al., 2021)
Initial condition	FV3-GFSv16 analysis (GDAS) using the local ensemble Kalman filter (LETKF) (Ott et al., 2004) with 4-dimensional incremental analysis update (4D-IAU)	FV3-GFSv16 analysis (GDAS)
Lateral Boundary Condition	N/A	FV3-GFSv16 analysis (GDAS)
Cloud Microphysics	GFDL six-category cloud microphysics scheme (Lin et al., 1983; Lord et al., 1984; Krueger et al., 1995; Chen and Lin, 2011; Chen and Lin, 2013)	Morrison 2-moment scheme (Morrison et al., 2009)
PBL Physics Scheme	Scale-aware (sa) turbulent kinetic energy (TKE) -based moist eddy-diffusivity mass-flux (EDMF) (sa-TKE-EDMF) (Han and Bretherton, 2019)	Yonsei University Scheme (Hong et al., 2006)
Shallow/Deep Cumulus Parameterization	SAS Scheme (Han et al. 2011; 2017)	Kain Fritsch multiscale (Kain, 2004)
Shortwave and Longwave Radiation	RRTMG (Mlawer et al. 1997; Clough et al. 2005; Iacono et al. 2008)	RRTMG (Iacono et al. 2008).
Land Surface Model	Noah Land Surface Model (Chen and Dudhia 2001; Ek et al. 2003; Tewari et al. 2004)	Noah (Tewari et al., 2004)
Surface Layer	Monin-Obukhov (Monin-Obukhov 1954; Grell et al. 1994; Jimenez et al. 2012)	Revised MM5 Scheme (Jimenez et al., 2012)
Other treatment		FDDA nudging is enabled for temperature and specific humidity whole domain, and for wind components (U, V) outside the PBL.

2



Variables	Obs Mean	GFS-CMAQ					WRF-CMAQ				
		MB	NMB	RMSE	R	Slope	MB	NMB	RMSE	R	Slope
Temperature (K)	295	0.979	0.332	2.04	0.988	1.13	1.16	0.393	2.28	0.989	1.17
RH (%)	35.6	-7.3	-20.5	11.8	0.781	0.717	-6.05	-17	12.6	0.677	0.598
Wind Speed (m/s)	4.81	0.758	15.8	3.25	0.432	0.473	-1.11	-23.1	2.4	0.666	0.524
O <sub>3</sub> (ppbv)	57.9	-10.7	-18.5	15	0.651	0.34	-10.4	-17.9	14.1	0.717	0.413
CO (ppbv)	134	-37.6	-28	53.2	0.654	0.573	-37.1	-27.7	52.9	0.652	0.572
NO <sub>x</sub> (ppbv)	1.11	0.507	45.6	2.9	0.704	1.15	0.345	31.1	2.86	0.695	1.12
NO <sub>y</sub> (ppbv)	2.56	-0.0418	-1.63	3.07	0.743	0.892	0.055	2.15	3.14	0.724	0.86
NO <sub>z</sub> (ppbv)	1.63	-0.465	-28.6	1.17	0.782	0.553	-0.125	-7.66	1.08	0.788	0.721
HONO (ppbv)	0.00432	0.012	279	0.0438	0.379	0.444	0.0134	311	0.0487	0.358	0.48
HNO <sub>3</sub> (ppbv)	0.291	0.154	53.1	0.421	0.683	1.34	0.337	116	0.65	0.708	1.89
PAN (ppbv)	0.399	-0.251	-63	0.416	0.675	0.221	-0.222	-55.6	0.386	0.681	0.284
NH <sub>3</sub> (ppbv)	3.55	-0.801	-22.6	5.26	0.0481	0.038	-1.58	-44.5	4.37	0.304	0.155
C <sub>2</sub> H <sub>4</sub> (ppbv)	0.121	0.0582	48.1	0.189	0.702	0.869	0.0385	31.9	0.187	0.682	0.836
C <sub>2</sub> H <sub>2</sub> (ppbv)	0.146	-0.0734	-50.3	0.137	0.784	0.496	-0.0696	-47.7	0.137	0.771	0.494
SO <sub>2</sub> (ppbv)	0.342	-0.235	-68.8	0.567	0.0238	0.00835	-0.221	-64.5	0.568	-1.26×10 <sup>-3</sup>	-0.00047
Acetone (ppbv)	2.74	-2.28	-83.1	2.45	0.686	0.192	-2.2	-80.4	2.38	0.668	0.199
HCHO (ppbv)	2.1	-0.972	-46.4	1.26	0.559	0.447	-0.909	-43.4	1.25	0.513	0.442
CH <sub>3</sub> CHO (ppbv)	0.736	-0.326	-44.2	0.538	0.647	0.386	-0.349	-47.4	0.554	0.643	0.38
Benzene (ppbv)	0.0449	-0.0193	-43	0.057	0.398	0.385	-0.0191	-42.6	0.0564	0.397	0.375
Toluene (ppbv)	0.039	0.0409	105	0.153	0.759	1.74	0.0352	90.1	0.14	0.762	1.63
Isoprene (ppbv)	0.073	0.0361	49.4	0.174	0.6	0.838	0.00661	9.06	0.145	0.648	0.797
EC (µg/std m <sup>3</sup> )	0.108	0.191	177	0.572	0.518	2.09	0.228	211	0.609	0.455	1.88
OA (µg/std m <sup>3</sup> )	10.9	-7.15	-65.7	9.72	0.565	0.263	-6.48	-59.5	9.45	0.495	0.243
Sulfate (µg/std m <sup>3</sup> )	1.31	-0.781	-59.7	1.11	0.0856	0.0188	-0.773	-59	1.11	0.0322	0.00677
NH <sub>4</sub> <sup>+</sup> (µg/std m <sup>3</sup> )	0.745	-0.615	-82.5	0.805	0.416	0.103	-0.596	-79.9	0.778	0.509	0.145
Nitrate (µg/std m <sup>3</sup> )	1.22	-1.08	-88.1	1.49	0.562	0.229	-1.04	-85.3	1.45	0.57	0.279
AOE (/Mm)	54.5	-29.3	-53.8	47	0.593	0.227	-27.4	-50.2	45.9	0.588	0.227

Table 2. Statistics of the two models compared to the observation for DC-8 flight segments with non-fire events below 3km (ASL) over west of -100°W. All aerosols are in submicron. The normalized mean bias (NMB) is in unit %.

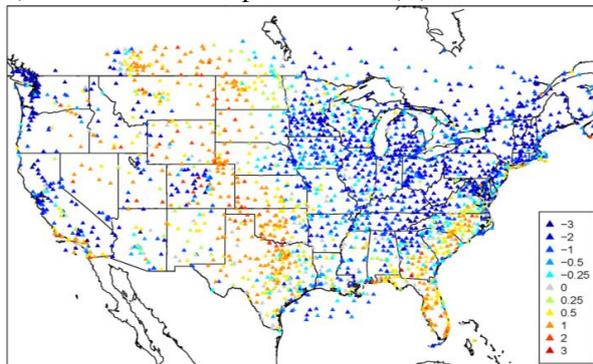


Variables	Obs Mean	GFS-CMAQ					WRF-CMAQ				
		MB	NMB	RMSE	R	Slope	MB	NMB	RMSE	R	Slope
Temperature (K)	287	-0.389	-0.135	0.702	0.995	1.01	-0.688	-0.24	0.863	0.997	1.04
RH (%)	27.8	-0.761	-2.74	7.84	0.712	0.553	4.3	15.5	11.1	0.556	0.534
Wind Speed (m/s)	5.42	0.766	14.1	2.16	0.612	0.616	-0.811	-15	2.12	0.604	0.556
O <sub>3</sub> (ppbv)	55.7	-6.61	-11.9	11.8	0.587	0.262	-7.01	-12.6	11.5	0.653	0.346
CO (ppbv)	486	-377	-77.6	873	0.596	0.0347	-383	-78.8	883	0.442	0.0242
NO <sub>x</sub> (ppbv)	2.63	0.06	2.28	6.41	0.465	0.231	-0.619	-23.5	7.02	0.31	0.153
NO <sub>y</sub> (ppbv)	7.32	-4.19	-57.3	13.3	0.507	0.123	-4.66	-63.7	14.2	0.31	0.073
NO <sub>z</sub> (ppbv)	5.7	-4.8	-84.3	10.2	-0.189	-0.0106	-4.68	-82	10.2	-0.204	-0.0121
HONO (ppbv)	0.283	-0.274	-96.8	1.18	0.355	0.0043	-0.274	-96.8	1.18	0.291	0.00457
HNO <sub>3</sub> (ppbv)	0.148	0.148	99.7	0.256	0.532	1.07	0.179	121	0.28	0.402	0.768
PAN (ppbv)	0.971	-0.793	-81.7	1.63	0.27	0.0195	-0.765	-78.8	1.61	0.279	0.026
NH <sub>3</sub> (ppbv)	17.7	-12.3	-69.3	28.3	0.379	0.0654	-13.7	-77.4	29.6	0.232	0.0386
C <sub>2</sub> H <sub>4</sub> (ppbv)	4.5	-4.34	-96.3	10.2	0.421	0.00498	-4.36	-96.8	10.2	0.14	0.0018
C <sub>2</sub> H <sub>2</sub> (ppbv)	1.04	-1.01	-96.9	2.08	0.534	0.00866	-1.01	-97	2.09	0.363	0.00623
SO <sub>2</sub> (ppbv)	0.699	-0.322	-46.1	1.38	0.589	0.198	-0.392	-56.1	1.5	0.429	0.132
Acetone (ppbv)	3.54	-3.2	-90.3	4.56	0.13	0.00862	-3.18	-89.7	4.55	0.135	0.0112
HCHO (ppbv)	8.17	-7.13	-87.3	17.8	0.232	0.0062	-7.19	-88	17.8	0.119	0.00303
CH <sub>3</sub> CHO (ppbv)	3.65	-3.18	-87.4	9.13	0.186	0.00547	-3.21	-88	9.2	-0.027	-0.00097
Benzene (ppbv)	0.683	-0.67	-98.1	1.84	0.54	0.00432	-0.672	-98.3	1.84	0.367	0.00275
Toluene (ppbv)	0.451	-0.436	-96.6	1.36	0.402	0.00491	-0.438	-97	1.36	0.195	0.00245
Isoprene (ppbv)	0.095	-7.9×10 <sup>-3</sup>	-8.29	0.234	0.123	0.0579	-0.033	-34.7	0.242	-0.014	-0.00541
EC (µg/std m <sup>3</sup> )	1.89	-0.53	-28	3.28	0.612	0.295	-0.787	-41.6	3.7	0.448	0.195
OA (µg/std m <sup>3</sup> )	156	-146	-93.4	420	0.612	0.0174	-147	-94.2	423	0.472	0.0122
Sulfate (µg/std m <sup>3</sup> )	0.791	-0.116	-14.7	0.676	0.415	0.184	-0.214	-27.1	0.728	0.322	0.13
NH <sub>4</sub> <sup>+</sup> (µg/std m <sup>3</sup> )	1	-0.591	-59.1	0.931	0.767	0.351	-0.615	-61.5	0.956	0.729	0.359
Nitrate (µg/std m <sup>3</sup> )	1.7	-0.56	-32.9	1.47	0.805	0.613	-0.634	-37.2	1.59	0.774	0.599
AOE (/Mm)	391	-350	-89.3	994	0.688	0.027	-357	-91.1	1010.	0.532	0.0152

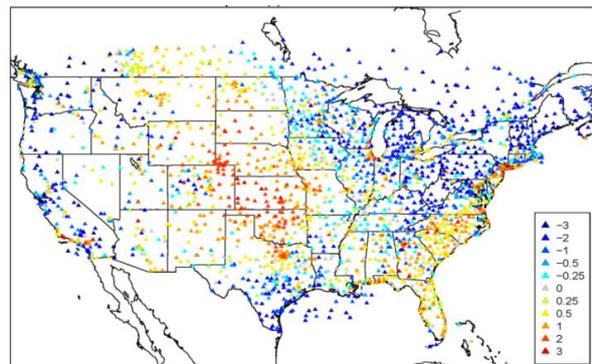
Table 3, same as Table 2 except for the wildfire affected flight segments.



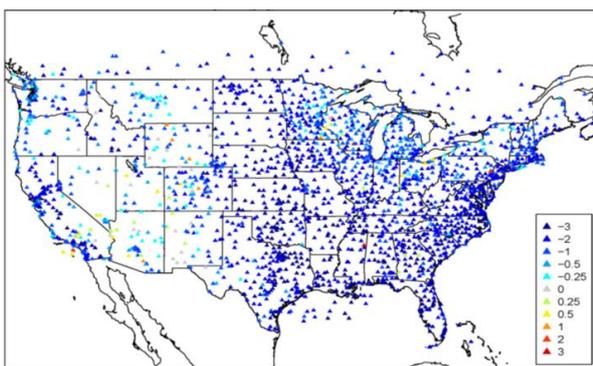
a) GFS mean 2m temperature bias (K), 08/2019



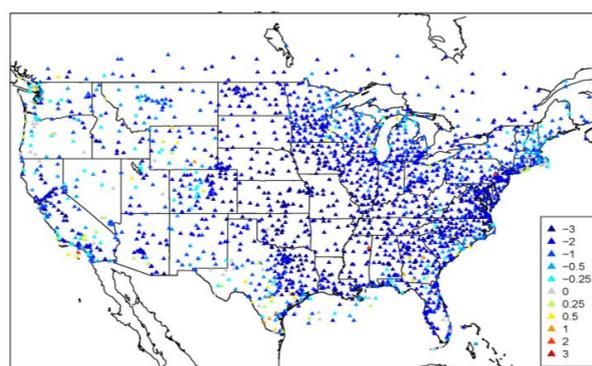
b) WRF mean 2m temperature bias (K), 08/2019



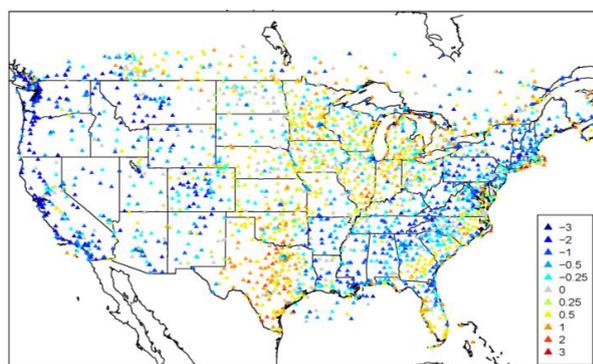
c) GFS mean 2m specific humidity bias (g/kg), 08/2019



d) WRF mean 2m specific humidity bias (g/kg), 08/2019



e) GFS mean 10m wind speed bias (m/s), 08/2019



f) WRF mean 10m wind speed bias (m/s), 08/2019

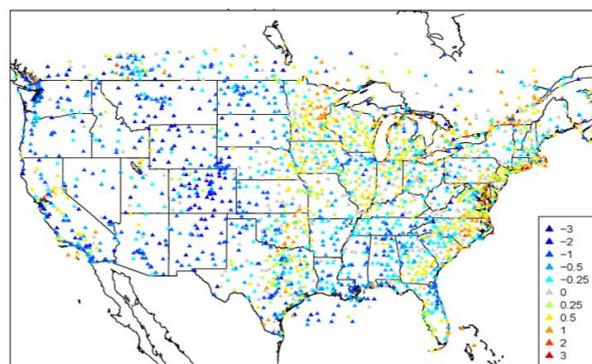
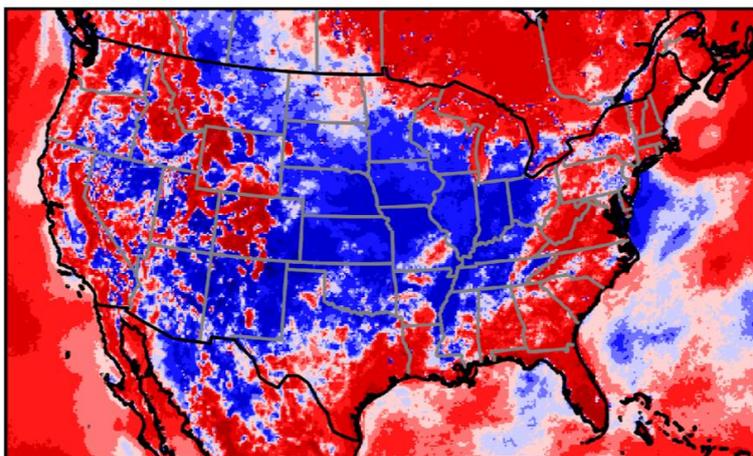


Figure 1. GFS and WRF surface meteorological biases for METAR (METeoro logical Aerodrome Report) stations averaged over August, 2019



a) Monthly Mean PBL Difference (GFS-WRF) at 18UTC



b) Monthly Mean PBL Difference (GFS-WRF) at 06UTC

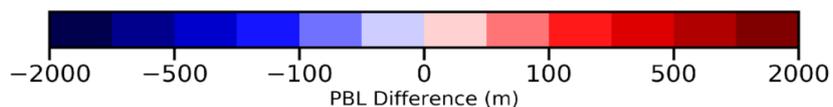
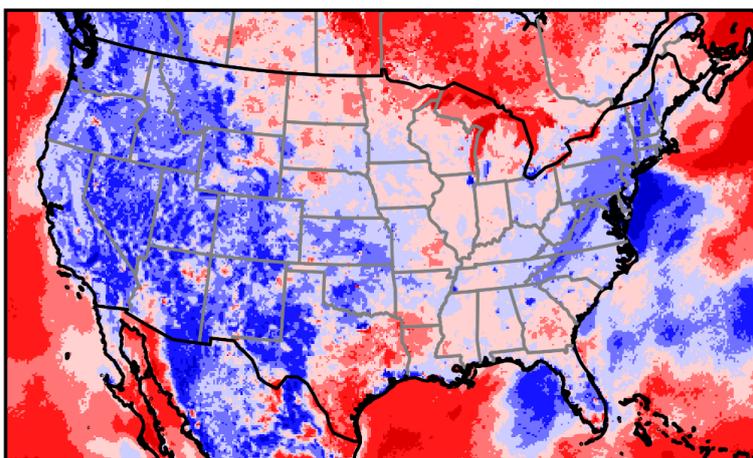


Figure 2. Monthly mean PBL height difference (GFS-WRF) for daytime (a) and nighttime (b), August, 2019.

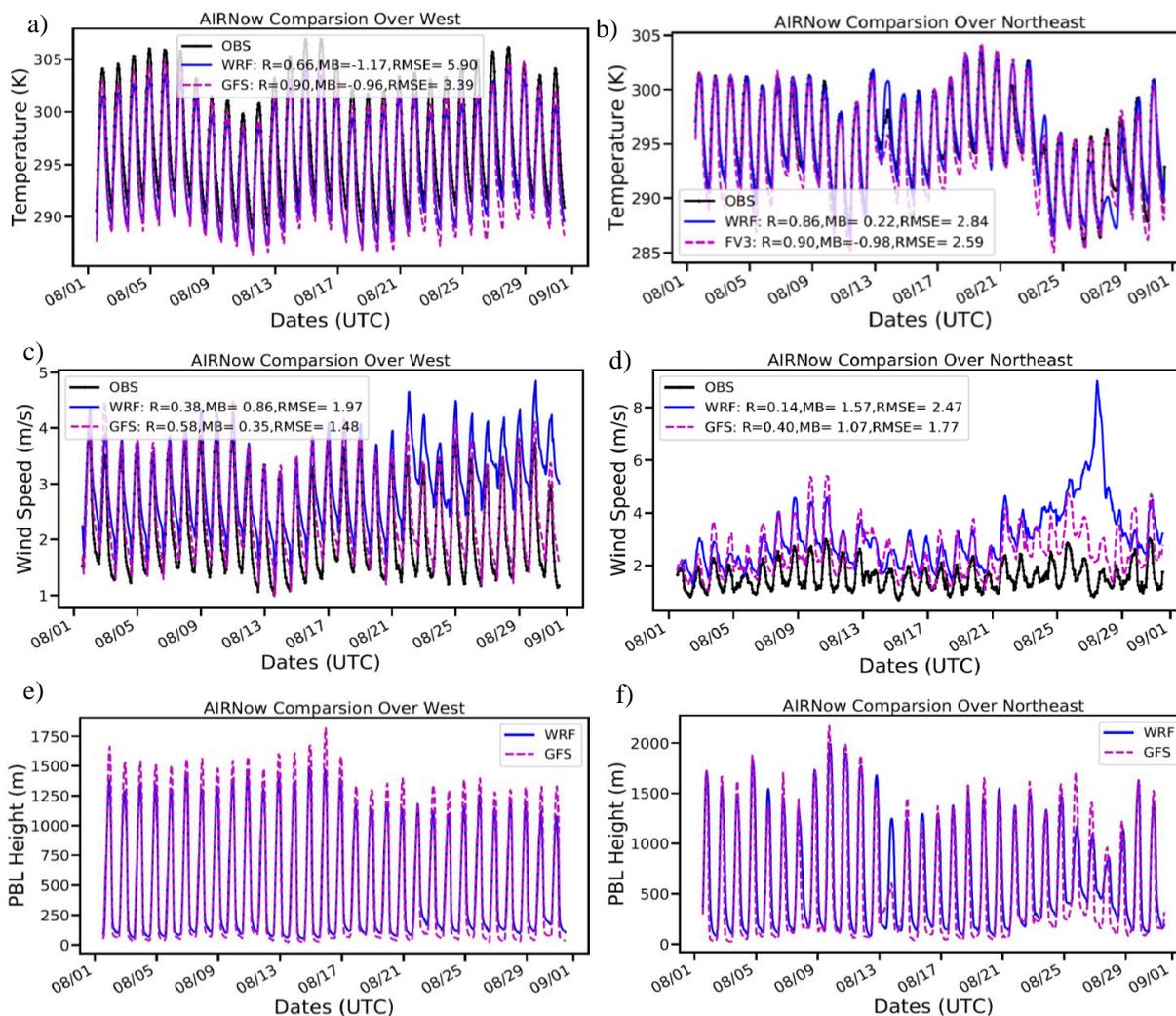


Figure 3. The WRF and GFS time-series comparison over AIRNow stations over the U.S. West and Northeast for 2m temperature (a, b), 10m wind speed (c, d), and PBL height (e,f).

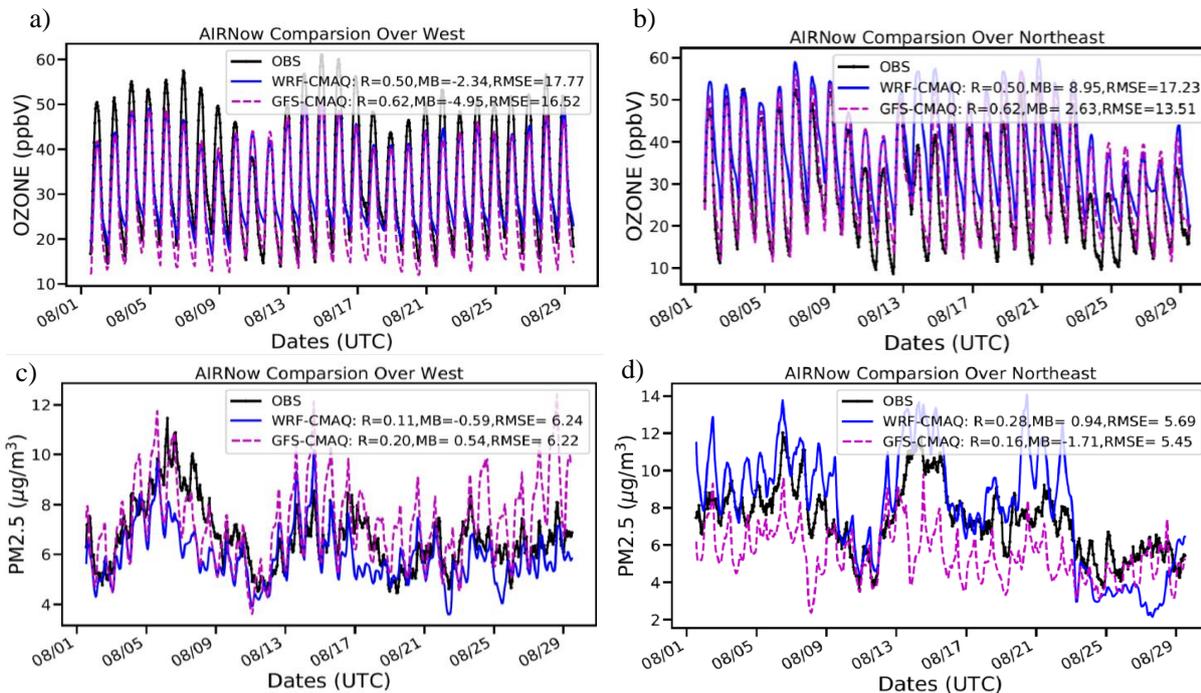


Figure 4. Same as Figure 3 but for ozone (a, b) and PM2.5 (c, d).

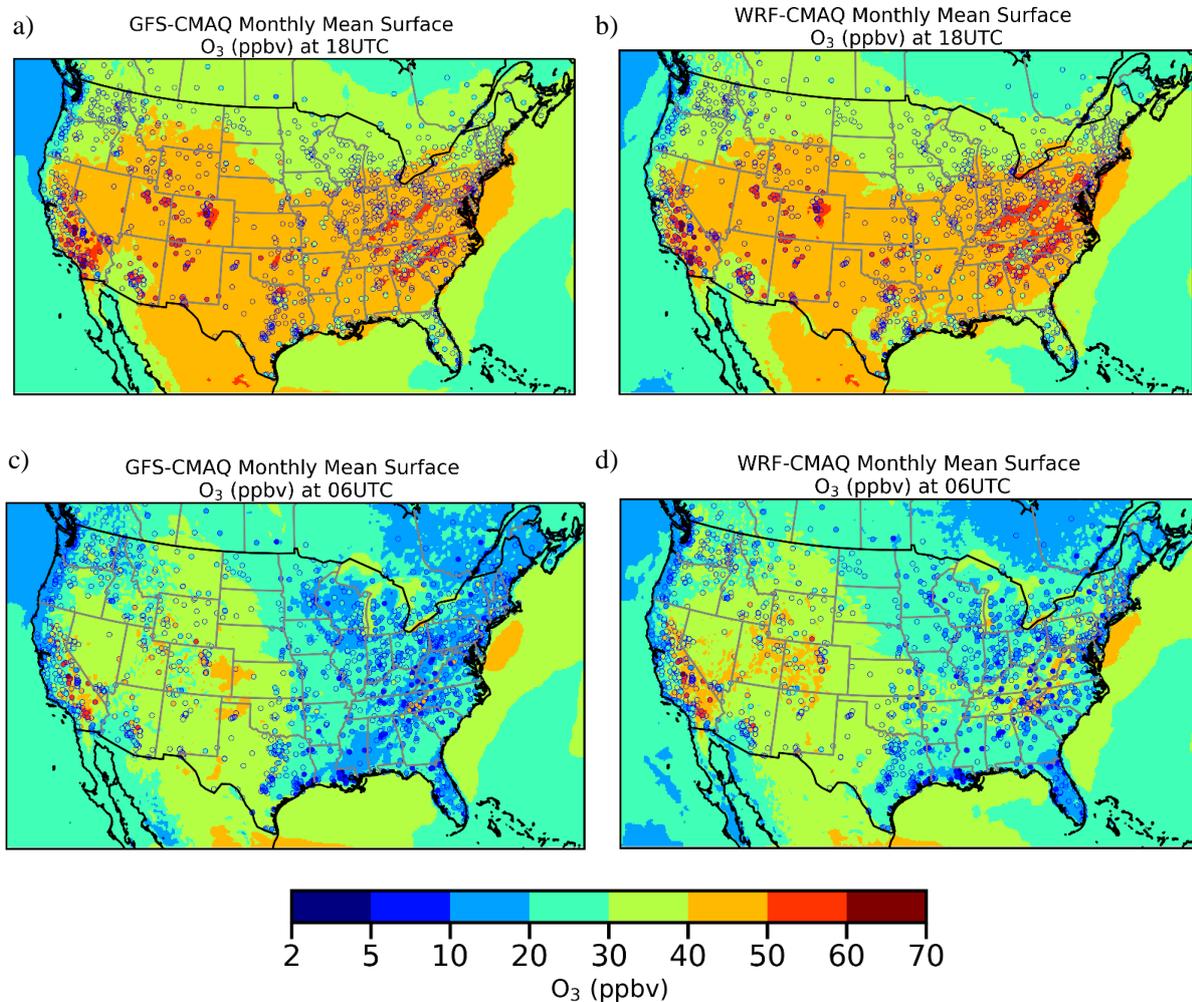


Figure 5. Monthly mean surface ozone predictions by GFS-CMAQ (left plots) and WRF-CMAQ (right plots) for daytime (top plots) and nighttime (bottom plots) compared to the corresponding AIRNow observations, August, 2019

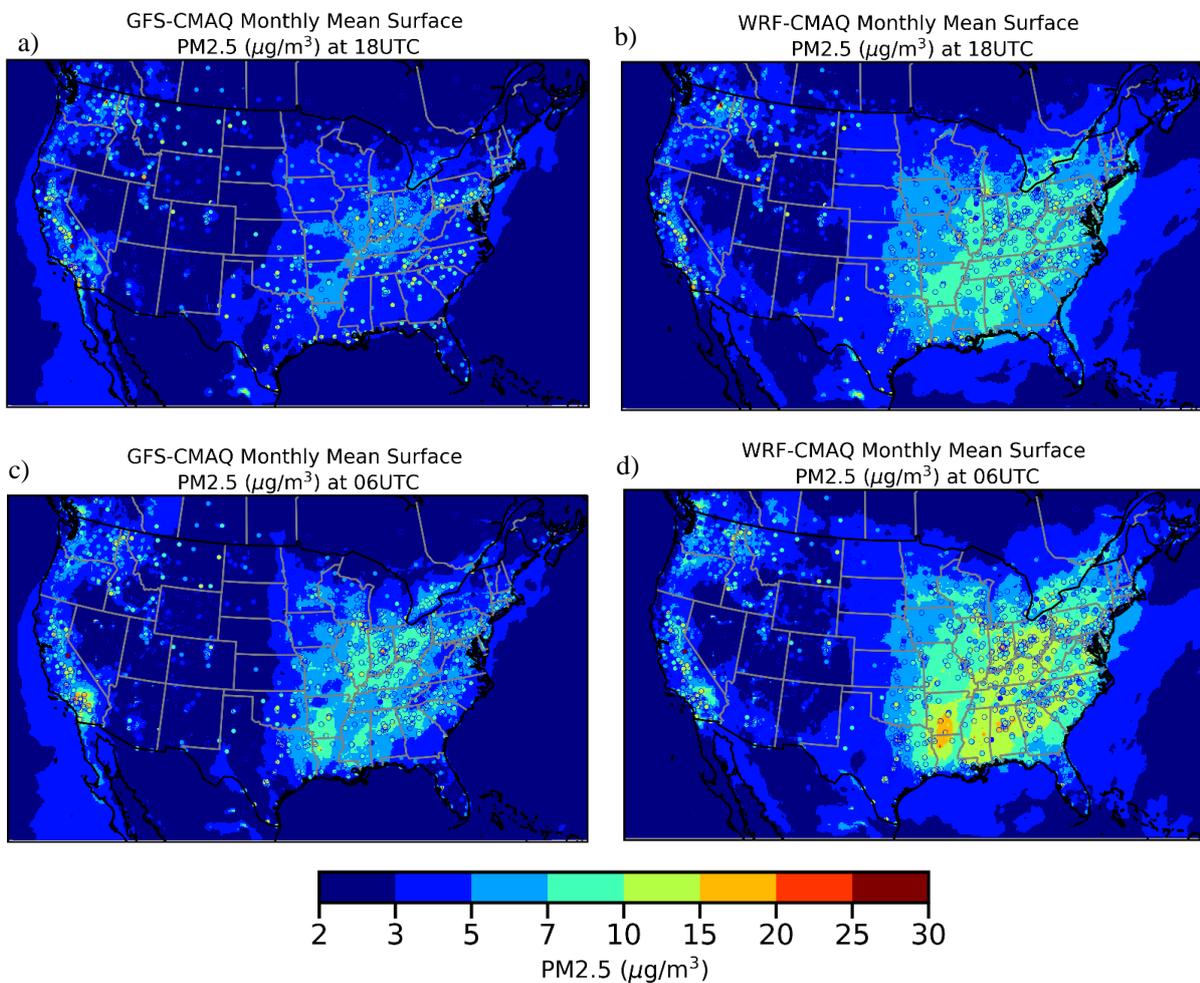


Figure 6, same as figure 5 but for surface PM2.5



a) FIREX-AQ DC-8 Flight Path on 07/22

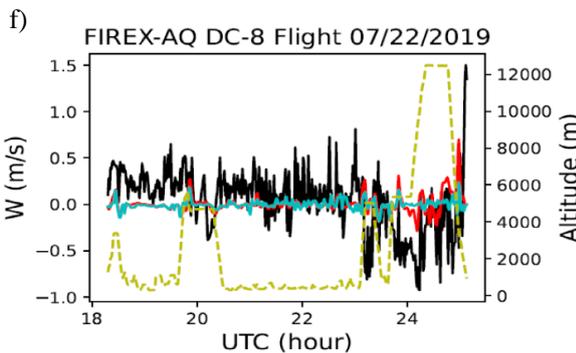
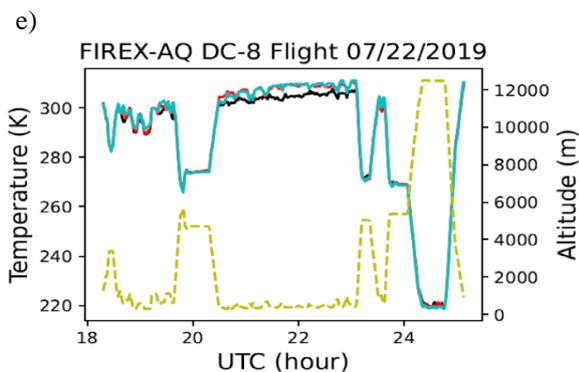
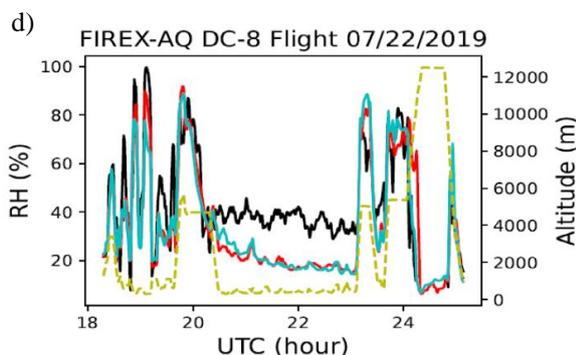
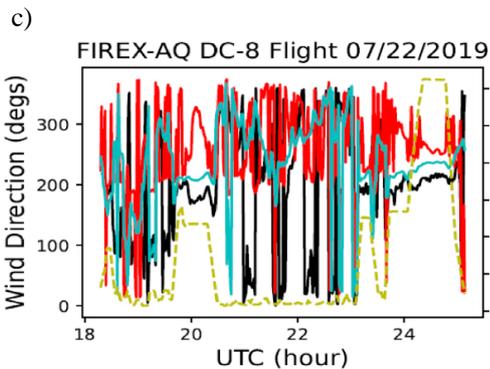
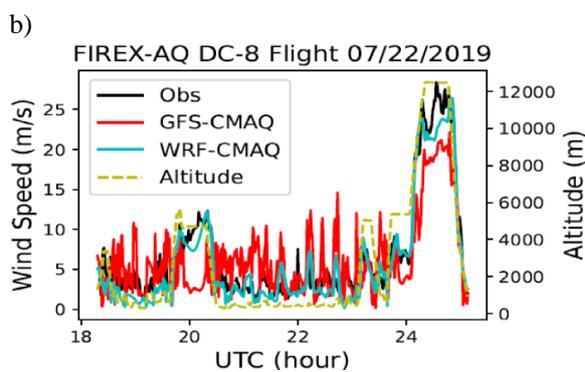
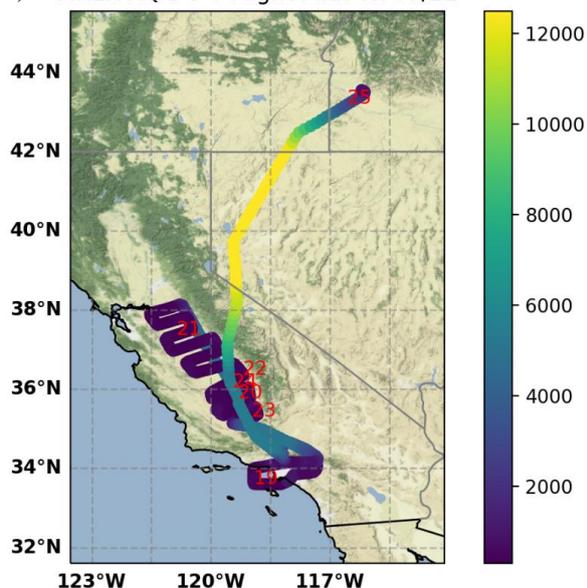


Figure 7. Modeled meteorological variables compared with observations for the DC-8 flight on 22 July, 2019 (b to f). The plot a shows the flight path colored in altitudes above sea level with UTC time in red text. Base map credits: © OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

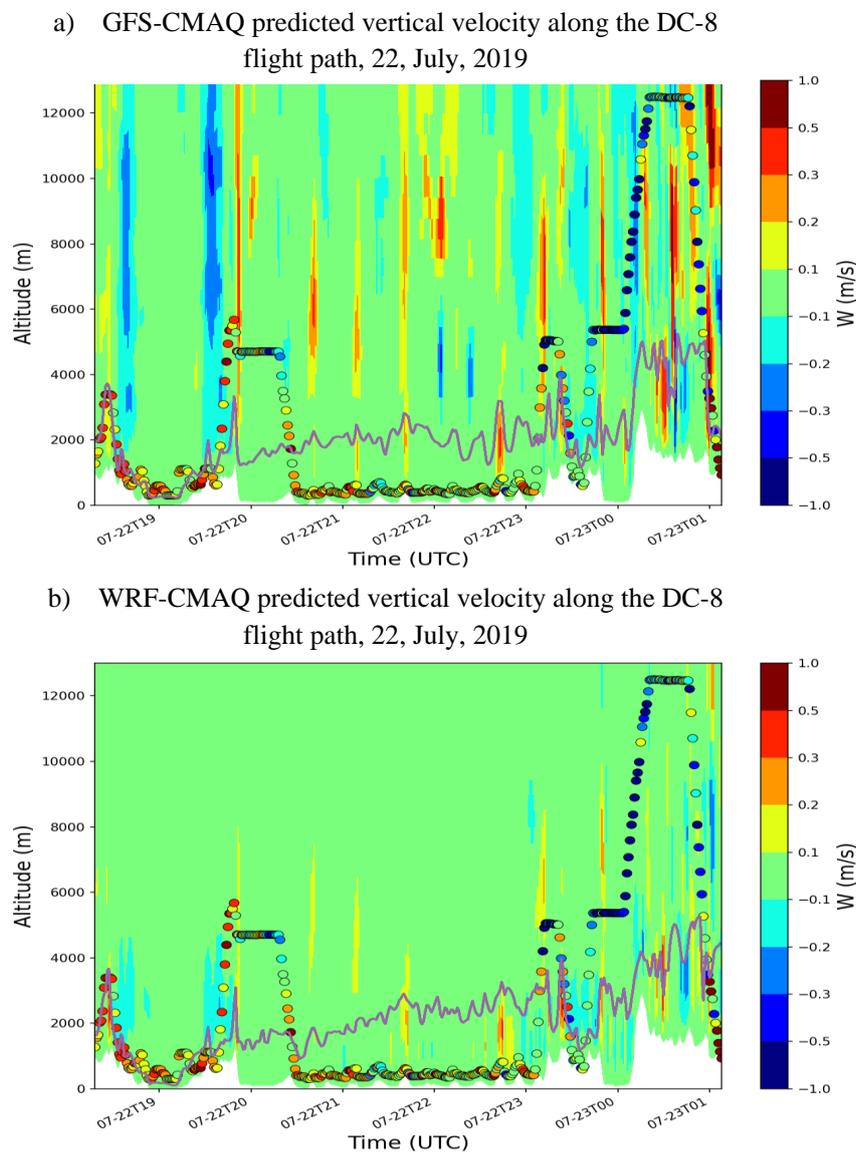


Figure 8, Curtain plots of the vertical velocity ( $W$ ) predicted by GFS-CMAQ (a) and WRF-CMAQ (b) along the DC-8 flight on 22 July, 2019. The colored dots showed the DC-8 measured vertical velocities. The solid lines showed the predicted PBL heights of these two models.

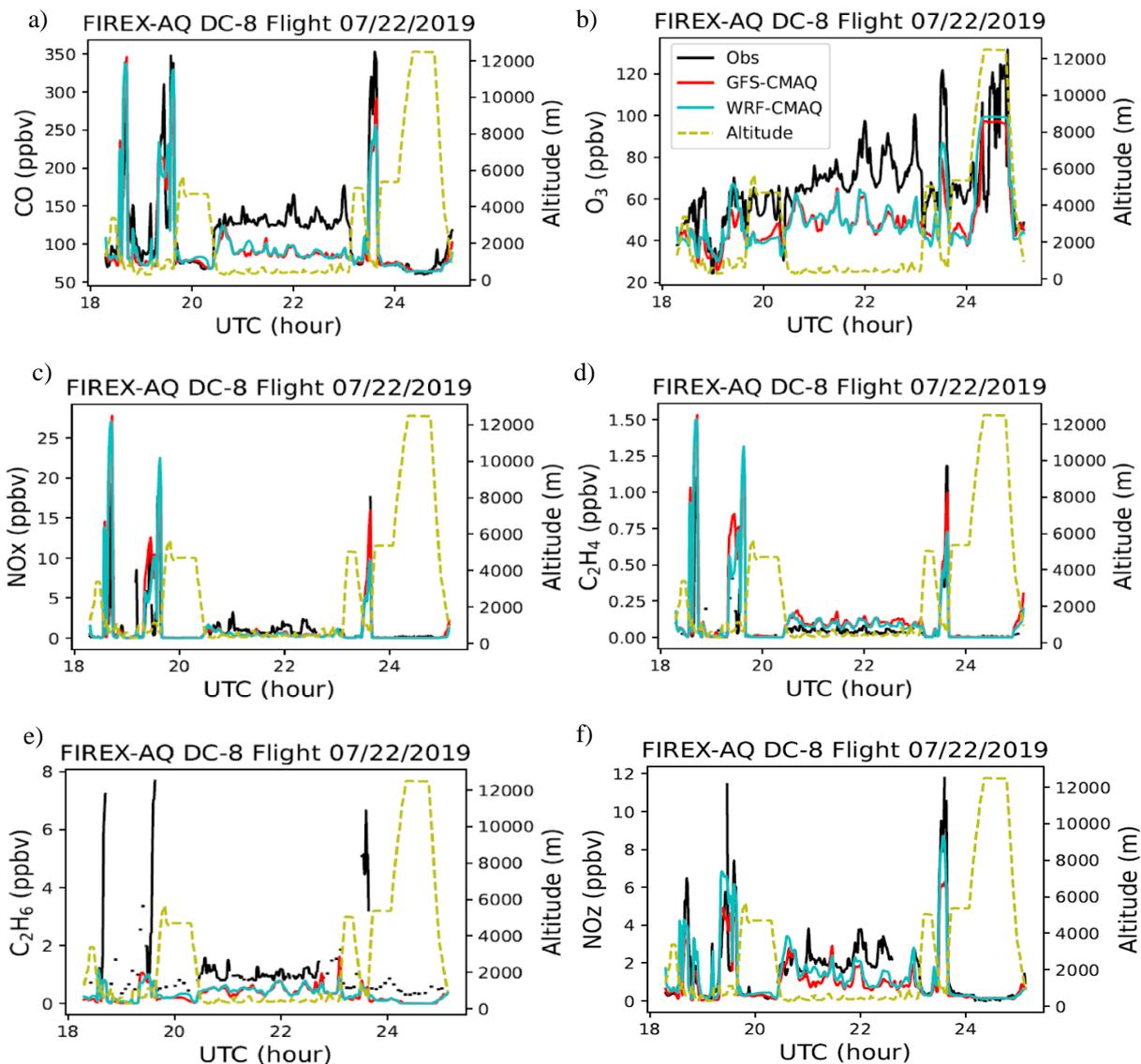


Figure 9. Model predicted chemical concentrations compared with observations along with the DC-8 flight on 22 July, 2019

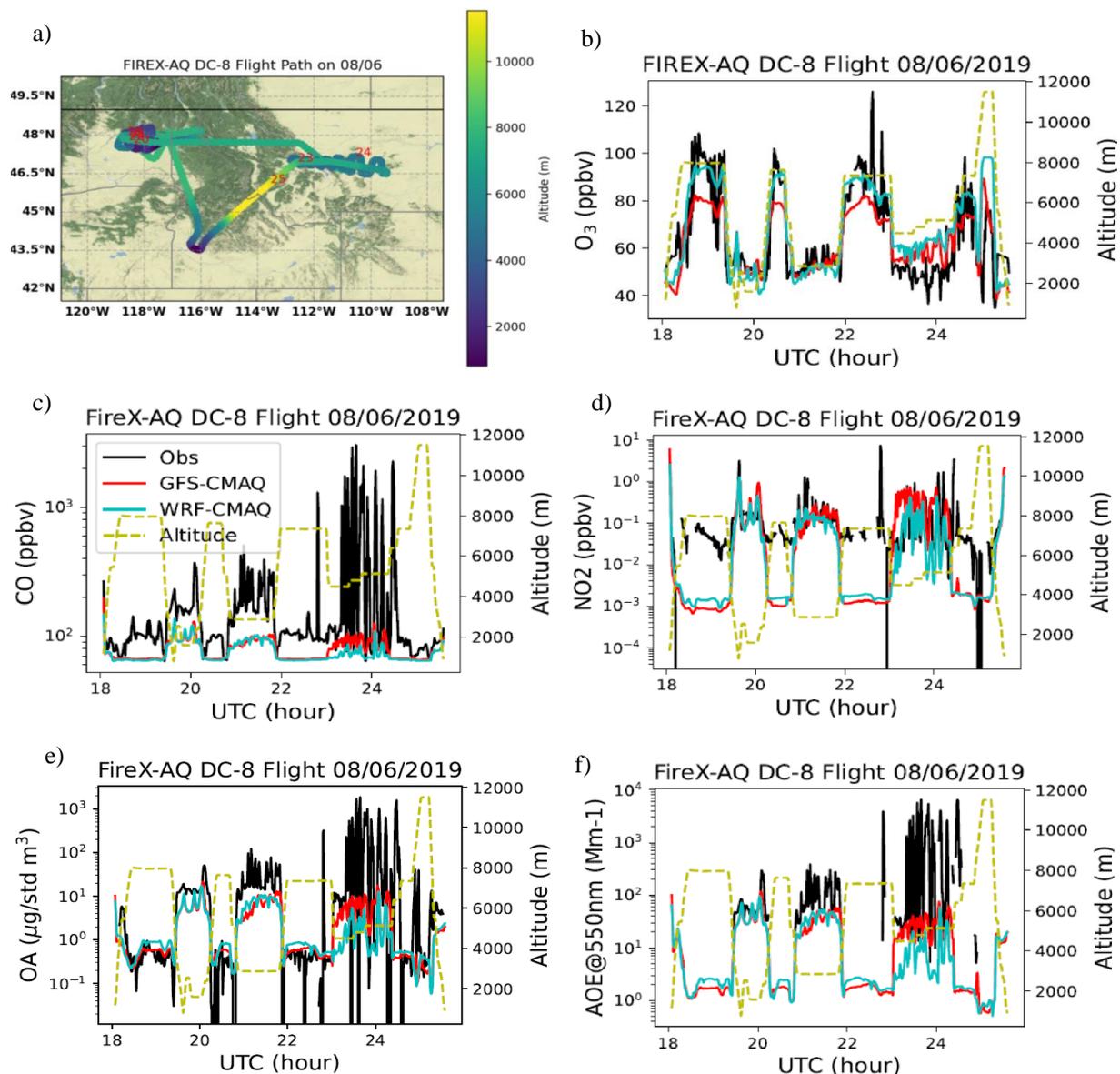


Figure 10. The DC-8 flight path (a), and model-observation comparisons for O<sub>3</sub> (b), CO (c), NO<sub>2</sub> (d), submicron organic aerosol (OA) (e) and aerosol optical extinction coefficient (AOE) at wavelength of 550nm (f) on 06 August, 2019. Base map credits: © OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

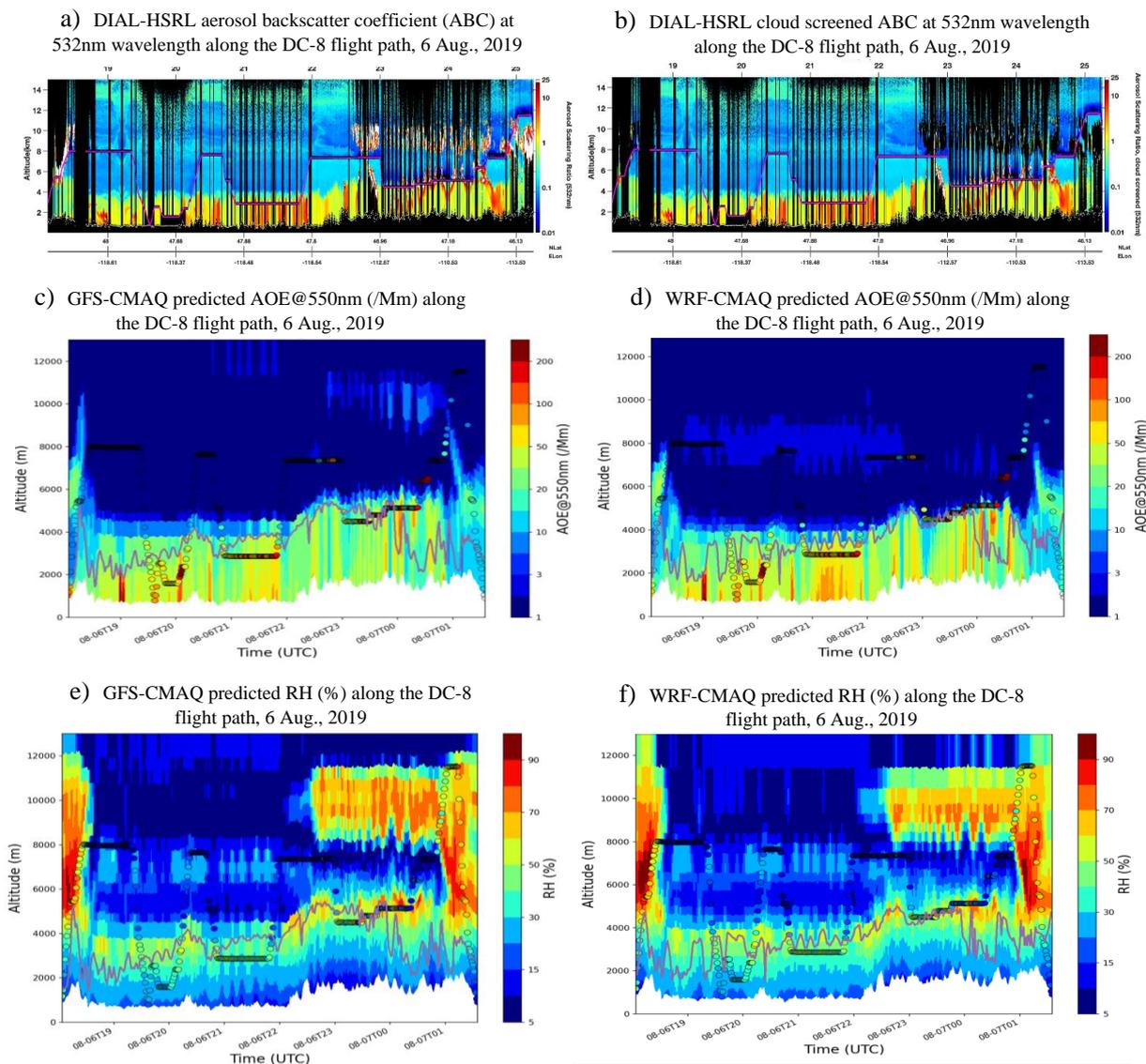


Figure 11, The Differential Absorption High Spectral Resolution Lidar (DIAL-HSRL) retrieved aerosol backscatter coefficient (ABC) at 532nm wavelength in unit /km/steradian (a) and cloud screened one (b); curtain plots of the AOE (b, c) and relative humidity (RH) (d, e) predicted by GFS-CMAQ (left) and WRF-CMAQ (right) along the DC-8 flight on 06 August, 2019. The colored dots showed the corresponding measured values. The solid lines showed the predicted PBL heights of these two models.