Air Quality Modeling Technical Support Document for the Proposed Revised Cross-State Air Pollution Rule Update

Office of Air Quality Planning and Standards
United States Environmental Protection Agency
October 2020

1. Introduction

In this technical support document (TSD) we describe the air quality modeling performed to support the proposed Revised Cross State Air Pollution Rule Update. For this proposed rule, the focus of the air quality modeling is to project ozone design values[1] at individual monitoring sites to 2021[2] and to estimate state-by-state contributions to those 2021 concentrations. The projected 2021 ozone design values are used to identify ozone monitoring sites that are projected to be nonattainment or have maintenance problems in 2021 for the 2008 ozone NAAQS. Ozone contribution information for 2021 is then used to quantify projected interstate contributions from emissions in each upwind state to ozone design values at projected nonattainment and maintenance sites in other states (i.e., in downwind states). This TSD also describes air quality modeling and results for the 2023 and 2028 projection years which were used to support the proposed rule.[3]

The remaining sections of this TSD are as follows. Section 2 describes the air quality modeling platform and the evaluation of model predictions using measured concentrations. Section 3 defines the procedures for projecting ozone design value concentrations and the approach for identifying monitoring sites projected to have nonattainment and/or maintenance problems in 2021. Section 4 describes (1) the source contribution (i.e., apportionment) modeling and (2) the procedures for quantifying contributions to individual monitoring sites including nonattainment and/or maintenance sites. For questions about the information in this TSD please contact Norm Possiel at possiel.norm@epa.gov.

[1] The ozone design value for a monitoring site is the 3-year average of the annual fourth-highest daily maximum 8-hour average ozone concentration.

[2] The rationale for using 2021 as the future analytic year for this transport assessment is described in the preamble for this proposed rule.

[3] The input and output data for the air quality modeling, as described in this TSD, can be found on data drives in the docket for this proposed rule. The contents of the data drives are listed in the following file which is in the docket: AQ Modeling Data Drives_Proposed Revised CSAPR Update.docx.

2. Air Quality Modeling Platform

The EPA used a 2016-based air quality modeling platform which includes emissions, meteorology and other inputs for 2016 as the base year for the modeling described in this document. The emissions were developed as part of the 2016 Platform Collaborative Project that included participation from EPA, Multi-State Jurisdictional Organizations (MJOs) and states. This process resulted in a common-use set of emissions data for a 2016 base year and 2023 and 2028 projection years that can be leveraged by EPA and states for regulatory air quality modeling. The 2016 modeling platform including the projected 2023 and 2028 emissions were used to drive the 2016 base year and 2023 and 2028 base case air quality model simulations for this proposed rule. Because emissions are not available for the 2021 analytic year, we used the 2016-Centered measured ozone design values coupled with 2023 model-predicted design values to estimate design values in 2021, based on linear interpolation between these two data points. To quantify ozone contributions in 2021 we applied modeling-based contributions in 2023 to the 2021 ozone design values. The methods for developing design values and contributions for 2021 are described in sections 3 and 4, below. In addition, we modeled the 2028 base case emissions to project ozone design values and contributions in that year. The projected design values and contribution data were used in Step 3 of the four-step transport framework, as described in the proposed rule. The Step 3 analysis is described in Ozone Transport Policy Analysis Technical Support Document.

2.1 Air Quality Model Configuration

The photochemical model simulations performed for this proposed rule used the Comprehensive Air Quality Model with Extensions (CAMx version 7beta 6).[4,5] CAMx is a three-dimensional grid-based Eulerian air quality model designed to simulate the formation and fate of oxidant precursors, primary and secondary particulate matter concentrations, and deposition over regional and urban spatial scales (e.g., the contiguous U.S.). Consideration of the different

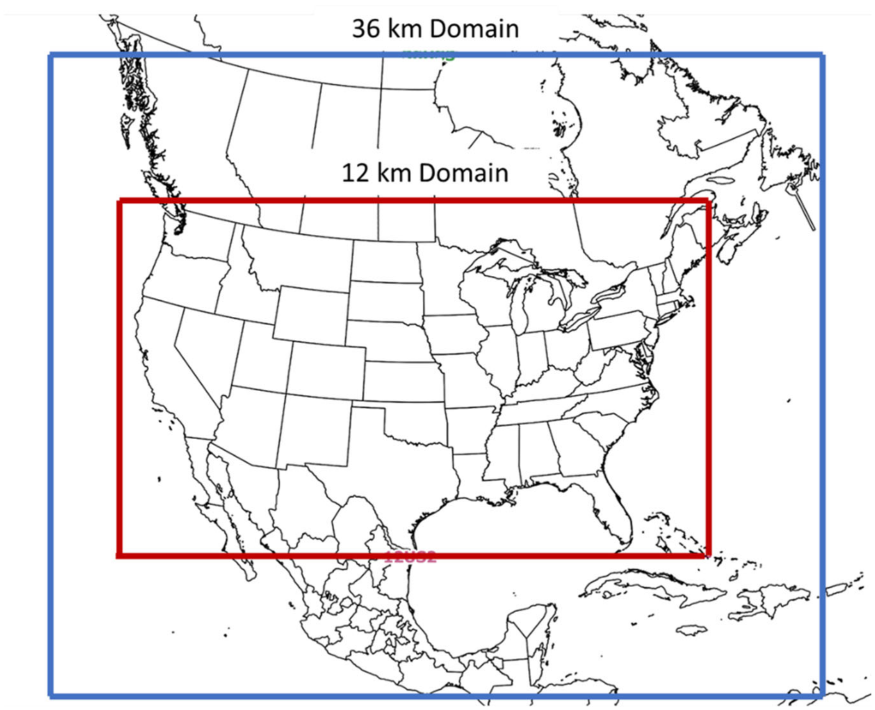
[4] Ramboll Environment and Health, May 2020, www.camx.com. Note that CAMx v7beta6 is a pre-lease of CAMx version 7 that was used by EPA because the official release of version 7 did not occur until May 2020, which was too late for use in the air quality modeling for this proposed rule.

[5] The scripts used for the CAMx model simulations can be found in the following file in the docket: CAMx Model Simulation Scripts.docx

processes (e.g., transport and deposition) that affect primary (directly emitted) and secondary (formed by atmospheric processes) pollutants at the regional scale in different locations is fundamental to understanding and assessing the effects of emissions on air quality concentrations.

Figure 2-1 shows the geographic extent of the modeling domains that were used for air quality modeling in this analysis. The large domain covers the 48 contiguous states along with most of Canada and all of Mexico with a horizontal resolution of 36 x 36 km. Air quality modeling for the 36 km domain was used to provide boundary conditions for the nested 12 km x 12 km domain for the 2016 and projection year emissions scenarios. Both modeling domains have 25 vertical layers with a top at about 17,550 meters, or 50 millibars (mb). The model simulations produce hourly air quality concentrations for each grid cell across each modeling domain.

Figure 2-1. Air quality modeling domains.



CAMx requires a variety of input files that contain information pertaining to the modeling domain and simulation period. These include gridded, hourly emissions estimates and meteorological data, and initial and boundary concentrations. Separate emissions inventories were prepared for the 2016 base year and the 2023 and 2028 projections. All other inputs (i.e. meteorological fields, initial concentrations, and boundary concentrations) were specified for the

2016 base year model application and remained unchanged for the projection-year model simulations.[6]

2.2 Meteorological Data for 2016

The 2016 meteorological data for the air quality modeling were derived from running Version 3.8 of the Weather Research Forecasting Model (WRF) (Skamarock, et al., 2008). The meteorological outputs from WRF include hourly-varying horizontal wind components (i.e., speed and direction), temperature, moisture, vertical diffusion rates, and rainfall rates for each grid cell in each vertical layer. Selected physics options used in the WRF simulations include Pleim-Xiu land surface model (Xiu and Pleim, 2001; Pleim and Xiu, 2003), Asymmetric Convective Model version 2 planetary boundary layer scheme (Pleim 2007a,b), Kain-Fritsch cumulus parameterization (Kain, 2004) utilizing the moisture-advection trigger (Ma and Tan, 2009), Morrison double moment microphysics (Morrison, et al., 2005; Morrison and Gettelman, 2008), and RRTMG longwave and shortwave radiation schemes (Iacono, et.al., 2008).

Both the 36 km and 12 km WRF model simulations utilize a Lambert conformal projection centered at (-97,40) with true latitudes of 33 and 45 degrees north. The 36 km domain contains 184 cells in the X direction and 160 cells in the Y direction. The 12 km domain contains 412 cells in the X direction and 372 cells in the Y direction. The atmosphere is resolved with 35 vertical layers up to 50 mb (see Table 2-1), with the thinnest layers being nearest the surface to better resolve the planetary boundary layer (PBL).

The 36 km WRF model simulation was initialized using the 0.25-degree GFS analysis and 3-hour forecast from the 00Z, 06Z, 12Z, and 18Z simulations. The 12 km model was initialized using the 12km North American Model (12NAM) analysis product provided by National Climatic Data Center (NCDC).[7] The 40km Eta Data Assimilation System (EDAS) analysis (ds609.2) from the National Center for Atmospheric Research (NCAR) was used where

---

[6] The CAMx annual simulations for 2016, 2023, and 2028 were each performed using two time segments (January 1 through April 30, 2011 with a 10-day ramp-up period at the end of December 2010 and May 1 through December 31, 2016 with a 10-day ramp-up period at the end of April 2011). The CAMx 2023 and 2028 contribution modeling was performed for the period May 1 through September 30, 2016 with a 10-day ramp-up period at the end of April 2016.

[7] https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-mesoscale-forecast-system-nam

12NAM data was unavailable.[8] Analysis nudging for temperature, wind, and moisture was applied above the boundary layer only. The model simulations were conducted continuously. The ‘ipxwrf’ program was used to initialize deep soil moisture at the start of the run using a 10-day spinup period (Gilliam and Pleim, 2010). Landuse and land cover data were based on the USGS for the 36NOAM simulation and the 2011 National Land Cover Database (NLCD 2011) for the 12US simulation. Sea surface temperatures were ingested from the Group for High Resolution Sea Surface Temperatures (GHRSST) (Stammer et al., 2003) 1 km SST data.

Additionally, lightning data assimilation was utilized to suppress (force) deep convection where lightning is absent (present) in observational data. This method is described by Heath et al. (2016) and was employed to help improve precipitation estimates generated by the model.

Table 2-1. Vertical layers and their approximate height above ground level.

WRF Layer	Height (m)	Pressure (mb)	Sigma
35	17,556	5000	0.000
34	14,780	9750	0.050
33	12,822	14500	0.100
32	11,282	19250	0.150
31	10,002	24000	0.200
30	8,901	28750	0.250
29	7,932	33500	0.300
28	7,064	38250	0.350
27	6,275	43000	0.400
26	5,553	47750	0.450
25	4,885	52500	0.500
24	4,264	57250	0.550
23	3,683	62000	0.600
22	3,136	66750	0.650
21	2,619	71500	0.700
20	2,226	75300	0.740
19	1,941	78150	0.770
18	1,665	81000	0.800
17	1,485	82900	0.820
16	1,308	84800	0.840
15	1,134	86700	0.860
14	964	88600	0.880
13	797	90500	0.900
12	714	91450	0.910
11	632	92400	0.920
10	551	93350	0.930
9	470	94300	0.940
8	390	95250	0.950

[8] https://www.ready.noaa.gov/edas40.php.

WRF Layer	Height (m)	Pressure (mb)	Sigma
7	311	96200	0.960
6	232	97150	0.970
5	154	98100	0.980
4	115	98575	0.985
3	77	99050	0.990
2	38	99525	0.995
1	19	99763	0.9975
Surface	0	100000	1.000

Details of the annual 2016 meteorological model simulation and evaluation are provided in a separate technical support document which can be found in the docket for this proposed rule.[9]

The meteorological data generated by the WRF simulations were processed using wrfcamx v4.7 (Ramboll 2019) meteorological data processing program to create model-ready meteorological inputs to CAMx. In running wrfcamx, vertical eddy diffusivities (Kv) were calculated using the Yonsei University (YSU) (Hong and Dudhia, 2006) mixing scheme. We used a minimum Kv of 0.1 m^2/sec except for urban grid cells where the minimum Kv was reset to 1.0 m^2/sec within the lowest 200 m of the surface in order to enhance mixing associated with the nighttime "urban heat island" effect. In addition, we invoked the subgrid convection and subgrid stratoform cloud options in our wrfcamx run for 2016.

2.3 Initial and Boundary Concentrations

The lateral boundary and initial species concentrations for the 36 km modeling domain are provided by a three-dimensional global atmospheric chemistry model, the Hemispheric version of the Community Multi-scale Air Quality Model (H-CMAQ) version 3.1.1. The H-CMAQ predictions were used to provide one-way dynamic boundary concentrations at one-hour intervals and an initial concentration field for the 36 km CAMx simulations. The air quality predictions from the 36 km CAMx simulations were used to provide boundary concentrations for the 12 km modeling. More information about the H-CMAQ model and other applications using this tool is available at: https://www.epa.gov/cmaq/hemispheric-scale-applications.

[9] Meteorological Modeling for 2016.docx.

2.4 Emissions Inventories

CAMx requires detailed emissions inventories containing temporally allocated (i.e., hourly) emissions for each grid-cell in the modeling domain for a large number of chemical species that act as primary pollutants and precursors to secondary pollutants. Annual emission inventories for 2016, 2023, and 2028 were preprocessed into CAMx-ready inputs using the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system (Houyoux et al., 2000). Information on the emissions inventories used as input to the CAMx model simulations can be found in the emissions inventory technical support document.[10]

2.5 Air Quality Model Evaluation

An operational model performance evaluation for ozone was conducted to examine the ability of the CAMx modeling system to simulate 2016 measured concentrations. This evaluation focused on graphical analyses and statistical metrics of model predictions versus observations. Details on the evaluation methodology, the calculation of performance statistics, and results are provided in Appendix A. Overall, the ozone model performance statistics for the CAMx 2016 simulation are within or close to the ranges found in other recent peer-reviewed applications (e.g., Simon et al, 2012 and Emory et al, 2017). As described in Appendix A, the predictions from the 2016 modeling platform correspond closely to observed concentrations in terms of the magnitude, temporal fluctuations, and geographic differences for 8-hour daily maximum (MDA8) ozone. Thus, the model performance results demonstrate the scientific credibility of our 2016 modeling platform. These results provide confidence in the ability of the modeling platform to provide a reasonable projection of expected future year ozone concentrations and contributions. Model performance statistics for individual monitoring sites for the period May through September are provided in a spreadsheet file in the docket for this proposed rule.[11]

[10] Preparation of Emissions Inventories for the 2016v1 North American Emissions Modeling Platform.docx.

[11] CAMx 2016 MDA8 O3 Model Performance Stats by Site.xls.

3. Identification of Future Nonattainment and Maintenance Receptors in 2021

3.1 Definition of Nonattainment and Maintenance Receptors

The ozone predictions from the 2016 base year and future case CAMx model simulations were used to calculate average and maximum ozone design values for the 2021 analytic year using the approach described in this section. Following the general approach in the CSAPR Update, we evaluated 2021 projected average and maximum design values in conjunction with the most recent measured ozone design values (i.e., 2019)[12] to identify sites that may warrant further consideration as potential nonattainment or maintenance sites in 2021. Those monitoring sites with 2021 average design values that exceed the NAAQS (i.e., 2021 average design values of 76 ppb or greater)[13] and that are currently measuring nonattainment are considered to be nonattainment receptors in 2021. Similarly, monitoring sites with a projected 2021 maximum design value that exceeds the NAAQS would be projected to be maintenance receptors in 2021. In the CSAPR Update approach, maintenance-only receptors include both those monitoring sites where the projected average design value is below the NAAQS, but the maximum design value is above the NAAQS, and monitoring sites with projected 2021 average design values that exceed the NAAQS, but for which current design values based on measured data do not exceed the NAAQS.

The procedures for calculating projected 2021 average and maximum design values are described below. The monitoring sites that we project to be nonattainment and maintenance receptors for the ozone NAAQS in the 2021 base case are used for assessing the contribution of emissions in upwind states to downwind nonattainment and maintenance of the 2008 ozone NAAQS as part of this proposed rule.

---

[12] The 2019 design values are the most current official design values available for use in this proposed rule. The 2019 ozone design values, by monitoring site, can be found in the following file in the docket: 2010 thru 2019 Ozone Design Values.xls.

[13] In determining compliance with the NAAQS, ozone design values are truncated to integer values. For example, a design value of 70.9 parts per billion (ppb) is truncated to 70 ppb which is attainment. In this manner, design values at or above 71.0 ppb are considered to be violations of the NAAQS.

3.2 Approach for Projecting Ozone Design Values

As noted above, the projected design values for 2021 are based on an interpolation of between the 2016-Centered average and maximum design values and the corresponding average and maximum design values projected for 2023.[14] In this section we describe the approach for projecting 2023 design values followed by the method for calculating design values in 2021.

The ozone predictions from the CAMx model simulations were used to project ambient (i.e., measured) ozone design values (DVs) to 2023 based on an approach that follows from EPA's guidance for attainment demonstration modeling (US EPA, 2018),[15] as summarized here. The modeling guidance recommends using 5-year weighted average ambient design values centered on the base modeling year as the starting point for projecting average design values to the future. Because 2016 is the base emissions year, we used the average ambient 8-hour ozone design values for the period 2014 through 2018 (i.e., the average of design values for 2014-2016, 2015-2017 and 2016-2018) to calculate the 5-year weighted average design values (i.e., 2016-Centered design values). The 5-year weighted average ambient design value at each site was projected to 2023 and 2028 using the Software for Model Attainment Test Software – Community Edition (SMAT-CE). This program calculates the 5-year weighted average design value based on observed data and projects future year values using the relative response predicted by the model. Equation (3-1) describes the recommended model attainment test in its simplest form, as applied for monitoring site *i*:

$$(\mathrm{DVF})_{\mathrm{i}} = (RRF)_i * (DVB)_i \qquad \text{Equation 3-1}$$

$\mathrm{DVF_i}$ is the estimated design value for the future year at monitoring site i; $\mathrm{RRF_i}$ is the relative response factor for monitoring site *i*; and $\mathrm{DVB_i}$ is the base period design value monitored at site *i*. The relative response factor for each monitoring site $(RRF)_i$ is the fractional change in MDA8 ozone between the base and future year. The RRF is based on the average ozone on model-predicted "high" ozone days in grid cells in the vicinity of the monitoring site. The modeling guidance recommends calculating RRFs based on the highest 10 modeled ozone days in the base year simulation at each monitoring site. Specifically, the RRF was calculated based on the 10 highest days in the 2016 base year modeling in the vicinity of each monitor location.

[14] The approach for projecting ozone design values in 2023 was also applied to project ozone design values in 2028.

[15] EPA's ozone attainment demonstration modeling guidance is referred to as "the modeling guidance" in the remainder of this document.

For cases in which the base year model simulation did not have 10 days with ozone values greater than or equal to 60 ppb at a site, we used all days with ozone >= 60 ppb, as long as there were at least 5 days that meet that criteria. At monitor locations with less than 5 days with modeled 2016 base year ozone >= 60 ppb, no RRF or DVF was calculated for the site and the monitor in question was not included in this analysis.

The modeling guidance recommends calculating the RRF using the base year and future year model predictions from the cells immediately surrounding the monitoring site along with the grid cell in which the monitor is located. In this approach the RRF was based on a 3 x 3 array of 12 km grid cells centered on the location of the grid cell containing the monitor.

In light of comments on the Notice of Data Availability (82 FR 1733; January 6, 2017) and other analyses, EPA also projected design values based on a modified version of the "3 x 3" approach for those monitoring sites located in coastal areas. In this alternative approach, EPA eliminated from the RRF calculations the modeling data in those grid cells that are dominated by water (i.e., more than 50 percent of the area in the grid cell is water) and that do not contain a monitoring site (i.e., if a grid cell is more than 50 percent water but contains an air quality monitor, that cell would remain in the calculation). The choice of more than 50 percent of the grid cell area as water as the criteria for identifying overwater grid cells is based on the treatment of land use in the Weather Research and Forecasting model (WRF).[16] Specifically, in the WRF meteorological model those grid cells that are greater than 50 percent overwater are treated as being 100 percent overwater. In such cases the meteorological conditions in the entire grid cell reflect the vertical mixing and winds over water, even if part of the grid cell also happens to be over land with land-based emissions, as can often be the case for coastal areas. Overlaying land-based emissions with overwater meteorology may be representative of conditions at coastal monitors during times of on-shore flow associated with synoptic conditions and/or sea-breeze or lake-breeze wind flows. But there may be other times, particularly with off-shore wind flow when vertical mixing of land-based emissions may be too limited due to the presence of overwater meteorology. Thus, for our modeling EPA calculated 2023 projected average and maximum design values at individual monitoring sites based on

[16] https://www.mmm.ucar.edu/weather-research-and-forecasting-model.

both the "3 x 3" approach as well as the alternative approach that eliminates overwater cells in the RRF calculation for near-coastal areas (i.e., "no water" approach).

For both the "3 x 3" approach and the "no water" approach, the grid cell with the highest base year MDA8 ozone concentration on each day in the applicable array of grid cells surrounding the location of the monitoring site[17] is used for both the base and future components of the RRF calculation. That is, the base and future year data are paired in space for the grid cell that has the highest MDA8 concentration on the given day.

The approach for calculating 2023 projected maximum design values is similar to the approach for calculating the projected average design values. To calculate the projected maximum design values we start with the highest (i.e., maximum) ambient design value from the 2016-Centered 5-year period (i.e., the maximum of design values from 2014-2016, 2014-2017, and 2016-2018). The base period maximum design value at each site was projected to 2023 using the site-specific RRFs, as determined using the procedures for calculating RRFs described above.

The 2023 average and maximum design values for both the "3x3" and "no water" approaches were then paired with the corresponding base period measured design values at each ozone monitoring site. Design values for 2021 for both the "3 x 3" and "no water" approaches were calculated by linearly interpolating between the 2016 base period and 2023 projected values. The steps in the interpolation process for estimating 2021 average and maximum design values are as follows:

(1) Calculate the ppb change in design values between the 2016 base period and 2023;

(2) Divide the ppb change by 7 to calculate the ppb change per year over the 7-year period between 2016 and 2023;

(3) Multiply the ppb per year value by five to calculate the ppb change in design values over the 5-year period between 2016 and 2021;

(4) Subtract the ppb change between 2016 to 2021 from the 2016 design values to produce the design values for 2021.

[17] For the "3 x 3" approach the applicable array contains the 9 grid cells that surround and include the grid cell containing the monitoring site. The applicable array for the "no water" approach includes the grid cell containing the monitoring site along with the subset of the "3 x 3" grid cells that are not classified as "water" grid cells using the criteria described in this TSD.

As noted in the preamble, EPA is soliciting public comment on the use of the "3 x 3" and "no water" approaches for this rulemaking. For the proposed rule, EPA is relying upon design values based on the "no water" approach for identifying nonattainment and maintenance receptors and for calculating contributions, as described in section 4, below.

Consistent with the truncation and rounding procedures for the 8-hour ozone NAAQS, the projected design values are truncated to integers in units of ppb.[18] Therefore, projected design values that are greater than or equal to 76 ppb are considered to be violating the 2008 ozone NAAQS. For those sites that are projected to be violating the NAAQS based on the average design values in 2021, we examined the preliminary measured design values for 2019, which are the most recent available measured design values at the time of this proposed rule. As noted above, we identify nonattainment receptors as those sites that are violating the NAAQS based on current measured air quality and also have projected average design values of 76 ppb or greater. Maintenance-only receptors include both (1) those sites with projected average design values above the NAAQS that are currently measuring clean data and (2) those sites with projected average design values below the level of the NAAQS, but with projected maximum design values of 76 ppb or greater.[19]

Table 3-1 contains the 2016-Centered base period average and maximum design values, the 2021 base case average and maximum design values[20], and the 2019 design values for the two sites that are projected to be nonattainment receptors in 2021 and the two sites that are projected to be maintenance-only receptors in 2021.[21,22]

[18] 40 CFR Part 50, Appendix P to Part 50 – Interpretation of the Primary and Secondary National Ambient Air Quality Standards for Ozone.

[19] In addition to the maintenance-only receptors, the 2021 ozone nonattainment receptors are also maintenance receptors because the maximum design values for each of these sites is always greater than or equal to the average design value.

[20] The design values for 2021 in this table are based on the "no water" approach.

[21] Using design values from the "3 x 3" approach does not change the total *number* of receptors in 2021. However, with the "3 x 3" approach the maintenance-only receptor in New Haven County, CT has a projected maximum design value of 75.5 ppb and would, therefore, not be a receptor using this approach. In contrast, monitoring site 090010017 in Fairfield County, CT has projected average and maximum design values of 75.7 and 76.3 ppb, respectively with the "3 x 3" approach and would, therefore, be a maintenance-only receptor with this approach.

[22] The projected 2021and 2023 design values using both the "3 x 3" and "no-water" approaches along with the 2016-Centered and 2019 design values at individual monitoring sites are provided in the following file which is in the docket for this proposed rule: Projected 2021_2023 3x3 & No Water O3 Design Values.xls.

Table 3-1. 2016-Centered, 2021 average and maximum design values, and 2019 design values at projected nonattainment and maintenance-only receptor sites in the East[23] (units are ppb).

Nonattainment Receptors							
Monitor ID	State	Site	2016-Centered Average	2016-Centered Maximum	2021 Average	2021 Maximum	2019
090013007	CT	Stratford	83.0	83	76.5	77.4	82
090019003	CT	Westport	82.7	83	78.5	78.9	82
Maintenance-Only Receptors							
Monitor ID	State	Site	2016-Centered Average	2016-Centered Maximum	2021 Average	2021 Maximum	2019
090099002	CT	Madison	79.7	82	74.0	76.1	82
482010024	TX	Houston	79.3	81	75.5	77.1	81

4. Ozone Contribution Modeling

The method for estimating contributions in 2021 is based, in part, on source apportionment for 2023. In this section we first describe the source apportionment modeling for 2023 followed by the method for using these data to calculate contributions in 2021 and 2023.

The EPA performed nationwide, state-level ozone source apportionment modeling using the CAMx Ozone Source Apportionment Technology/Anthropogenic Precursor Culpability Analysis (OSAT/APCA) technique[24] to provide data on the expected contribution of 2023 base case NO_X and VOC emissions from all sources in each state.

In the source apportionment model run, we tracked the ozone formed from each of the following contribution categories (i.e., "tags"):

- States – anthropogenic NO_X and VOC emissions from each of the contiguous 48 states and the District of Columbia tracked individually (emissions from all anthropogenic sectors in a given state were combined);
- Biogenics – biogenic NO_X and VOC emissions domain-wide (i.e., not by state);

[23] In this analysis the East includes all states from Texas northward to North Dakota and eastward to the East Coast.

[24] As part of this technique, ozone formed from reactions between biogenic VOC and NO_X with anthropogenic NO_X and VOC are assigned to the anthropogenic emissions.

- Initial and Boundary Concentrations – air quality concentrations used to initialize the 12 km model simulation and air quality concentrations transported into the 12 km modeling domain from the lateral boundaries;
- Tribes – the emissions from those tribal lands for which we have point source inventory data in the 2016 emissions platform (we did not model the contributions from individual tribes);
- Canada and Mexico – anthropogenic emissions from sources in the portions of Canada and Mexico included in the 12 km modeling domain (contributions from Canada and Mexico were not modeled separately);
- Fires – combined emissions from wild and prescribed fires domain-wide within the 12 km modeling domain (i.e., not by state); and
- Offshore – combined emissions from offshore marine vessels and offshore drilling platforms (i.e., not by state).

The source apportionment modeling provided hourly contributions for 2023 to ozone from anthropogenic NO_X and VOC emissions in each state, individually to ozone concentrations in each model grid cell. The contributions to ozone from chemical reactions between biogenic NO_X and VOC emissions were modeled and assigned to the "biogenic" category. The contributions from wild fire and prescribed fire NO_X and VOC emissions were modeled and assigned to the "fires" category. The contributions from the "biogenic", "offshore", and "fires" categories are not assigned to individual states nor are they included in the state contributions.

CAMx OSAT/APCA model run was performed for the period May 1 through September 30 using the projected 2023 base case emissions and 2016 meteorology for this time period. The hourly contributions[25] from each tag were processed to calculate an 8-hour average contribution metric value for each tag at each monitoring site. The contribution metric values at each individual monitoring site are calculated using model predictions for the grid cell containing the monitoring site. The process for calculating the average contribution metric uses the source apportionment outputs in a "relative sense" to apportion the projected average design value at each monitoring location into contributions from each individual tag. This process is similar in

[25] Contributions from anthropogenic emissions under "NO_X-limited" and "VOC-limited" chemical regimes were combined to obtain the net contribution from NO_X and VOC anthropogenic emissions in each state.

concept to the approach described above for using model predictions to calculate future year ozone design values.

The basic approach used to calculate the average contribution metric values for 2021 and 2023[26] is described by the following steps:

(1) For the model grid cells containing an ozone monitoring site, calculate the 8-hour average contribution from each source tag to each monitoring site for the time period of the 8-hour daily maximum modeled (i.e., MDA8) concentration on each day;

(2) Average the MDA8 concentrations for each of the top 10 modeled ozone concentration days in 2023 and average the 8-hour contributions for each of these same days for each tag;

(3) Divide the 10-day average contribution for each tag by the corresponding 10-day average concentration to obtain a Relative Contribution Factor (RCF) for each tag for each monitoring site;

(3) Multiply the 2021 and 2023 average design values by the corresponding RCF to produce the average contribution metric values at each monitoring site in 2021 and 2023, respectively.

The contribution metric values calculated from step 3 are truncated to two digits to the right of the decimal (e.g., a calculated contribution of 0.78963… is truncated to 0.78 ppb). As a result of truncation, the tabulated contributions may not always sum to the 2021 and 2023 average design values. The details on how this approach is applied in the computer code to perform the contribution calculations is provided in Appendix B.

4.2 Contribution Modeling Results

The contribution metric values from each state and the other source tags at individual nonattainment and maintenance-only sites in the East in 2021 are provided in Appendix C. The largest contribution values from each state subject to this proposed rule to 2021 downwind nonattainment sites and to downwind maintenance-only sites are provided in Table 4-1.[27]

[26] The approach described for calculating contributions in 2023 was also applied to the 2028 modeling to calculate contributions for 2028.

[27] The 2021, 2023, and 2028 contribution metric values from each state and from the other source tags to individual monitoring sites nationwide are provided in a file in the docket for this proposed rule: Ozone Design Values & Contributions_Proposed Revised CSAPR Update.xls

Table 4-1. Largest contribution from each state to downwind nonattainment and maintenance-only Receptors in 2021 (units are ppb).

Upwind State	Largest Downwind Contribution to Nonattainment Receptors for Ozone	Largest Downwind Contribution to Maintenance-Only Receptors for Ozone
Alabama	0.11	0.27
Arkansas	0.18	0.15
Illinois	0.81	0.80
Indiana	1.26	1.08
Iowa	0.17	0.22
Kansas	0.13	0.11
Kentucky	0.87	0.79
Louisiana	0.27	4.68
Maryland	1.21	1.56
Michigan	1.71	1.62
Mississippi	0.10	0.37
Missouri	0.36	0.33
New Jersey	8.62	5.71
New York	14.44	12.54
Ohio	2.55	2.35
Oklahoma	0.20	0.14
Pennsylvania	6.86	5.64
Texas	0.59	0.36
Virginia	1.30	1.69
West Virginia	1.49	1.55
Wisconsin	0.23	0.23

4.4 Upwind/Downwind Linkages

In CSAPR and the CSAPR Update, the EPA used a contribution screening threshold of 1 percent of the NAAQS to identify upwind states that may significantly contribute to downwind nonattainment and/or maintenance problems and which warrant further analysis to determine if emissions reductions might be required from each state to address the downwind air quality problem. The EPA determined that 1 percent was an appropriate threshold to use in the analysis for those rulemakings because there were important, even if relatively small, contributions to identified nonattainment and maintenance receptors from multiple upwind states mainly in the eastern U.S. The agency has historically found that the 1 percent threshold is appropriate for identifying interstate transport linkages for states collectively contributing to downwind ozone nonattainment or maintenance problems because that threshold captures a high percentage of the total pollution transport affecting downwind receptors.

Based on the approach used in CSAPR and the CSAPR Update, upwind states that contribute ozone in amounts at or above the 1 percent of the NAAQS threshold to a particular downwind nonattainment or maintenance receptor are considered to be "linked" to that receptor in Step 2 of the CSAPR framework for purposes of further analysis in Step 3 to determine whether and what emissions from the upwind state contribute significantly to downwind nonattainment and interfere with maintenance of the NAAQS at the downwind receptors. For the 2008 ozone NAAQS the value of a 1 percent threshold is 0.75 ppb. The individual upwind state to downwind receptor "linkages" and contributions based on a 0.75 ppb threshold are identified in Table 4-2. In summary, Indiana, Kentucky, Maryland, Michigan, New Jersey, New York, Ohio, Pennsylvania, Virginia, and West Virginia are each linked to the nonattainment receptors in Westport and Stratford, and the maintenance-only receptor in Madison, Connecticut; Illinois is linked to the nonattainment receptor in Westport and the maintenance-only receptor in Madison; and Louisiana is linked to the maintenance-only receptor in Houston, Texas.

As noted above, when applying the CSAPR framework, an upwind state's linkage to a downwind receptor alone does not determine whether the state significantly contributes to nonattainment or interferes with maintenance of a NAAQS to a downwind state. The determination of significant contribution is made in Step 3 as part of a multi-factor analysis, as described in the Ozone Transport Policy Analysis Technical Support Document.

Table 4-2. Contributions from upwind states that are "linked" to each downwind nonattainment and maintenance receptor in the East.[28]

	Nonattainment Receptors			Maintenance-Only Receptors
Upwind State	Stratford, CT	Westport, CT	Upwind State	Madison, CT
Illinois	0.69	0.81	Illinois	0.80
Indiana	0.99	1.26	Indiana	1.08
Kentucky	0.78	0.87	Kentucky	0.79
Maryland	1.21	1.20	Maryland	1.56
Michigan	1.16	1.71	Michigan	1.62

[28] Note that for the purpose of completeness we have included the contribution from Illinois to the receptor in Stratford, CT, even though Illinois is not linked to this receptor.

	Nonattainment Receptors			Maintenance-Only Receptors
New Jersey	7.70	8.62	New Jersey	5.71
New York	14.42	14.44	New York	12.54
Ohio	2.34	2.55	Ohio	2.35
Pennsylvania	6.72	6.86	Pennsylvania	5.64
Virginia	1.29	1.30	Virginia	1.69
West Virginia	1.45	1.49	West Virginia	1.55
				Houston, TX
			Louisiana	4.68

5. References

Emery, C., Z. Liu, A. Russell, M. T. Odom, G. Yarwood, and N. Kumar, 2017. Recommendations on Statistics and Benchmarks to Assess Photochemical Model Performance. *J. Air and Waste Management Association*, 67, 582-598.

Gilliam, R.C. and J.E. Pleim, 2010. Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW. *J. Appl. Meteor. Climatol.*, **49**, 760–774.

Henderson, B.H., F. Akhtar, H.O.T. Pye, S.L. Napelenok, W.T. Hutzell, 2014. A Database and Tool for Boundary Conditions for Regional Air Quality Modeling: Description and Evaluations, *Geoscientific Model Development*, **7**, 339-360.

Hong, S-Y, Y. Noh, and J. Dudhia, 2006. A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Mon. Wea. Rev.*, 134, 2318–2341.

Houyoux, M.R., Vukovich, J.M., Coats, C.J., Wheeler, N.J.M., Kasibhatla, P.S.,2000. Emissions Inventory Development and Processing for the Seasonal Model for Regional Air Quality (SMRAQ) project, *Journal of Geophysical Research – Atmospheres*, 105(D7), 9079-9090.

Iacono, M.J., J.S. Delamere, E.J. Mlawer, M.W. Shephard, S.A Clough, and W.D. Collins, 2008. Radiative Forcing by Long-Lived Greenhouse Gases: Calculations with the AER Radiative Transfer Models, *J. Geophys. Res.,* 113, D13103.

Kain, J.S., 2004. The Kain-Fritsch Convective Parameterization: An Update, *J. Appl. Meteor.,* 43, 170-181.

Ma, L-M. and Tan Z-M, 2009. Improving the Behavior of Cumulus Parameterization for Tropical Cyclone Prediction: Convective Trigger, *Atmospheric Research*, 92, 190-211.

Morrison, H.J., A. Curry, and V.I. Khvorostyanov, 2005. A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description, *J. Atmos. Sci.*, 62, 1665–1677.

Morrison, H. and A. Gettelman, 2008. A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and Numerical Tests, *J. Climate,* 21, 3642-3659.

Pleim, J.E. and A. Xiu, 2003. Development of a Land-Surface Model. Part II: Data Assimilation, *J. Appl. Meteor.*, 42, 1811–1822

Pleim, J.E., 2007a. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing, *J. Appl. Meteor. Climatol.*, 46, 1383–1395.

Pleim, J.E., 2007b. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part II: Application and Evaluation in a Mesoscale Meteorological Model, *J. Appl. Meteor. Climatol.*, 46, 1396–1409.

Ramboll Environ, 2020. User's Guide Comprehensive Air Quality Model with Extensions version 7, www.camx.com. Ramboll Environ International Corporation, Novato, CA.

Ramboll Environ, 2019. wrfcamx version 4.7 Release Notes. February 26, 2019. www.camx.com. Ramboll Environ International Corporation, Novato, CA.

Skamarock, W.C., J.B. Klemp, J. Dudhia, et al., 2008. A Description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR. http://wwww.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf

Simon, H., K.R. Baker, and S.B. Phillips, 2012. Compilation and Interpretation of Photochemical Model Performance Statistics Published between 2006 and 2012, *Atmospheric Environment*, 61, 124-139.

Stammer, D., F.J. Wentz, and C.L. Gentemann, 2003. Validation of Microwave Sea Surface Temperature Measurements for Climate Purposes, *J. of Climate*, 16(1), 73-87.

U.S. Environmental Protection Agency, 2018. Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze, Research Triangle Park, NC. https://www3.epa.gov/ttn/scram/guidance/guide/O3-PM-RH-Modeling_Guidance-2018.pdf

Xiu, A., and J.E. Pleim, 2001, Development of a Land Surface Model. Part I: Application in a Meso scale Meteorological Model, *J. Appl. Meteor.,* 40, 192-209.

Yantosca, B. 2004. GEOS-CHEMv7-01-02 User's Guide, Atmospheric Chemistry Modeling Group, Harvard University, Cambridge, MA.

Appendix A

2016 Model Performance Evaluation

An operational model evaluation was conducted for the 2016 base year CAMx v7beta6 simulation performed for the 12 km U.S. modeling domain. The purpose of this evaluation is to examine the ability of the 2016 air quality modeling platform to represent the magnitude and spatial and temporal variability of measured (i.e., observed) ozone concentrations within the modeling domain. The evaluation presented here is based on model simulations using the 2016 emissions platform (i.e., scenario name 2016fh_16j)). The model evaluation for ozone focuses on comparisons of model predicted 8-hour daily maximum concentrations to the corresponding observed data at monitoring sites in the EPA Air Quality System (AQS). The locations of the ozone monitoring sites in this network are shown in Figure A-1.

Included in the evaluation are statistical measures of model performance based upon model-predicted versus observed concentrations that were paired in space and time. Model performance statistics were calculated for several spatial scales and temporal periods. Statistics were calculated for individual monitoring sites, and in aggregate for monitoring sites within each state and within each of nine climate regions of the 12 km U.S. modeling domain. The regions include the Northeast, Ohio Valley, Upper Midwest, Southeast, South, Southwest, Northern Rockies, Northwest and West[1,2], which are defined based upon the states contained within the National Oceanic and Atmospheric Administration (NOAA) climate regions (Figure A-2)[3] as defined in Karl and Koss (1984).

[1] The nine climate regions are defined by States where: Northeast includes CT, DE, ME, MA, MD, NH, NJ, NY, PA, RI, and VT; Ohio Valley includes IL, IN, KY, MO, OH, TN, and WV; Upper Midwest includes IA, MI, MN, and WI; Southeast includes AL, FL, GA, NC, SC, and VA; South includes AR, KS, LA, MS, OK, and TX; Southwest includes AZ, CO, NM, and UT; Northern Rockies includes MT, NE, ND, SD, WY; Northwest includes ID, OR, and WA; and West includes CA and NV.

[2] Note most monitoring sites in the West region are located in California (see Figures 2A-2a and 2A-2b), therefore statistics for the West will be mostly representative of California ozone air quality.

[3] NOAA, National Centers for Environmental Information scientists have identified nine climatically consistent regions within the contiguous U.S., http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php.

For maximum daily average 8-hour (MDA8) ozone, model performance statistics were created for the period May through September.[4] The aggregate statistics by state and by climate region are presented in this appendix. Model performance statistics for MDA8 ozone at individual monitoring sites based on days with observed values $\geq$ 60 ppb can be found in the docket in the file named "2016v1 CAMx Ozone Model Performance Statistics by Site".

In addition to the above performance statistics, we prepared several graphical presentations of model performance for MDA8 ozone. These graphical presentations include:
(1) maps that show the mean bias and error as well as normalized mean bias and error calculated for MDA8 $\geq$ 60 ppb for May through September at individual AQS and CASTNet monitoring sites;
(2) bar and whisker plots that show the distribution of the predicted and observed MDA8 ozone concentrations by month (May through September) and by region and by network; and
(3) time series plots (May through September) of observed and predicted MDA8 ozone concentrations for selected monitoring sites.

The Atmospheric Model Evaluation Tool (AMET) was used to calculate the model performance statistics used in this document (Gilliam et al., 2005). For this evaluation we have selected the mean bias, mean error, normalized mean bias, and normalized mean error to characterize model performance, statistics which are consistent with the recommendations in Simon et al. (2012) and the draft photochemical modeling guidance (U.S. EPA, 2014a).

Mean bias (MB) is the average of the difference (predicted – observed) divided by the total number of replicates (*n*). Mean bias is given in units of ppb and is defined as:

MB = $\frac{1}{n}\sum_{1}^{n}(P - O)$, where P = predicted and O = observed concentrations

Mean error (ME) calculates the absolute value of the difference (predicted - observed) divided by the total number of replicates (*n*). Mean error is given in units of ppb and is defined as:

[4] In calculating the ozone season statistics we limited the data to those observed and predicted pairs with observations that are $\geq$ 60 ppb in order to focus on concentrations at the upper portion of the distribution of values.

$$ME = \frac{1}{n}\sum_1^n |P - O|$$

Normalized mean bias (NMB) is the average the difference (predicted - observed) over the sum of observed values. NMB is a useful model performance indicator because it avoids over inflating the observed range of values, especially at low concentrations. Normalized mean bias is given in percentage units and is defined as:

$$NMB = \frac{\sum_1^n (P-O)}{\sum_1^n (O)} * 100$$

Normalized mean error (NME) is the absolute value of the difference (predicted - observed) over the sum of observed values. Normalized mean error is given in percentage units and is defined as:

$$NME = \frac{\sum_1^n |P-O|}{\sum_1^n (O)} * 100$$

As described in more detail below, the model performance statistics indicate that the 8-hour daily maximum ozone concentrations predicted by the 2016 CAMx modeling platform closely reflect the corresponding 8-hour observed ozone concentrations in each region of the 12 km U.S. modeling domain. The acceptability of model performance was judged by considering the 2016 CAMx performance results in light of the range of performance found in recent regional ozone model applications (Emery et al., NRC, 2002; Phillips et al., 2007; Simon et al., 2012; U.S. EPA, 2005; U.S. EPA, 2009; U.S. EPA, 2010.[5] These other modeling studies

[5] Christopher Emery, Zhen Liu, Armistead G. Russell, M. Talat Odman, Greg Yarwood & Naresh Kumar (2017) Recommendations on statistics and benchmarks to assess photochemical model performance, Journal of the Air & Waste Management Association, 67:5, 582-598, DOI: 10.1080/10962247.2016.1265027

National Research Council (NRC), 2002. Estimating the Public Health Benefits of Proposed Air Pollution Regulations, Washington, DC: National Academies Press.

U.S. Environmental Protection Agency; Technical Support Document for the Final Clean Air Interstate Rule: Air Quality Modeling; Office of Air Quality Planning and Standards; RTP, NC; March 2005 (CAIR Docket OAR-2005-0053-2149).

U.S. Environmental Protection Agency, Proposal to Designate an Emissions Control Area for Nitrogen Oxides, Sulfur Oxides, and Particulate Matter: Technical Support Document. EPA-420-R-007, 329pp., 2009. (http://www.epa.gov/otaq/regs/nonroad/marine/ci/420r09007.pdf) represent a wide range of modeling analyses that cover various models, model configurations, domains, years and/or episodes, chemical mechanisms, and aerosol modules. Overall, the ozone model performance results for the 2016 CAMx simulations are within the range found in other recent peer-reviewed and regulatory applications. The model performance results, as described in this document, demonstrate that the predictions from the 2016 modeling platform correspond closely to observed concentrations in terms of the magnitude, temporal fluctuations, and geographic differences for 8-hour daily maximum ozone.

The 8-hour ozone model performance bias and error statistics by network for the period May-September for each region and each state are provided in Tables A-1 and A-2, respectively. The statistics shown were calculated using data pairs on days with observed 8-hour ozone of ≥ 60 ppb. The distributions of observed and predicted 8-hour ozone by month in the period May through September for each region are shown in Figures A-3 through A-11. Spatial plots of the mean bias and error as well as the normalized mean bias and error for individual monitors are shown in Figures A-12 through A-15.

Time series plots of observed and predicted MDA 8-hour ozone during the period May through September for 2021 nonattainment and/or maintenance sites are provided in Figure A-16, (a) through (d).

As indicated by the statistics in Table A-1, the base year 2016 modeling tends to under predict MDA8 ozone, although the bias and error are relatively low in each region. Generally, mean bias for 8-hour ozone ≥ 60 ppb during the period May through September is close to or within ± 10 ppb[6] in nearly all of the regions. The mean error is less than 10 ppb in the Northeast, Ohio Valley, Southeast, South, and Southwest. Normalized mean bias is within ± 10 percent for

---

Phillips, S., K. Wang, C. Jang, N. Possiel, M. Strum, T. Fox, 2007. Evaluation of 2002 Multi-pollutant Platform: Air Toxics, Ozone, and Particulate Matter, 7th Annual CMAS Conference, Chapel Hill, NC, October 6-8, 2008. (http://www.cmascenter.org/conference/2008/agenda.cfm).

U.S. Environmental Protection Agency, 2010, Renewable Fuel Standard Program (RFS2) Regulatory Impact Analysis. EPA-420-R-10-006. February 2010. Sections 3.4.2.1.2 and 3.4.3.3. Docket EPA-HQ-OAR-2009-0472-11332. (http://www.epa.gov/oms/renewablefuels/420r10006.pdf)

Simon, H., Baker, K.R., and Phillips, S. (2012) Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmospheric Environment* **61**, 124-139.

[6] Note that "within ± 5 ppb" includes values that are greater than or equal to -5 ppb and less than or equal to 5 ppb.

sites in the Northeast, Ohio Valley, Southeast, and Southwest with somewhat larger values in the other regions where the normalized mean bias is less than 20 percent. The normalized mean error is less than 15 percent for the Northeast, Ohio Valley, Southeast, South, and Southwest and less that 20 percent in the Upper Midwest, Northern Rockies, Northwest, and West regions.

The monthly distributions of MDA8 model-predicted ozone for each region are provided in Figures A-3 through A-11. In the Northeast, Ohio Valley, and Upper Midwest, the model under predicts in May and June followed by over prediction in the remainder of the ozone season. In the Southeast, the distribution of predictions generally corresponds well with that of the observed concentrations in May and June with over prediction during the remainder of the ozone season. The distribution of predicted concentrations tends to be close to that of the observed data at the 25^{th} percentile, median and 75th percentile values in the South with a tendency for under-prediction in the Southwest and Northern Rockies. In the Northwest modeled MDA8 ozone under predicts in May and June, but then closely tracks the observed values in July, August, and September. Measured MDA8 ozone is under predicted in the West region.

Figures A-12 through A-15 show the spatial variability in bias and error at monitor locations for MDA8 ozone on days with measured concentrations $\geq$ 60 ppb. Mean bias, as seen from Figure A-12, is within $\pm$ 5 ppb at many sites from portions of Texas northeastward to the Northeast Corridor. In this area, the normalized mean bias is within $\pm$ 10 percent, the mean error is mainly between 4 and 8 ppb and the normalized mean error is between 5 to 15 percent. At most monitoring sites across the remainder of the East the model under predicts by 5 to 10 ppb, the normalized mean bias is between 5 and 10 percent, the mean error is in the range of 8 to 12 ppb, and normalized mean error of 5 to 10 percent. The exceptions are at some monitoring sites in mainly the interior parts of Michigan, Wisconsin, and Upstate New York where the magnitude of under prediction is 10 to 15 ppb, the normalized mean bias is -10 to 30 percent, the mean error is 12 to 16 ppb, and the normalized mean error is 15 to 25 percent.

Elsewhere in the U.S., mean bias is generally in the range of -5 to -10 ppb. The most notable exceptions are in portions of Arizona, California, and Wyoming where the mean bias is in the range of -10 to -15 ppb and up to -15 to 20 ppb at some sites in the Central Valley of California. At monitoring sites in the vicinity of Denver Las Vegas, Phoenix, San Francisco, and along the California coastline the normalized mean bias is within ± 10 percent. Model predictions at monitoring sites in these areas also have the lowest mean error (e.g., 6 to 10 ppb) and the lowest normalized mean error (e.g., ≤ 15 percent) in the western U.S.

In addition to the above analysis of overall model performance, we also examine how well the modeling platform replicates day to day fluctuations in observed 8-hour daily maximum concentrations for the four monitoring sites that are projected to be receptors in 2021 (i.e., Stratford, CT, Westport, CT, New Haven-Madison, CT, and Houston-Aldine, TX). For this site-specific analysis we present the time series of observed and predicted 8-hour daily maximum concentrations by site over the period May through September. The results, as shown in Figures A-16 (a) through (d), indicate that the modeling platform generally replicates the day-to-day variability in ozone during this time period at these sites. That is, days with high modeled concentrations are generally also days with high measured concentrations and, conversely, days with low modeled concentrations are also days with low measured concentrations in most cases. For example, model predictions at these sites not only accurately capture the day-to-day variability in the observations, but also appear to capture the timing and magnitude of multi-day high ozone episodes as well as time periods of relatively low concentrations.

Model performance statistics for MDA8 ozone ≥ 60 ppb during the period May through September at each of the four receptor sites are provided in Table A-2. These statistics indicate that, overall, the model predictions are close in magnitude to the corresponding measurements. As evident from the mean bias and normalized mean bias, the model under predicts the corresponding measured data to some extent. The magnitude of the performance statistics is consistent across these sites. The general range of mean bias 4 to 6 ppb, normalized mean is -6 to -8 ppb, mean error is 7 to 9 ppb, and the normalized mean error is less than 10 to 13%.

In summary, the ozone model performance statistics for the CAMx 2016 simulation are within or close to the ranges found in other recent peer-reviewed applications (e.g., Simon et al, 2012 and Emory et al, 2017). As described in this appendix, the predictions from the 2016 modeling platform correspond closely to observed concentrations in terms of the magnitude, temporal fluctuations, and geographic differences for 8-hour daily maximum ozone. Thus, the model performance results demonstrate the scientific credibility of our 2016 modeling platform.

These results provide confidence in the ability of the modeling platform to provide a reasonable projection of expected future year ozone concentrations and contributions.

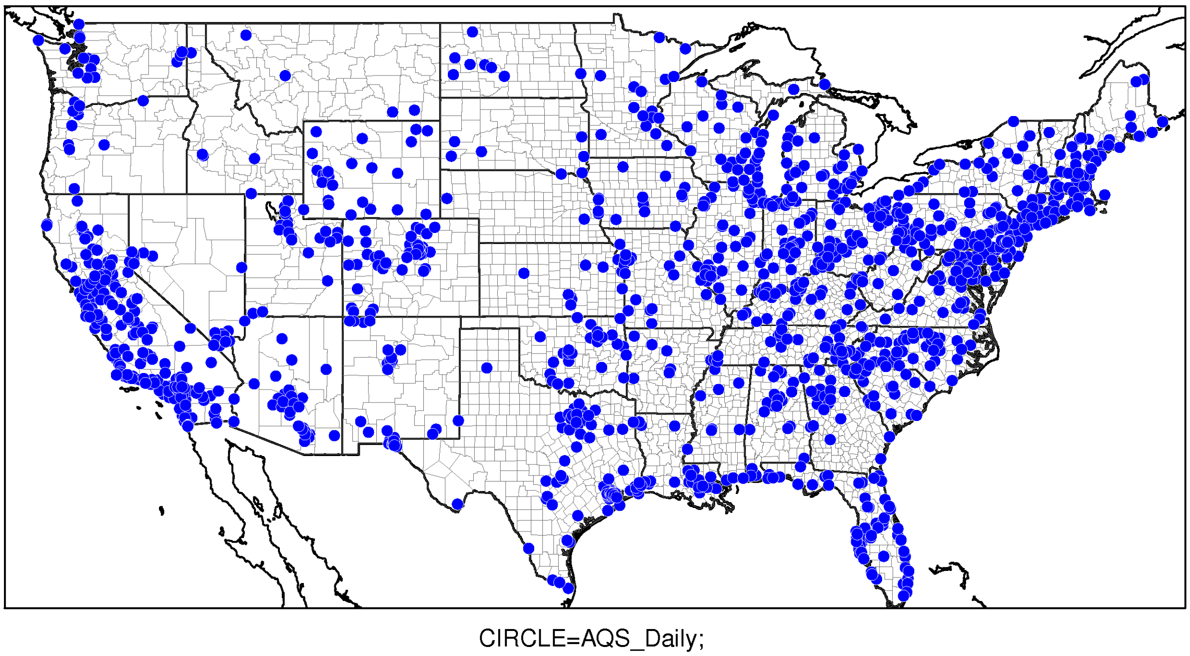


Figure A-1a. AQS ozone monitoring sites.

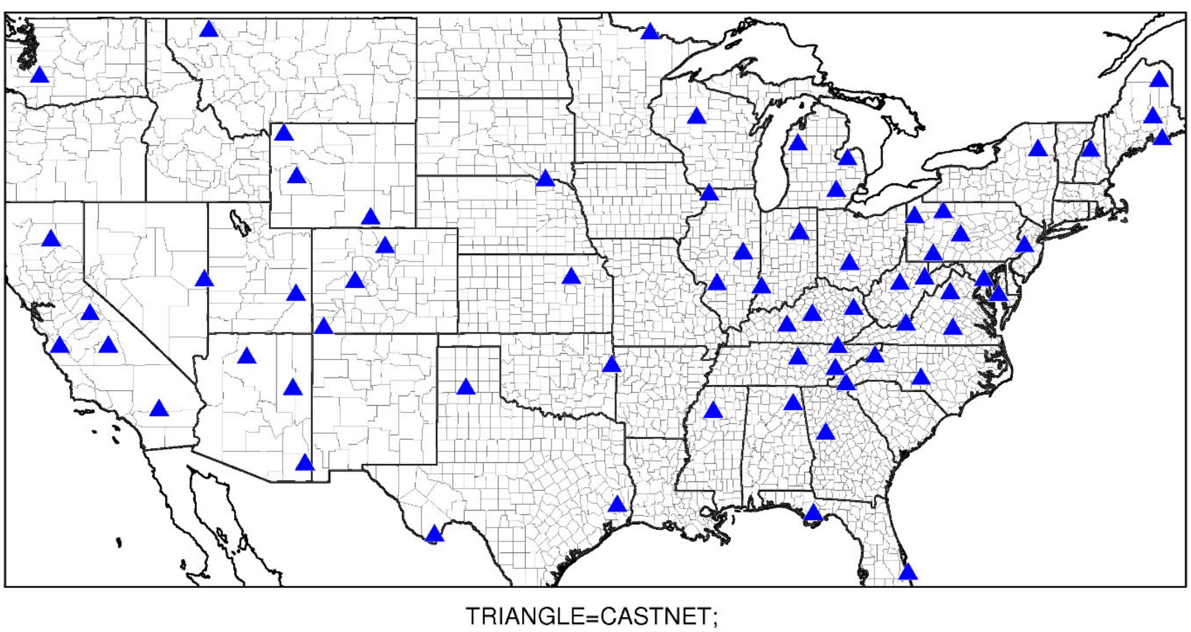


Figure A-1b. CASTNet ozone monitoring sites.

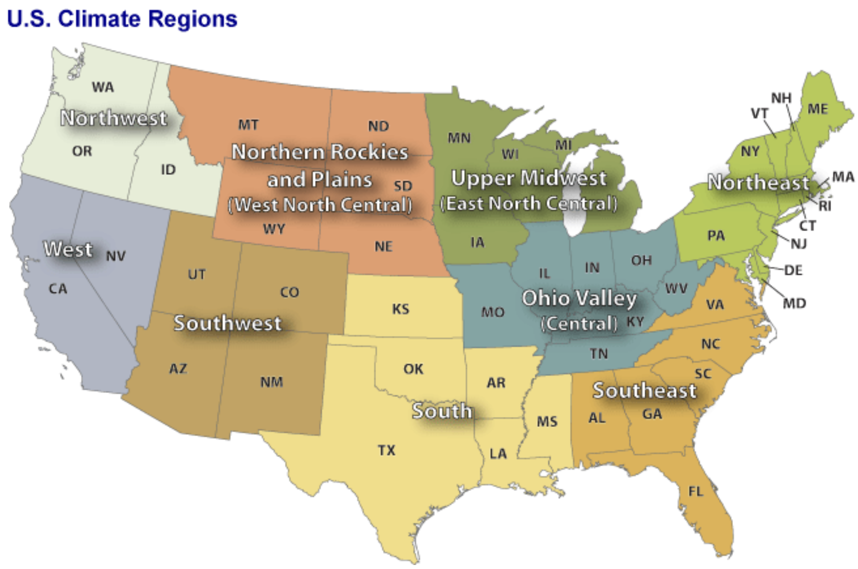


Figure A-2. NOAA climate regions (source: http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php#references)

Table A-1. Performance statistics for MDA8 ozone ≥ 60 ppb for May through September by climate region.

Climate Region	Number of Days ≥ 60 ppb	MB (ppb)	ME (ppb)	NMB (%)	NME (%)
Northeast	2962	-3.7	7.2	-5.6	10.7
Ohio Valley	3201	-5.3	7.9	-8.1	12.0
Upper Midwest	1134	-10.3	11.0	-15.6	16.6
Southeast	1401	-3.8	6.6	-5.8	10.2
South	983	-6.2	8.2	-9.6	12.6
Southwest	3076	-7.8	9.3	-12.0	14.3
Northern Rockies	206	-11.3	11.7	-18.0	18.6
Northwest	84	-7.9	11.0	-12.1	17.0
West	8274	-10.9	11.8	-15.4	16.7

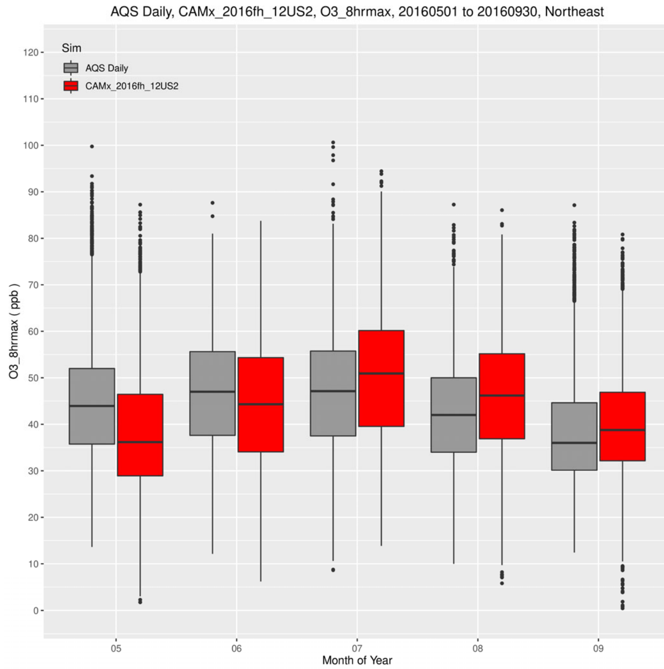


Figure A-3. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Northeast region, [symbol = median; top/bottom of box = 75th/25th percentiles; top/bottom dots = peak/low values]

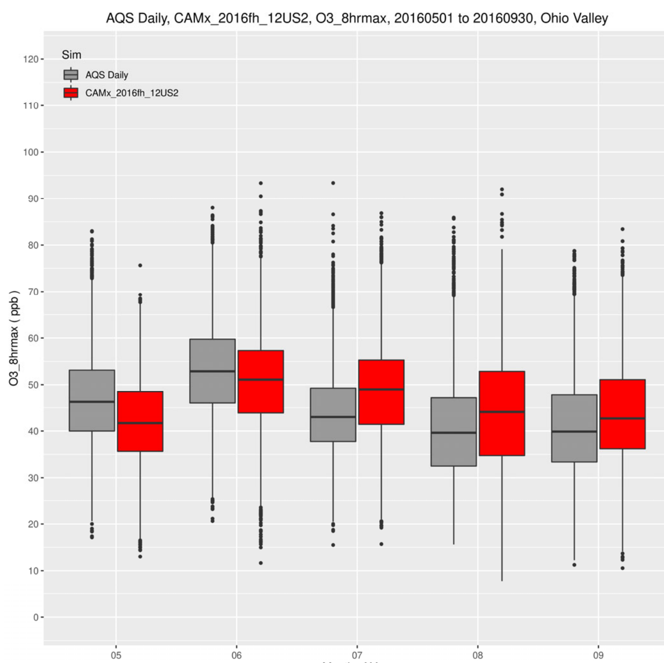


Figure A-4. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Ohio Valley region.

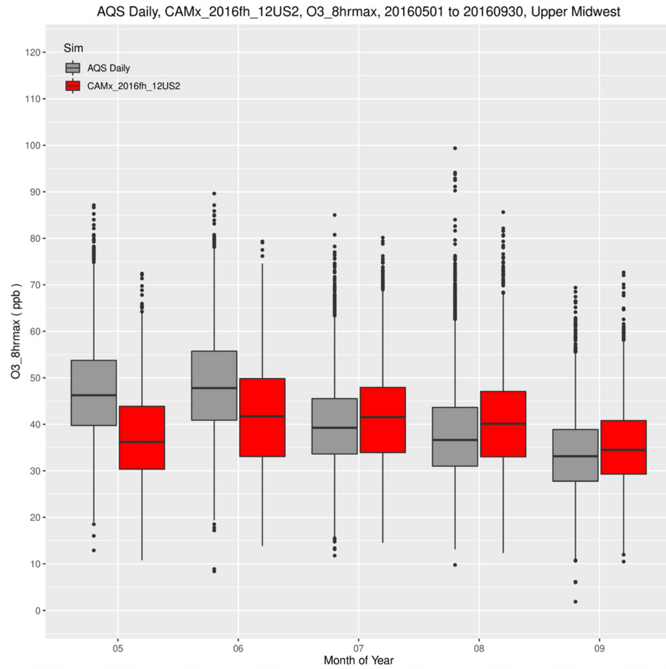


Figure A-5. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Upper Midwest region.

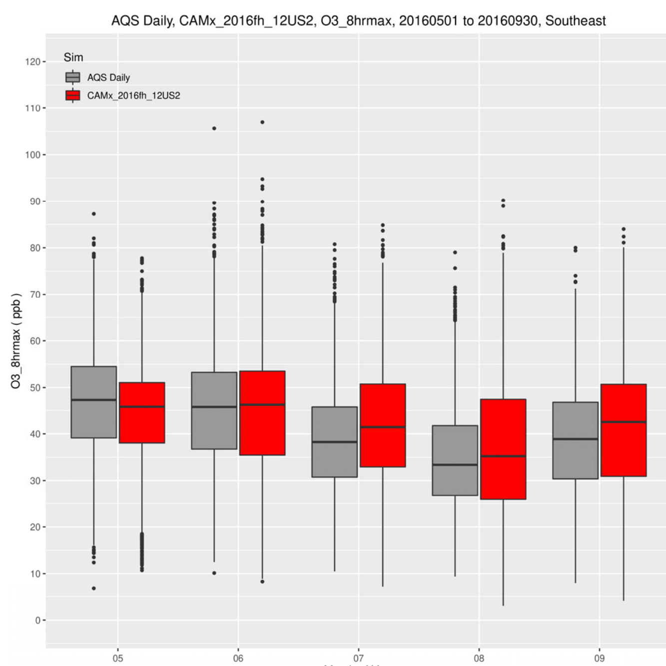


Figure A-6. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Southeast region.

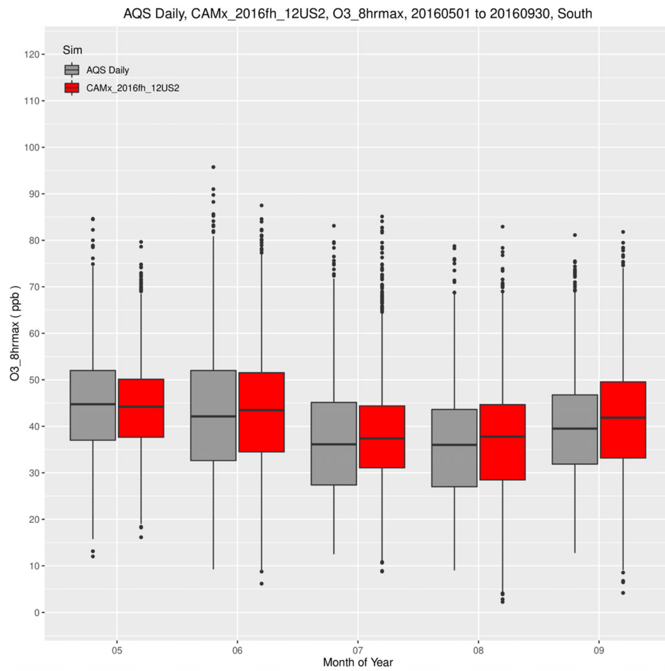


Figure A-7. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the South region.

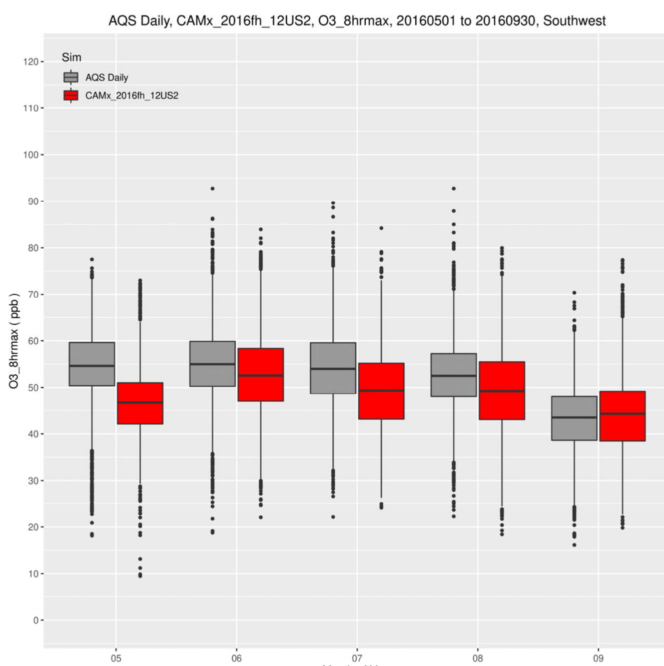


Figure A-8. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Southwest region.

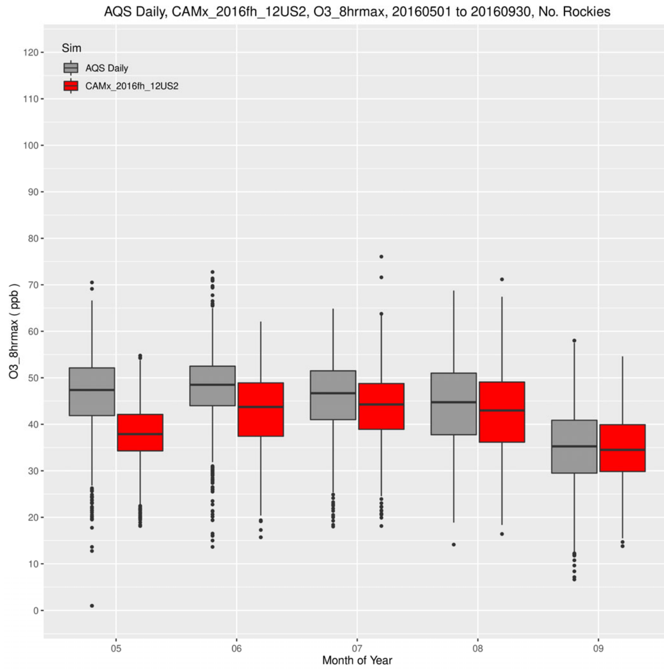


Figure A-9. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Northern Rockies region, AQS Network (left) and CASTNet (right).

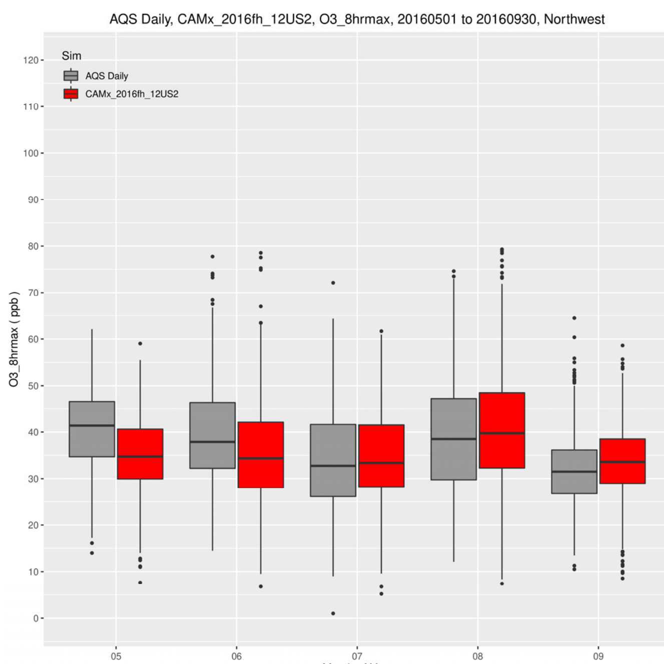


Figure A-10. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the Northwest region.

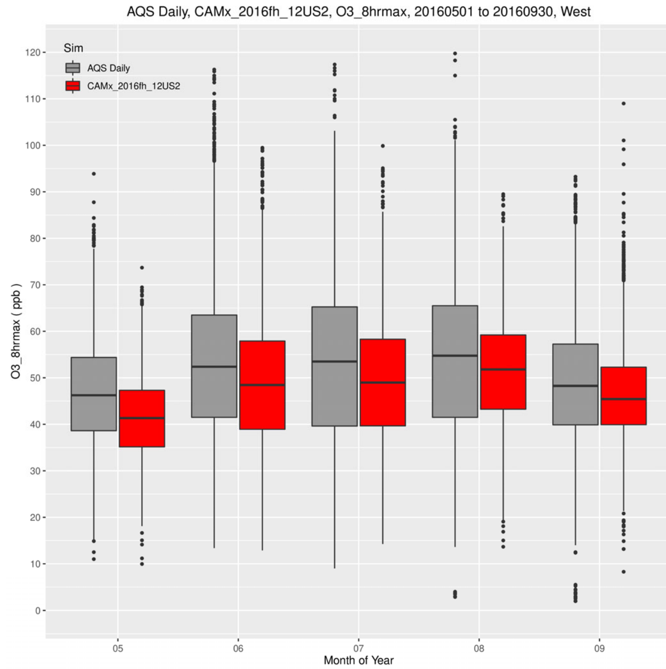


Figure A-11. Distribution of observed and predicted MDA8 ozone by month for the period May through September for the West region.

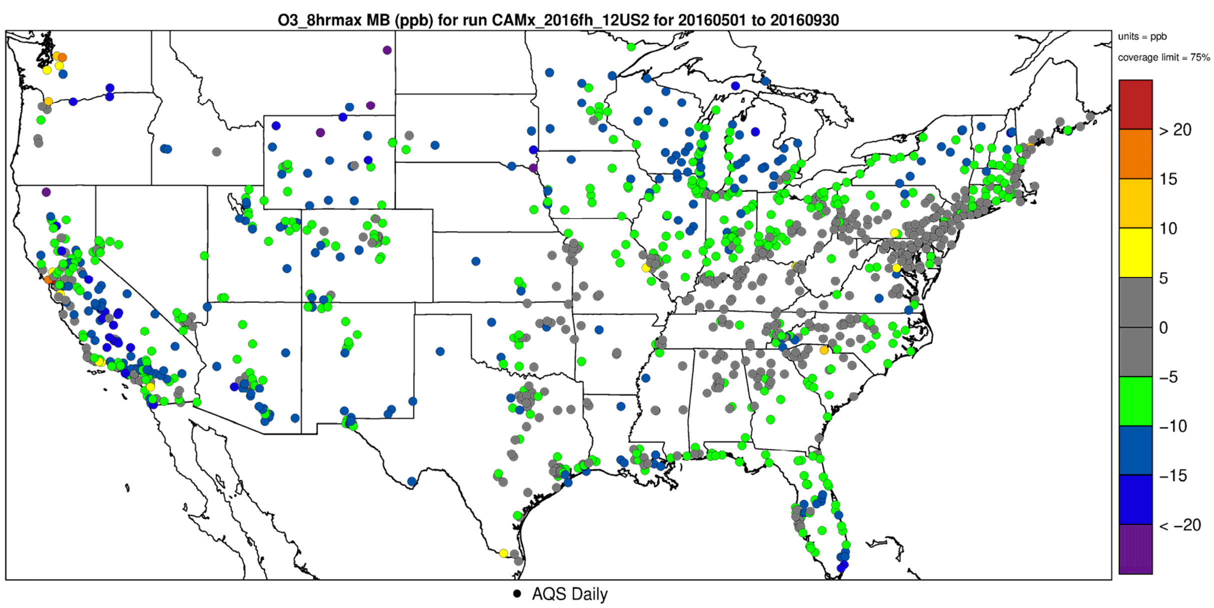


Figure A-12. Mean Bias (ppb) of MDA8 ozone ≥ 60 ppb over the period May-September.

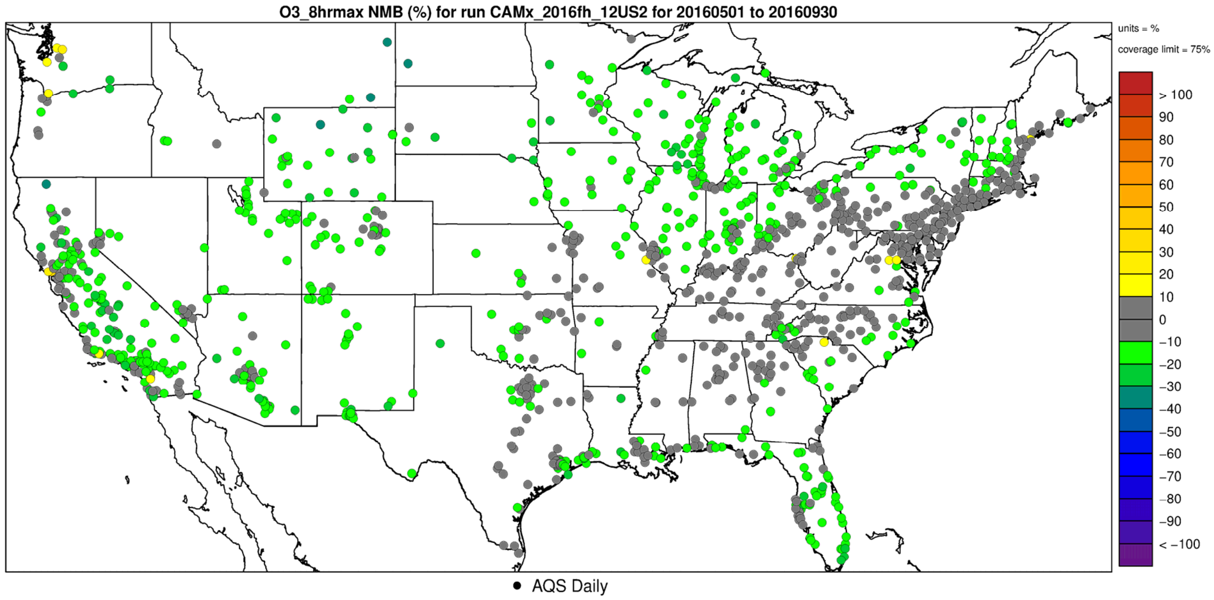


Figure A-13. Normalized Mean Bias (%) of MDA8 ozone ≥ 60 ppb over the period May-September 2016.

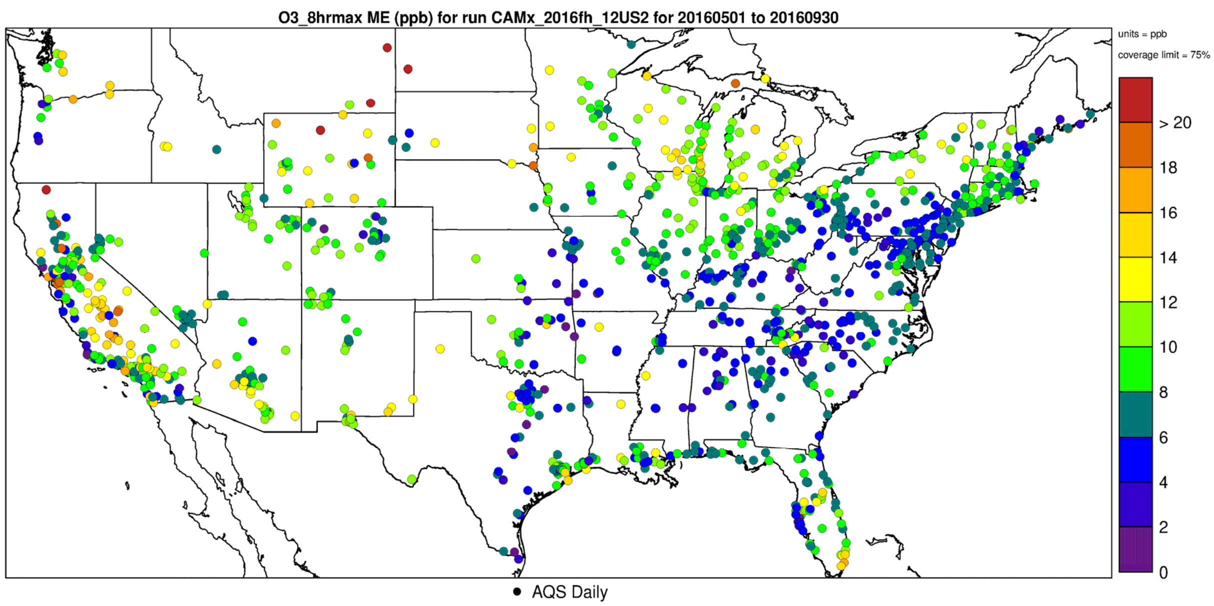


Figure A-14. Mean Error (ppb) of MDA8 ozone ≥ 60 ppb over the period May-September 2016.

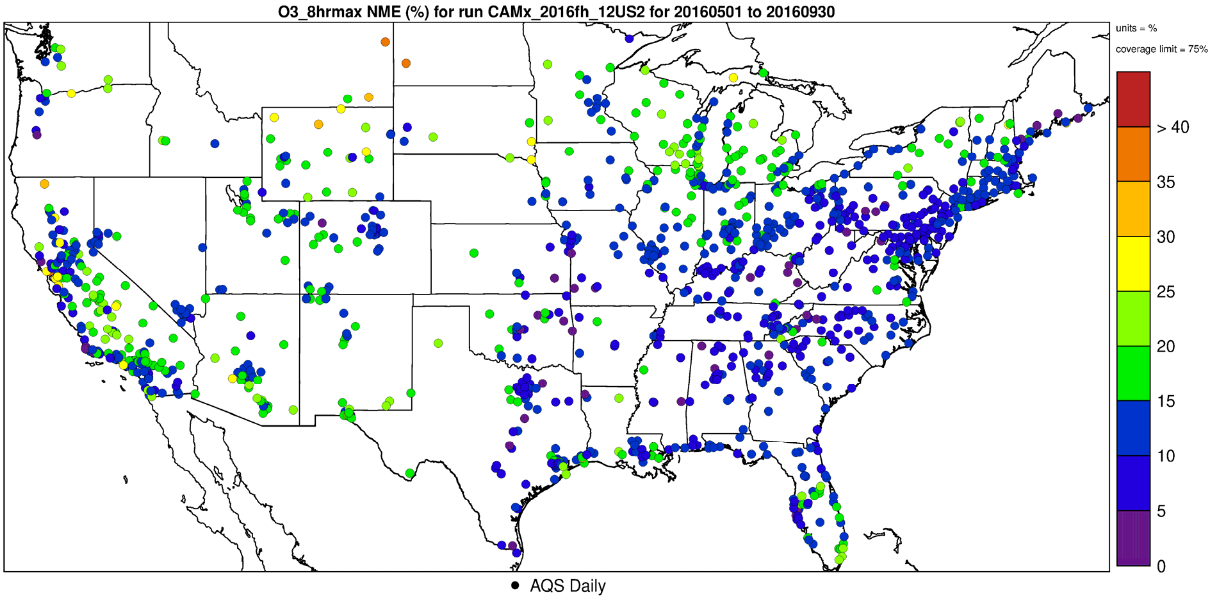


Figure A-15. Normalized Mean Error (%) of MDA8 ozone ≥ 60 ppb over the period May-September 2016.

Table A-2. Performance statistics for MDA8 ozone ≥ 60 ppb for May through September for monitoring sites in Stratford, CT, Westport, CT, New Haven-Madison, CT, and Houston-Aldine, TX.

State	Site Name	Number of Days ≥ 60 ppb	MB (ppb)	ME (ppb)	NMB (%)	NME (%)
CT	Stratford	36.0	-4.6	9.1	-6.4	12.9
CT	Westport	29.0	-5.7	9.2	-7.8	12.7
CT	New Haven-Madison	29.0	-4.6	7.3	-6.5	10.4
TX	Houston-Aldine	15.0	-4.2	8.8	-6.5	13.4

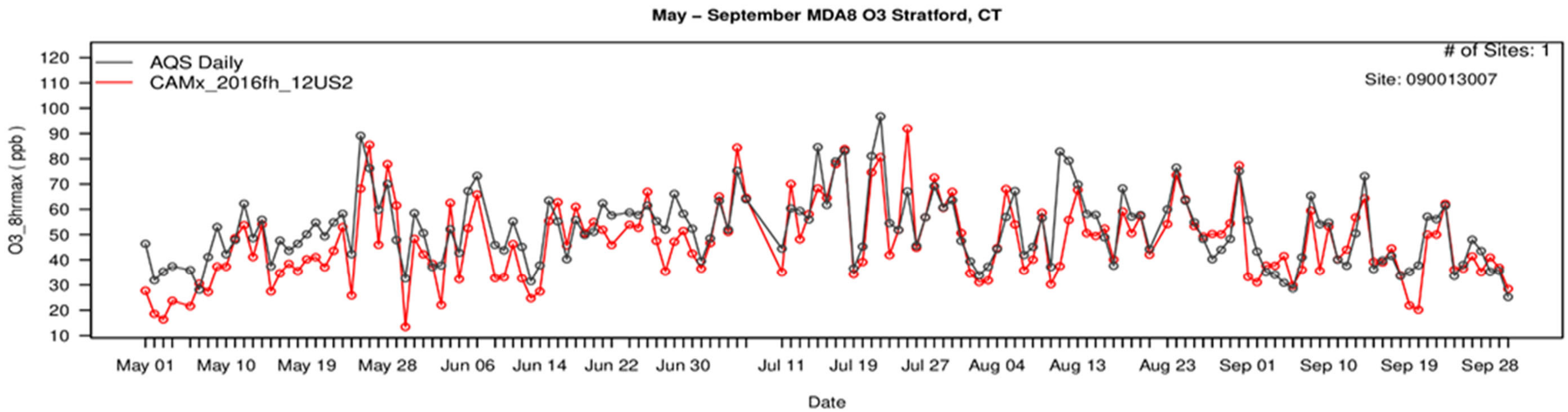


Figure A-16a. Time series of observed (black) and predicted (red) MDA8 ozone for May through September 2016 at site 090013007 in Stratford, Fairfield Co., Connecticut.

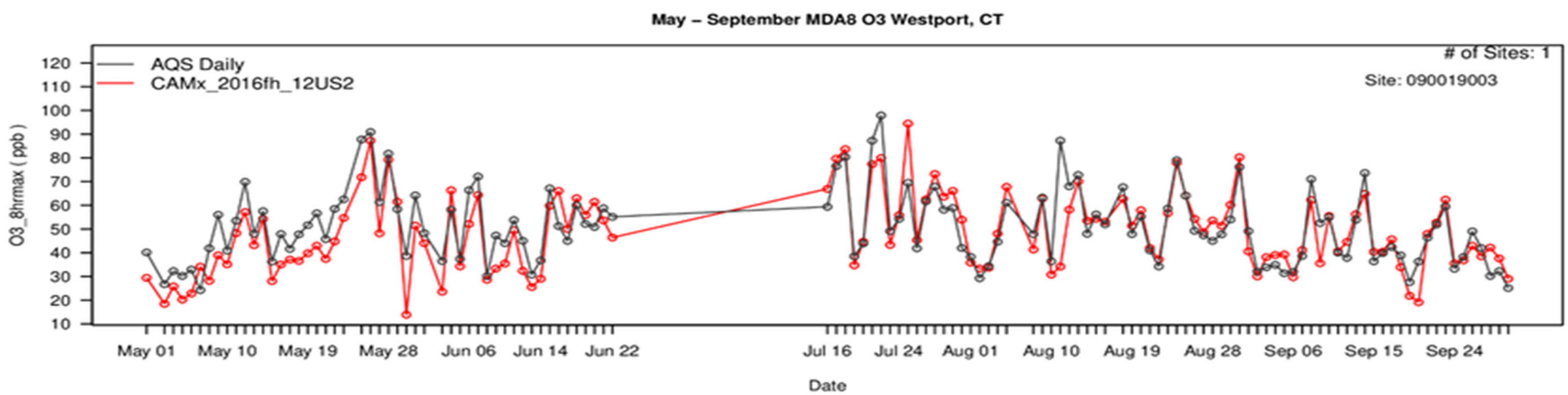


Figure A-16b. Time series of observed (black) and predicted (red) MDA8 ozone for May through September 2016 at site 090019003 in Westport, Fairfield Co., Connecticut.

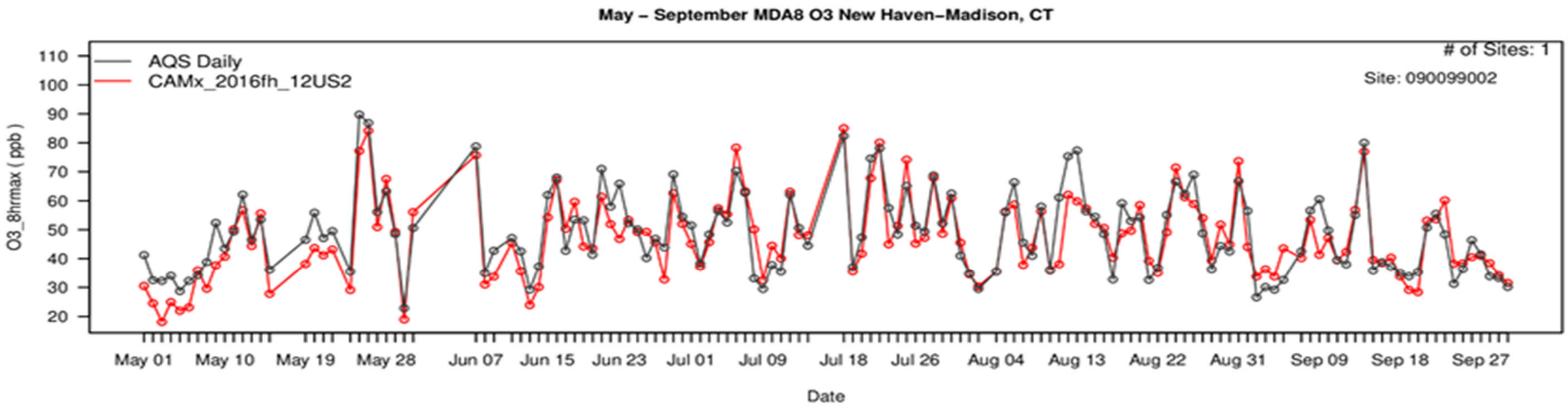


Figure A-16c. Time series of observed (black) and predicted (red) MDA8 ozone for May through September 2016 at site 090099002 in Madison, New Haven Co., Connecticut.

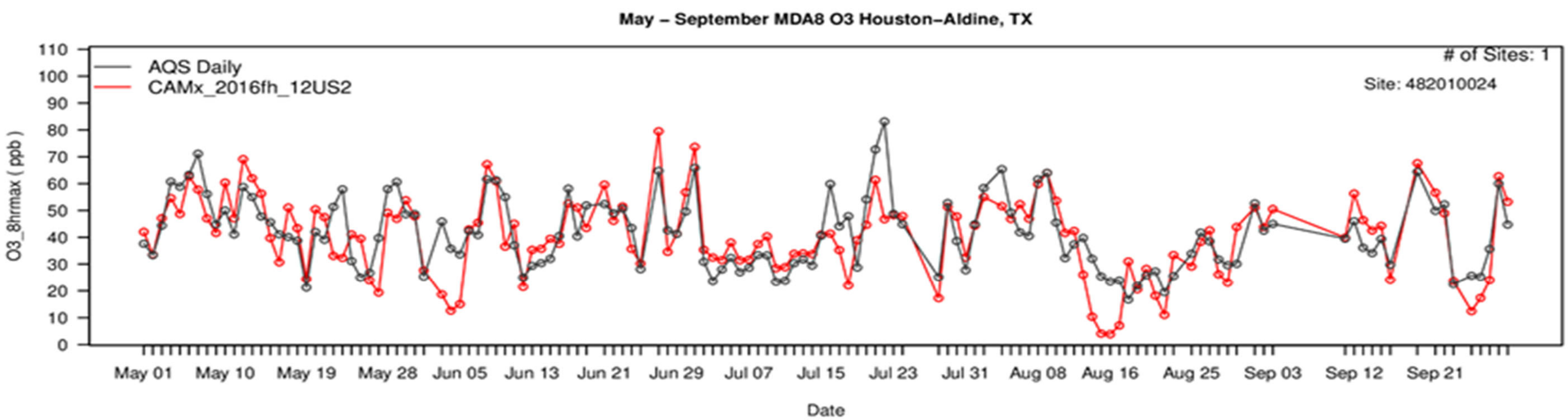


Figure A-16d. Time series of observed (black) and predicted (red) MDA8 ozone for May through September 2016 at site 482010024 in Harris Co., Texas.

Appendix B

Computation Steps for Calculating the Average Contribution Metric

Step 1. Modeled hourly ozone concentrations are used to calculate the 8-hour daily maximum ozone (MDA8) concentration in each grid cell on each day.

Step 2. The gridded hourly ozone contributions from each tag are subtracted from the corresponding gridded hourly total ozone concentrations to create a "pseudo" hourly ozone value for each tag for each hour in each grid cell.

Step 3. The hourly "pseudo" concentrations from Step 2 are used to calculate 8-hour average "pseudo" concentrations for each tag for the time period that corresponds to the MDA8 concentration from Step 1. Step 3 results in spatial fields of 8-hour average "pseudo" concentrations for each grid cell for each tag on each day.

Step 4. The 8-hour average "pseudo" concentrations for each tag and the MDA8 concentrations are extracted for those grid cells containing ozone monitoring sites. We used the data for the 10 days with the highest MDA8 modeled concentrations in 2023 (i.e., top 10 2023 modeled concentration days) in the downstream calculations. If there were fewer than 52023 exceedance days at a particular monitoring site then the data from the top five 2023 MDA8 concentration days are extracted and used in the calculations.[1]

Step 5. For each monitoring site and each tag, the 8-hour "pseudo" concentrations are then averaged across the days selected in Step 4 to create a multi-day average "pseudo" concentration for tag at each site. Similarly, the MDA8 concentrations were average across the days selected in Step 4.

Step 6. The multi-day average "pseudo" concentration and the corresponding multi-day average MDA8 concentration are used to create a Relative Contribution Factor (RCF) for each tag at each monitoring site. The RCF is the difference between the MDA8 concentration and the corresponding "pseudo" concentration, normalized by the MDA8 concentration.

---

[1] If there were fewer than 5 days with a modeled 2023 MDA8 concentration ≥ 60 ppb for the location of a particular monitoring site, then contributions were not calculated at that monitor.

Step 7. The RCF for each tag is multiplied by the 2023 average ozone design value to create the ozone contribution metrics for each tag at each site. Note that the sum of the contributions from each tag equals the 2023 average design value for that site.

Step 8. The contributions calculated from Step 7 are truncated to two digits to the right of the decimal (e.g., a calculated contribution of 0.78963… is truncated to 0.78 ppb). As a result of truncation the tabulated contributions may not always sum to the 2023 average design value.

Appendix C

Ozone Contributions to 2021 Nonattainment & Maintenance-Only Receptors

The tables in this appendix provide the contribution metric data from each state and the other source tags to the 2021 nonattainment and maintenance-only receptors. The table also contains the 2016-Centered and 2021 projected ozone design values at each site. The contributions and design values are in units of ppb.

A spreadsheet file with the 2021, 2023, and 2028 contributions to monitoring sites nationwide can be found in the following file in the docket for this proposed rule: Ozone Design Values & Contributions_Proposed Revised CSAPR Update.xls. Note that not all monitoring sites are included in the data sets for all three projection years because of the criteria used in the calculation of projected design values and contributions as described in this TSD.

								Contributions					
AQS Site ID	State	County	Location	2016-Centered Average DV	2016-Centered Maximum DV	2021 Average DV	2021 Maximum DV	AL	AZ	AR	CA	CO	CT
90013007	CT	Fairfield	Stratford	82.0	83	76.5	77.4	0.11	0.01	0.18	0.03	0.06	4.16
90019003	CT	Fairfield	Westport	82.7	83	78.5	78.8	0.11	0.01	0.17	0.03	0.06	2.73
90099002	CT	New Haven	Madison	79.7	82	73.9	76.1	0.07	0.01	0.15	0.02	0.05	3.96
482010024	TX	Harris	Houston	79.3	81	75.5	77.1	0.27	0.00	0.08	0.00	0.01	0.00

				Contributions											
AQS Site ID	State	County	Location	DE	DC	FL	GA	ID	IL	IN	IA	KS	KY	LA	ME
90013007	CT	Fairfield	Stratford	0.43	0.04	0.07	0.16	0.03	0.69	0.99	0.15	0.13	0.78	0.27	0.01
90019003	CT	Fairfield	Westport	0.43	0.04	0.07	0.16	0.02	0.81	1.26	0.17	0.13	0.87	0.27	0.00
90099002	CT	New Haven	Madison	0.53	0.05	0.02	0.08	0.02	0.80	1.08	0.22	0.11	0.79	0.15	0.01
482010024	TX	Harris	Houston	0.00	0.00	0.19	0.05	0.00	0.02	0.02	0.01	0.01	0.02	4.68	0.00

				Contributions											
AQS Site ID	State	County	Location	MD	MA	MI	MN	MS	MO	MT	NE	NV	NH	NJ	NM
90013007	CT	Fairfield	Stratford	1.21	0.35	1.16	0.16	0.10	0.36	0.08	0.07	0.01	0.10	7.70	0.03
90019003	CT	Fairfield	Westport	1.20	0.08	1.71	0.19	0.10	0.36	0.07	0.07	0.01	0.01	8.62	0.03
90099002	CT	New Haven	Madison	1.56	0.16	1.62	0.27	0.07	0.33	0.08	0.09	0.00	0.02	5.71	0.02
482010024	TX	Harris	Houston	0.00	0.00	0.00	0.00	0.37	0.02	0.00	0.00	0.00	0.00	0.00	0.03

				Contributions											
AQS Site ID	State	County	Location	NY	NC	ND	OH	OK	OR	PA	RI	SC	SD	TN	TX
90013007	CT	Fairfield	Stratford	14.42	0.56	0.10	2.34	0.20	0.03	6.72	0.04	0.17	0.04	0.31	0.58
90019003	CT	Fairfield	Westport	14.44	0.56	0.08	2.55	0.19	0.02	6.86	0.01	0.18	0.04	0.32	0.59
90099002	CT	New Haven	Madison	12.54	0.57	0.12	2.35	0.14	0.02	5.64	0.01	0.08	0.06	0.24	0.36
482010024	TX	Harris	Houston	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.05	32.68

				Contributions												
AQS Site ID	State	County	Location	UT	VT	VA	WA	WV	WI	WY	TRIBAL	CN & MX	Offshore	Fires	IC/BC	Biogenics
90013007	CT	Fairfield	Stratford	0.03	0.02	1.29	0.06	1.45	0.21	0.08	0.00	2.35	0.76	0.26	19.93	4.60
90019003	CT	Fairfield	Westport	0.03	0.01	1.30	0.05	1.49	0.23	0.08	0.00	2.58	0.68	0.35	21.07	4.78
90099002	CT	New Haven	Madison	0.02	0.01	1.69	0.06	1.55	0.23	0.07	0.00	3.02	1.07	0.25	20.84	4.72
482010024	TX	Harris	Houston	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.25	3.60	1.14	29.65	2.07