Attainment Demonstration Modeling for the 2008 Ozone National Ambient Air Quality Standard

Technical Support Document

Lake Michigan Air Directors Consortium

9501 W. Devon Ave., Suite 701

Rosemont, IL 60018

Please direct question/comments to adelman@ladco.org

November 19, 2020

This page is intentionally left blank

Contents

Document Change Log

Version	Date	Comments/Changes
Draft 1	July 1, 2020	First draft to LADCO states
Draft 2	July 29, 2020	Integrated comments from WI DNR
Draft 3	October 22, 2020	Integrated comments from U.S. EPA Region 5; updated modeling results with LADCO simulation CAMx_LADCO_2016aa2a
Final	November 19, 2020	

Errata/Known Issues

#	Description

Figures

Tables

Executive Summary

LADCO prepared this Technical Support Document to support the development of the O_3 NAA SIPs for the states of Wisconsin, Illinois, and Indiana pursuant to the 2008 O_3 NAAQS. LADCO used the Comprehensive Air Quality Model with Extensions (CAMx) v7.0 beta 4 to support these analyses. The LADCO CAMx modeling results are used here to identify O_3 monitoring sites that may have nonattainment or maintenance problems for the 2008 O_3 NAAQS by the attainment date in July 2021. Because the attainment deadline occurs during the 2021 O_3 season, the effective attainment deadline is the end of the 2020 O_3 season and thus resulted in the selection of 2020 as the projection year for this modeling application. LADCO used 2016 as the base modeling year from which we projected air quality in 2020.

LADCO based our 2020 O_3 air quality and NAA attainment forecasts on the CAMx modeling platform released by the U.S. EPA in September 2019 to support regional haze progress assessments. LADCO estimated 2020 emissions for most of the anthropogenic inventory sectors by interpolating between the 2016 and 2023 Inventory Collaborative 2016v1 inventories. We used linear interpolation for the emissions because 2020 inventories were not readily available for all of the sectors at the time that this application initiated. LADCO replaced the Electricity Generating Unit (EGU) emissions in the U.S. EPA 2016fh_16j platform with 2020 EGU forecasts estimated with the ERTAC EGU Tool version 16.1. ERTAC EGU 16.1 integrated state-reported information on EGU operations and forecasts as of December 2019. Overall both the NOx and VOC ozone season day emissions are projected to decrease in 2020 relative to 2016 in all of the LADCO states. The NOx reductions range from 10%-17%, driven primarily by reductions in mobile source emissions. The VOC reductions are more modest, at around 1% in each of the LADCO states and are also driven by reductions in mobile sources. For the anthropogenic sectors only (i.e., excluding biogenics), the ozone season day VOC emissions reductions are closer to 5% in each of the LADCO states.

The LADCO 2020 CAMx simulation predicted lower seasonal maximum O_3 concentrations across the majority of the modeling domain with the largest reductions occurring in the southeast U.S., east Texas, and the Central Valley of California. CAMx predicts that in 2020 the seasonal maximum daily maximum 8-hour average (MDA8) O_3 concentrations will decline along the western Lake Michigan shoreline in the range of 1-5 ppb compared with 2016.

The LADCO 2020 CAMx simulation predicts that no monitor in the region will have an average future year design value (DV_{2020}) that exceeds the 2008 O_3 NAAQS. The O_3 relative reduction factors (RRFs) in the Chicago NAA are in the range of 4-5%. The modest changes to the DVs in 2020 are due primarily to the short time period between the base and future years.

Excluding water cells in the attainment test calculation results in lower DVs_{2020} for the lakeshore monitors in the LADCO region.

1 Introduction

The Lake Michigan Air Directors Consortium (LADCO) was established by the states of Illinois, Indiana, Michigan, and Wisconsin in 1989. The four states and EPA signed a Memorandum of Agreement (MOA) that initiated the Lake Michigan Ozone Study and identified LADCO as the organization to oversee the study. Additional MOAs were signed by the states in 1991 (to establish the Lake Michigan Ozone Control Program), January 2000 (to broaden LADCO's responsibilities), and June 2004 (to update LADCO's mission and reaffirm the commitment to regional planning). In March 2004, Ohio joined LADCO. Minnesota joined the Consortium in 2012. LADCO consists of a Board of Directors (i.e., the State Air Directors), a technical staff, and various workgroups. The main purposes of LADCO are to provide technical assessments for and assistance to its member states, to provide a forum for its member states to discuss regional air quality issues, and to facilitate training for staff in the member states.

On March 12, 2008, the U.S. EPA revised the primary and secondary National Ambient Air Quality Standard (NAAQS) for ozone (O_3), strengthening the standard to a level of 0.075 parts per million (ppm) for a maximum daily 8-hour average. The form of the 8-hour O_3 NAAQS remained the same as the previous standard, the annual fourth-highest daily maximum averaged over three consecutive years. When U.S. EPA adopts a new or revises an existing NAAQS, it is required by Section 107(d)(1) of the Clean Air Act (CAA) to designate areas as nonattainment, attainment, or unclassifiable. Accordingly, on May 21, 2012, U.S. EPA designated Sheboygan County in eastern Wisconsin as a "marginal" O_3 nonattainment area (NAA) based on 2008-2010 ambient air quality data. On June 11, 2012, U.S. EPA designated the Chicago metropolitan area, including all or portions of eight counties in Illinois, two counties in northwest Indiana (Lake and Porter), and one partial county in southeast Wisconsin (Kenosha) as a "marginal" O_3 NAA based on monitoring data from 2009-2011. The attainment deadline for marginal NAAs to meet the 2008 O_3 NAAQS was July 20, 2015.

On April 11, 2016, U.S. EPA determined that the Chicago metropolitan area failed to attain the 2008 O_3 NAAQS by the applicable attainment date and thus reclassified the area as a “moderate” O_3 NAA. On September 28, 2016, U.S. EPA made a similar determination for Sheboygan County. The attainment deadline for moderate NAAs to meet the 2008 O_3 NAAQS was July 20, 2018.

On August 23, 2019, U.S. EPA determined that the entire Chicago metropolitan area again failed to attain the NAAQS and thus reclassified the area as a “serious” O_3 NAA. On July 15, 2019 EPA approved a revision to the Sheboygan County designation that splits the county into two distinct O_3 NAAs: shoreline and inland. In this same action, U.S. EPA approved a clean data determination for inland Sheboygan County. On July 10, 2020 the U.S. EPA officially redesignated both inland and shoreline Sheboygan County areas to attainment of the 2008 O_3 NAAQS.

The Chicago and Sheboygan nonattainment areas are shown in Figure 1. As a result of the actions for the Chicago NAA described above, the states of Illinois, Indiana, and Wisconsin must submit State Implementation Plans (SIPs) that meet the requirements applicable to “serious” O_3 NAAs. The NAA SIPs, or attainment demonstrations, must include a demonstration which identifies emissions reduction strategies sufficient to achieve the NAAQS by July 20, 2021, the attainment date for serious NAAs. Because the attainment deadline occurs during the 2021 O_3 season, the effective attainment deadline is the end of the 2020 O_3 season.

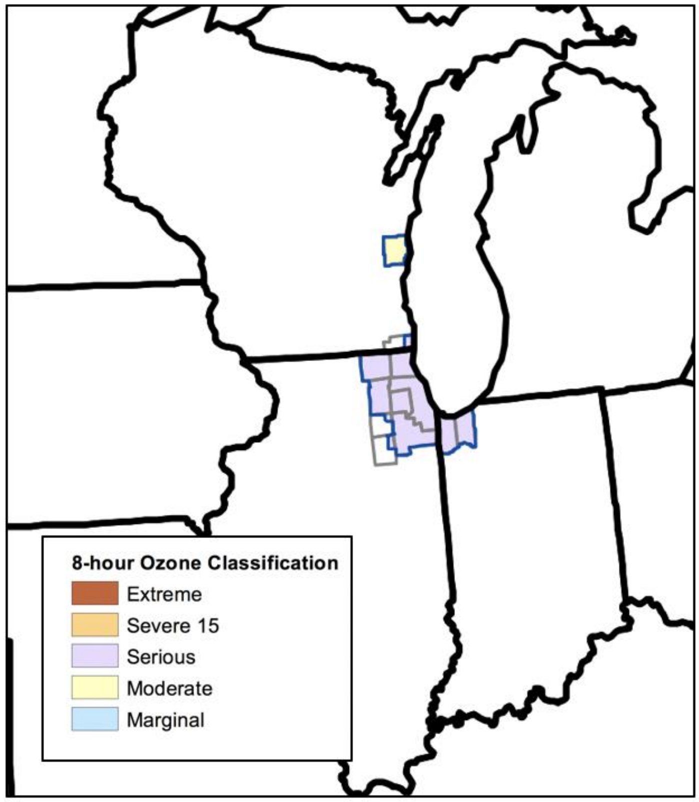

Figure 1. Nonattainment areas in the Lake Michigan region for the 2008 O_3 NAAQS (Source: U.S. EPA, May 2020).

One of LADCO's responsibilities is to provide technical air quality modeling guidance and support to the LADCO states. LADCO prepared this Technical Support Document (TSD) to support the development of the O_3 NAA SIPs for the states of Wisconsin, Illinois, and Indiana pursuant to the 2008 O_3 NAAQS. The analyses prepared by LADCO include preparation of modeling emissions inventories for the base year (2016) and the projected year of attainment (2020), evaluation and application of meteorological and photochemical grid models, analysis of ambient monitoring data, and a modeled attainment test for surface O_3 monitors in the Chicago NAA.

1.1 Project Overview

LADCO conducted regional air quality modeling to support the statutory obligations of the LADCO states under Clean Air Section 172**.** These SIP revisions are plans that describe how states with designated NAAs will bring those areas back into attainment of the NAAQS. LADCO used the Comprehensive Air Quality Model with Extensions (CAMx[1]) to support these analyses. In particular, LADCO used CAMx version 7.0 beta 4 to predict O_3 concentrations in 2020 to determine if current emissions control programs in the region will lead to attainment of the 2008 O_3 NAAQS.

This document describes how LADCO used CAMx modeling to project air quality from a 2016 base year to 2020, and to evaluate if the 2008 O_3 NAAQS NAAs in the LADCO region are predicted to attain the standard. The CAMx modeling outputs of this work are being presented to the IL, IN, and WI state air programs to support their 2008 O_3 NAAQS NAA SIP revisions that are due to EPA on August 3, 2020.

1.2 Organization of the Technical Support Document

This technical support document (TSD) is presented to the LADCO member states for estimating year 2020 O_3 future design values (DVFs). The TSD is organized into the following sections.

- Section 2 describes current surface O_3 conditions in the LADCO region and trends in O_3 concentrations over the past decade
- Section 3 describes the 2016 base year modeling and performance evaluation methods.
- Section 4 describes the 2020 CAMx air quality modeling platform that LADCO used to predict surface O_3 concentrations in 2020.

[1] www.camx.com

- Section 5 describes the approach used for estimating the O_3 DVFs. This section also includes a discussion on the methods used for identifying sites that are forecast to have O_3 NAAQS attainment problems.
- Section 6 presents a discussion of the performance evaluation and modeling results that the LADCO states can use to support their 2008 O_3 NAAQS NAA SIPs.
- The TSD concludes with a summary of significant findings and observations from the LADCO modeling.

2 2016 Ambient Air Quality Data Analysis

LADCO retrieves and conducts analysis on surface O_3 data collected at routine and special-purpose ambient monitors throughout the region. The current monitored O_3 design values (DVs), or the three-year average of the 4th highest daily maximum, 8-hour average O_3 concentrations, are presented in this section along with a discussion of trends in O_3 DVs and other metrics for tracking the changes in surface O_3 concentrations in the region. Design values are labeled by the last year of the three year average. For example, the 2019 O_3 DV is the average of the annual 4th highest daily maximum 8-hour average O_3 concentrations for the years 2017-2019.

2.1 Current Conditions

Figure 2 and Figure 3 are maps of the 2019 and 2020 O_3 design values (DVs) for the surface monitors around Lake Michigan. In Figure 2 warm colors represent O_3 concentrations approaching the 2008 O_3 NAAQS of 75 ppb; sites that are colored red in these plots indicate a violation of the 2008 standard. The 2019 DVs are based on validated data reported to the U.S. EPA. The 2020 DVs plot uses a different color scale, and these data are preliminary and based on unvalidated data reported through July 29, 2020. Note that several months remain in the O_3 season in 2020 and the values will change before the 2020 data become official. Table 1 and Table 2 show the same DVs data in tabulated form. Table 1 shows the annual DVs by 2008 O_3 NAAQS NAA from 2013 to present; the NAA DV is a reading from the "controlling" monitor, or the monitor with the highest 3-year DV in the entire NAA. Table 2 shows the annual DVs for key monitors in the Chicago 2008 O_3 NAAQS NAAs from 2013 to present.

The DV tables and figures show that no monitors in the Chicago NAA have 2019 3-year DVs that violate the 2008 O_3 NAAQS. Through September 30, 2020 one monitor in the Chicago NAA (Northbrook, IL) has a 2020 3-year DV that violates the standard. Between 2013 and 2018, 2015 was the last DV year in which there were no violations of the 2008 O_3 NAAQS in the Chicago NAA.

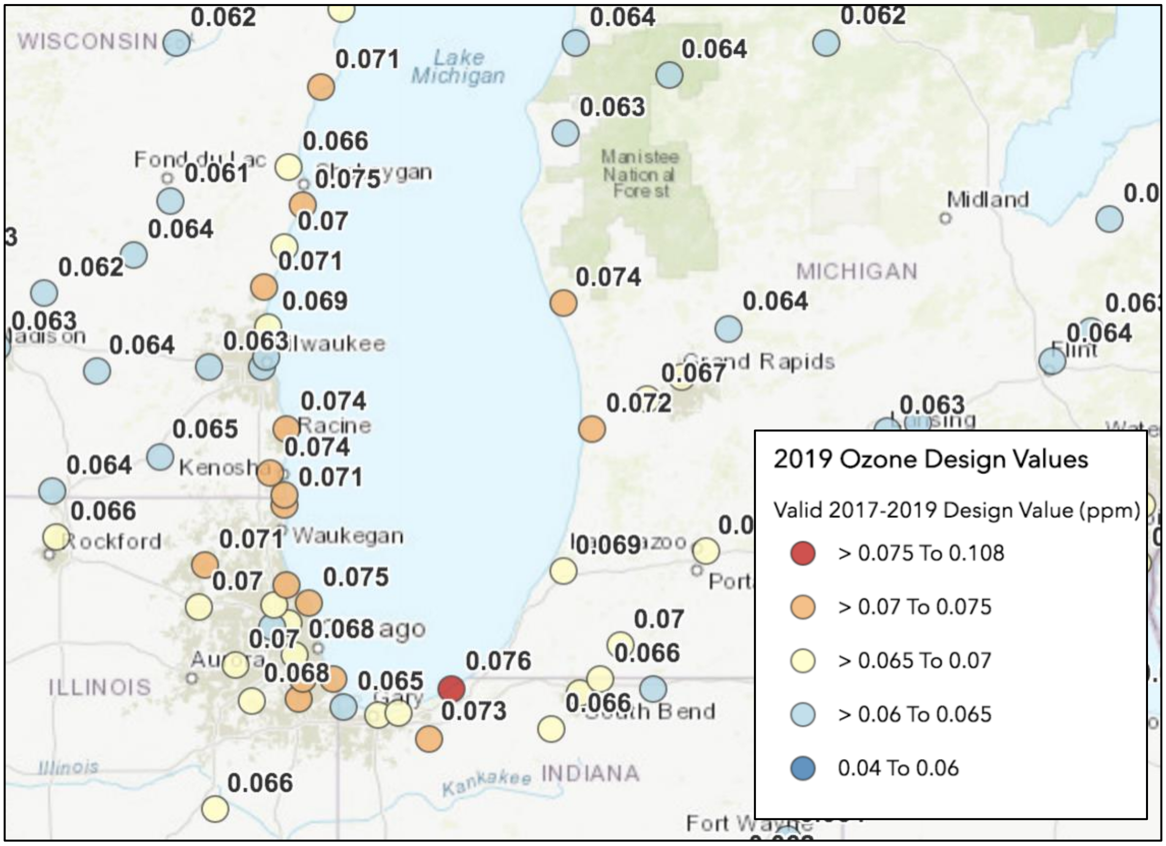

Figure 2. 2019 O_3 design values

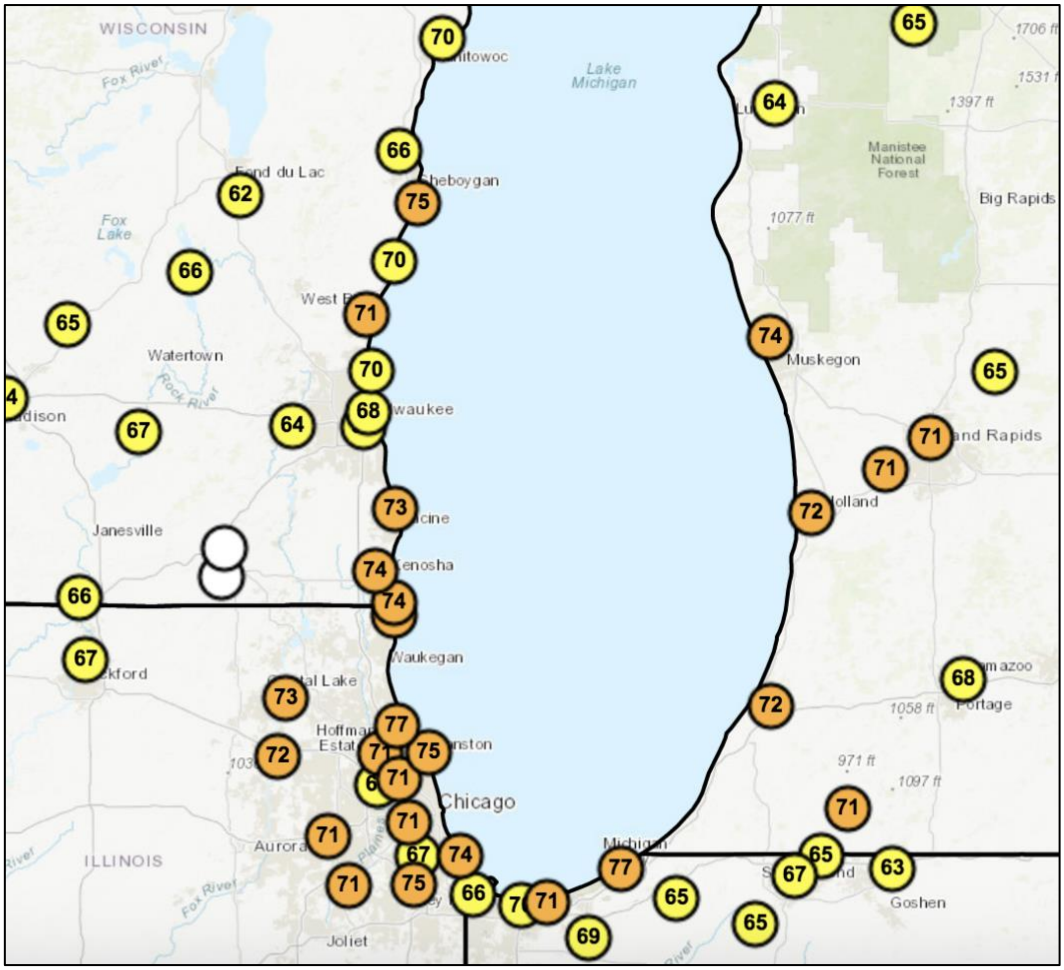

Figure 3. 2020 O_3 design values through September 30, 2020

Table 1. Chicago 2008 O_3 NAAQS NAA design values (ppb) [Source: U.S. EPA Green Book, May 2020]

Designated Area	2013	2014	2015	2016	2017	2018	2019	2020*
Chicago-Naperville, IL-IN-WI	82	81	75	77	78	79	75	77

* 2020 data are preliminary and incomplete; they were retrieved September 30, 2020 from AirNow Tech.

Table 2. Chicago 2008 O_3 NAAQS NAA monitor design values (ppb) [Source: U.S. EPA

Green Book, May 2020]

State	County	AQS Site ID	2013	2014	2015	2016	2017	2018	2019	2020*
NAA: Chicago, IL-IN-WI										
IL	Cook	170310001	71	69	65	69	73	77	75	71
IL	Cook	170310032	80	76	68	70	72	75	73	74
IL	Cook	170310076	72	70	64	69	72	75	72	67
IL	Cook	170311003	70		66	69	67	69	67	67
IL	Cook	170311601	71	71	66	69	69	70	68	68
IL	Cook	170313103			61	62	62	64	63	63
IL	Cook	170314002	72	69	62	66	68	72	68	68
IL	Cook	170314007	68	69	68	71	71	74	70	71
IL	Cook	170314201	77	74	68	71	72	77	74	75
IL	Cook	170317002	80	78	70	72	73	77	75	75
IL	DuPage	170436001	68	67	64	68	70	71	70	66
IL	Kane	170890005	69	68	65	68	69	71	70	72
IL	Lake	170971007	80	79	71	73	73	75	71	71
IL	Will	171971011	64	65	63	64	65	67	66	64
IN	Lake	180890022	69	69	65	67	68	70	68	70
IN	Lake	180892008			63	65		66	65	65
WI	Kenosha	550590019	82	81	75	77	78	79	75	72
WI	Kenosha	550590025			69	71	73	77	74	72

* 2020 data are preliminary and incomplete; they were retrieved September 30, 2020 from AirNow-Tech.

2.2 Meteorology and Transport

Ozone concentrations are significantly influenced by meteorological factors. Ozone production is driven by high temperatures and sunlight, as well as precursor concentrations. Ozone concentrations at a given location are also dependent on wind direction, which governs which sources or source regions are upwind. Wind-drive transport in turn affects how much ozone and ozone precursors impact a given area.

Qualitatively, O_3 episodes in the region are associated with hot weather, clear skies (sometimes hazy), low wind speeds, high solar radiation, and winds with a southerly component. These conditions are often a result of a slow-moving high pressure system to the east of the region. The relative importance of various meteorological factors is discussed later in this section. Transport of O_3 and its precursors is a significant factor and occurs on several spatial scales. Regionally, over a multi-day period, somewhat stagnant summertime conditions can lead to the build-up in O_3 and O_3 precursor concentrations over a large spatial area. This polluted air mass can be transported long distances, resulting in elevated O_3 levels in locations far downwind. Locally, emissions from urban areas add to the regional background leading to O_3 concentration hot spots downwind. Depending on the synoptic wind patterns (and local land-lake breezes), different downwind areas are affected.

The following key findings related to transport can be made:

- Ozone transport is an issue affecting many portions of the eastern U.S. The Lake Michigan area (and other areas in the LADCO region) both receives high levels of incoming (transported) O_3 and O_3 precursors from upwind source areas on many hot summer days, and contributes to the high levels of O_3 and O_3 precursors affecting downwind receptor areas.
- The presence of Lake Michigan influences the formation and transport of O_3 in the region, particularly at sites within a few kilometers of the shoreline . Depending on large-scale synoptic winds and local-scale lake breezes, different parts of the area experience high O_3 concentrations. For example, during southerly flow, high O_3 can occur in eastern Wisconsin, and during southwesterly flow, high O_3 can occur in western Michigan.
- Downwind shoreline areas around Lake Michigan are affected by transport of O_3 from major cities in the Lake Michigan area and from areas further upwind.

2.3 Ozone Trends

Figure 4 illustrates the 19-year trends in 3-year O_3 DVs at individual surface monitors in the Chicago NAA. The red horizontal lines mark the 2015 and 2008 O_3 NAAQS. After the decadal high year in 2012, surface O_3 concentrations have declined through 2019. While there has been an increasing trend in O_3 concentrations in the Chicago NAA monitors since the decadal low year in 2015, 2018 was the only year since 2015 that monitored 2008 O_3 NAAQS violations at multiple sites in the NAA.

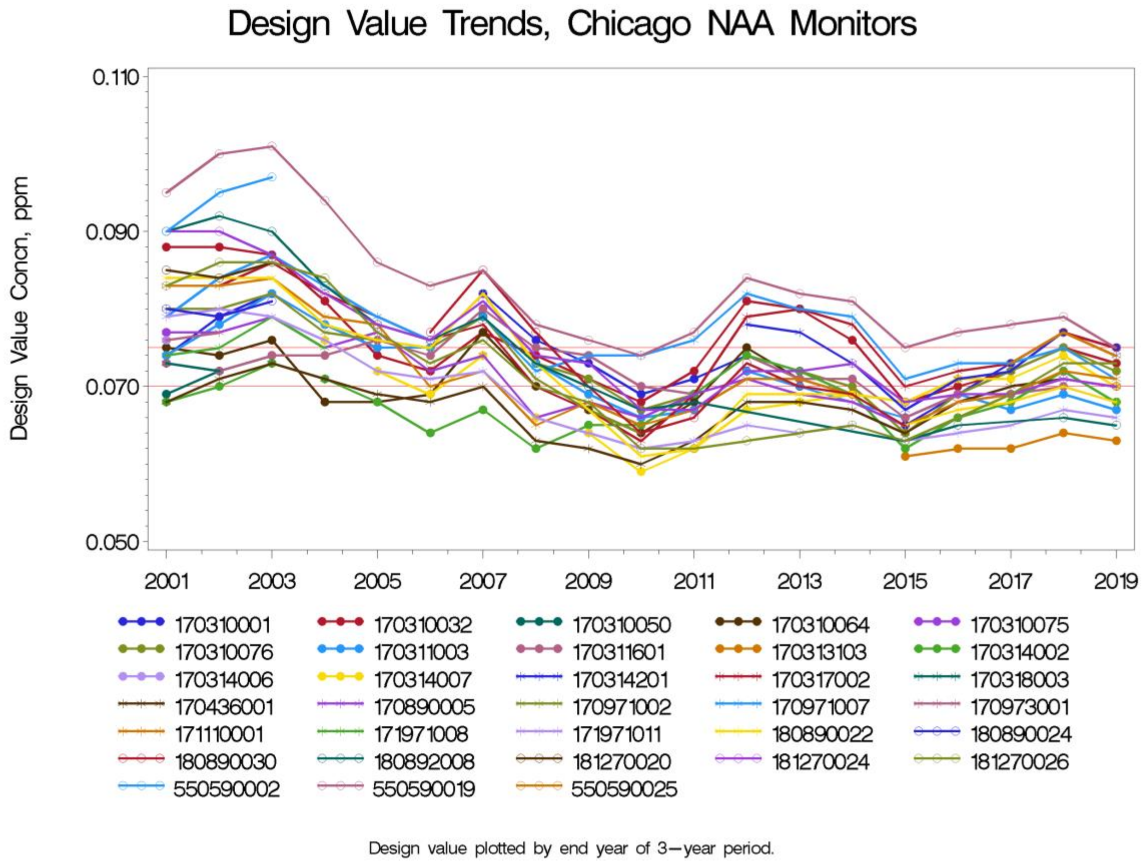

Figure 4. 3-year O_3 design value trends from 2001 to 2019 at all monitors in the Chicago NAA

Given the effect of meteorology on ambient O_3 levels, year-to-year variations in meteorology can make it difficult to assess short term (e.g. – less than 10 years) trends in O_3 concentrations. One approach to adjust the trends in O_3 concentrations for meteorological influences is through the use of Classification and Regression Trees (CART). CART is a statistical technique which partitions data sets into similar groups (Breiman et al., 1984). LADCO performed a CART analysis using data for the period 2005-

2018 for urban and downwind monitors in the 2008 O_3 NAAQS NAAs. The CART model searches through over thirty National Weather Service meteorological variables collected at airports[2] to determine which are most efficient in predicting O_3. Although the exact selection of predictive variables changes from site to site, the most common predictors of high surface O_3 concentrations during the period we analyzed are temperature, wind direction, and relative humidity. Only occasionally were upper air variables, transport time or distance, lake breeze, or other variables significant as predictors.

For each group of monitors in the NAAs we analyzed, LADCO developed regression trees that classify each summer day (May-September) by its meteorological conditions. Similar days are assigned to nodes, which are equivalent to branches of the regression tree. By grouping days with similar meteorology, the influence of meteorological variability on changes in O_3 concentrations is partially controlled for in the trend; the remaining trend is presumed to be due to trends in precursor emissions or other non-meteorological influences.

Trends over the 13-year period ending in 2018 were found to be declining for each monitor or composite area noted. These plots reflect long term trends and are not meant to depict trends over shorter time periods.

2.3.1 Northern Chicago NAA CART Analysis

LADCO used O_3 data from the Zion, IL and Chiwaukee, WI monitoring sites to identify trends in the surface concentrations downwind of Chicago using CART. Meteorological surface and aloft data used in this analysis are from the National Climatic Data Center's Integrated Surface Database and Integrated Radiosonde Archive; we used HYSPLIT trajectories to develop transport vectors.

Figure 5 shows the distribution of O_3 among Zion and Chiwaukee CART nodes. Each boxplot represents a group of days with common meteorological conditions. Node U

[2] National Climatic Data Center Integrated Surface Database

identifies the predictor variables that are associated with the highest mean observed O_3 concentrations at these monitors during the period of analysis (2005-2018). The days captured by this node have an average daily maximum O_3 concentration of 74 ppb and the following meteorological conditions:

- 24-hr southerly transport vector distance is >39 km
- average relative humidity is <70%
- afternoon wind direction is <211 deg
- max temperature is >85 F

Node T identifies the predictor variables that are associated with the second highest mean observed O_3 concentrations at these monitors during the period of analysis. Node T captures days with an average daily maximum O_3 concentration of 65 ppb and the following meteorological conditions:

- 24-hr southerly transport vector is >39 km
- average relative humidity is <70%
- afternoon wind direction is < 211 deg
- max temperature is <85 F and >78 F

CART identifies that the most significant predictors of high O_3 concentrations at Zion and Chiwaukee are warm and dry conditions with southerly flow. Daily maximum temperature is the only meteorological difference between nodes T and U. With all transport variables being equal, the cooler conditions represented by node T group days with an average O_3 concentration that is 9 ppb lower than the warmer days (>85 F) captured in node U.

Figure 6 shows the Zion, IL and Chiwaukee, WI O_3 trends by CART node. The node associated with the highest O_3 concentrations (node U) shows a distinct downward trend in O_3 concentrations during the 13 year CART analysis period. By controlling for the meteorological influence on O_3 concentrations during the most polluted days, this trend indicates that O_3 concentrations in the northern part of the Chicago 2008 O_3

NAAQS NAA are declining as the result of changes to emissions and other non-meteorological predictors.

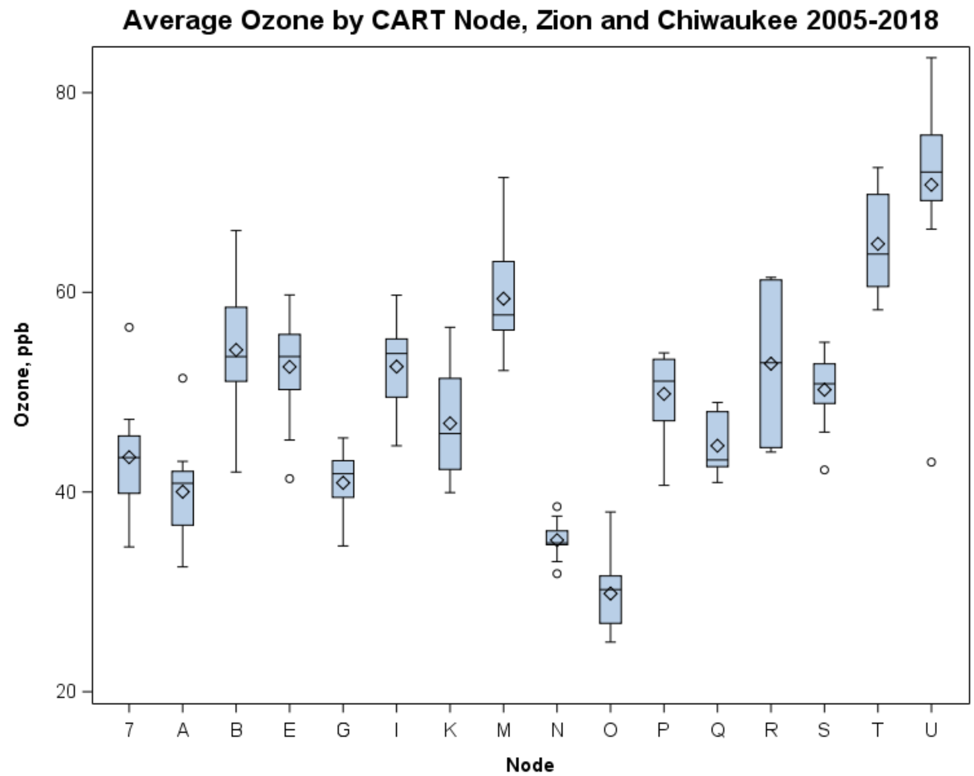

Figure 5. Northern Chicago NAA ozone concentrations by CART node.

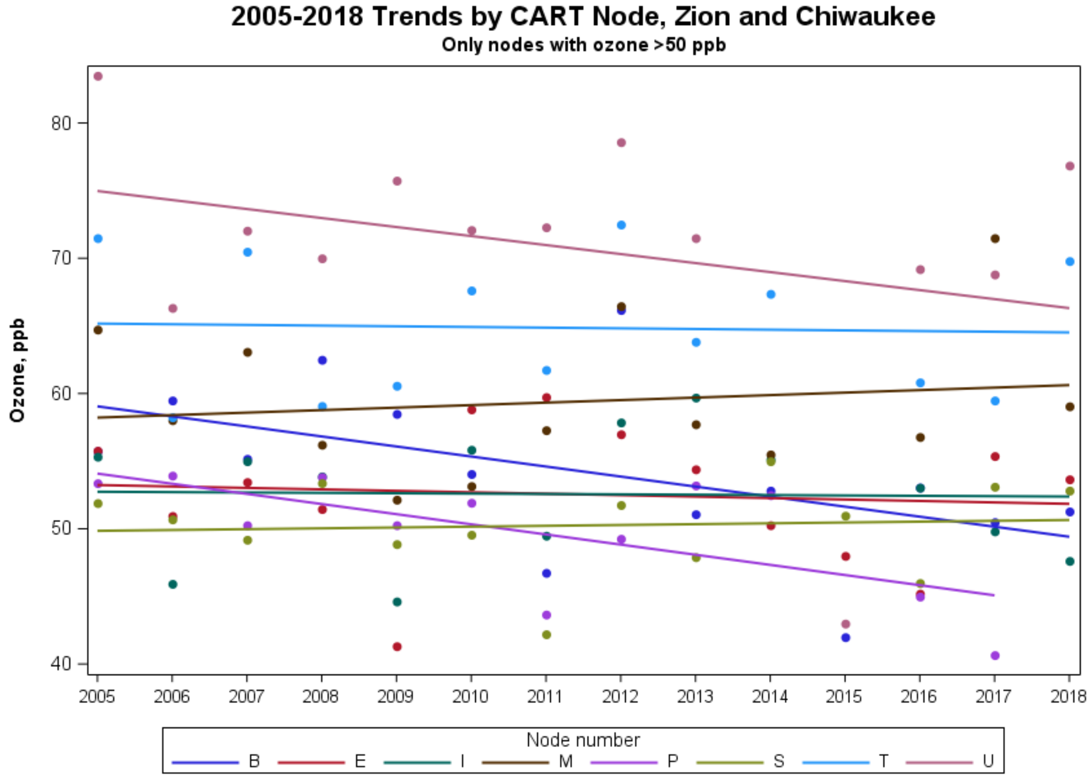

Figure 6. Northern Chicago NAA O_3 trends by CART node

2.4 Conceptual Model of Ozone in the Chicago NAA

A conceptual model is a qualitative summary of the physical, chemical, and meteorological processes that control the formation and distribution of pollutants in a given region. Based on the data and analyses presented above, and of previous conceptual models and technical support documents developed for the Lake Michigan region, a conceptual model of the behavior, meteorological influences, and causes of high O_3 in the Chicago NAA is summarized below:

- Monitoring data show that as of 2019 all of the surface O_3 monitoring sites in the western Lake Michigan region were meeting the 2008 8-hour O_3 NAAQS. Historical O_3 data show a downward trend over the past 19 years, due likely to federal and state emission control programs. Concentrations declined sharply from 2002 through 2010, and again from 2012 through 2015. Although ozone concentrations at the “controlling” monitors have been on the rise since 2015, there were no 3-year DVs in violation of the 2008 O_3 NAAQS at any monitor in the region in 2019.

- Ozone concentrations are strongly influenced by meteorological conditions, with more high O_3 days and higher O_3 levels during summers with above normal temperatures. Nevertheless, meteorologically adjusted trends at the controlling monitors show that concentrations have declined even on hot days, providing strong evidence that emission reductions of O_3 precursors have been effective.
- Inter- and intra-regional transport of O_3 and O_3 precursors affects many portions of the LADCO states, and is the principal cause of nonattainment in some areas far from population or industrial centers.
- The presence of Lake Michigan influences the formation, transport, and duration of elevated O_3 concentrations along its shoreline. Depending on large-scale synoptic winds and local-scale lake breezes, different parts of the area experience high O_3 concentrations. For example, under southerly flow, high surface O_3 concentrations can occur in eastern Wisconsin, and under southwesterly flow, high surface O_3 can occur in western Michigan.
- A natural lake-land breeze circulation pattern is a major cause of the high O_3 concentrations observed along the lakeshore. This pattern is driven by surface temperature gradients between the lake and the land. At night and in the early morning a land breeze (land –> lake) forms when the lake surface is warmer than the land surface. The land breeze transports O_3 precursors from industrial and mobile sources on land out over the lake. When the sun rises, the O_3 precursors over the lake begin to rapidly react to form O_3, and high over-lake concentrations are often observed during the summer. A lake breeze (lake –> land) forms when the land surface becomes warmer than the lake, typically in the early afternoon during the summer. The lake breeze transports the concentrated O_3 and precursors from the lake, inland to a narrow band along the lake shore. The O_3 concentrations observed along the lakeshore that violate the NAAQS are often associated with lake-land breeze patterns.

- Areas in closer proximity to the Lake shoreline display the most frequent and most elevated O_3 concentrations.

3 2016 Air Quality Modeling and Model Performance Evaluation

3.1 2016 Modeling Platform

LADCO based our 2016 O_3 air quality predictions on the 2016v1 National Emission Inventory Collaborative emissions inventory[3] and the U.S. EPA 2016ff CAMx modeling platform. The meteorology and initial and boundary conditions came from the U.S. EPA 2016ff CAMx modeling platform. LADCO processed most of the 2016 emissions using the U.S. EPA 2016fh_16j Sparse Matrix Operator Kernel Emissions (SMOKE) modeling platform. The CAMx inputs, including the meteorology data simulated with the Weather Research Forecast (WRF) model, emissions data, and boundary conditions represent year 2016 conditions. LADCO used the majority of the data and software provided by U.S. EPA for this platform, with a few exceptions described below.

3.2 Modeling Year Justification

LADCO selected 2016 as a modeling year for this study because at the initiation of this project in late 2019 CAMx input data for 2016 were widely available and they represented the state-of-the-science for emissions and meteorology data. In 2017 a group of multi-jurisdictional organizations (MJOs), states, and EPA established 2016 as the new base year for a national air quality modeling platform[4]. The group concluded that if only one recent year could be selected, that 2016 would serve as a good base year because of fairly typical O_3 conditions and average wildfire conditions. Following from the base year recommendations from that group, several modeling centers,

[3] http://views.cira.colostate.edu/wiki/wiki/10202
[4] Base Year Selection Workgroup Final Report

including U.S. EPA and LADCO, developed data and capabilities for simulating and evaluating air quality in 2016.

Following from the selection of 2016 as the base year for a national modeling platform, starting in late 2017, the MJOs, states, and EPA formed the National Emissions Inventory Collaborative to develop a 2016 emissions inventory and modeling platform. Over 200 participants collaborated across 12 workgroups to develop base and future year emissions to support upcoming regulatory modeling applications. This effort was designed to involve a broad group of emissions experts in the development of a new national emissions modeling platform. LADCO used the 2016 and 2023 inventories developed by the Collaborative for the modeling presented here as they were the most recent inventory data available at the initiation of this project.

LADCO selected 2020 as the future projection year because it aligns with the last O_3 season that will be used to determine attainment of the 2008 O_3 NAAQS.

3.3 Air Quality Model Configuration

LADCO based our CAMx air quality modeling platform for this application on the configuration that the U.S. EPA used for recent regional haze modeling (US EPA, 2019). LADCO used CAMx 7.0 beta 4 (Ramboll, 2018) as the photochemical grid model for this application. CAMx is a three-dimensional, Eulerian air quality model that simulates the chemical transformation and physical transport processes of air pollutants in the troposphere. It includes capabilities to estimate the concentrations of primary and secondary gas and particle phase air pollutants, and dry and wet deposition, from urban to continental spatial scales. As CAMx associates source-level air pollution emissions estimates with air pollution concentrations, it can be used to design and assess emissions reduction strategies pursuant to NAAQS attainment goals.

LADCO selected CAMx for this study because it is a component of recent U.S. EPA modeling platforms for investigating the drivers of ground level O_3 in the U.S. CAMx is a three-dimensional, Eulerian air quality model that simulates the chemical transformation and physical transport processes of air pollutants in the troposphere. It includes capabilities to estimate the concentrations of primary and secondary gas and particle phase air pollutants, and dry and wet deposition, from urban to continental spatial scales. As CAMx associates source-level air pollution emissions estimates with air pollution concentrations, it can be used to design and assess emissions reduction strategies pursuant to NAAQS attainment goals. As CAMx is a component of U.S. EPA studies with a similar scope to this project, LADCO was able to leverage the data and software elements that are distributed with recent U.S. EPA regulatory modeling platforms. Using these elements saved LADCO significant resources relative to building a modeling platform from scratch.

Figure 7 shows the U.S. EPA transport modeling domain for the continental U.S. A 12-km uniform grid (12US2) covers all of the continental U.S. and includes parts of Southern Canada and Northern Mexico. The domain has 36 vertical layers with a model top at about 17,550 meters (50 mb). LADCO used the same U.S. EPA 12-km domain for this project because it supported the use of meteorology, initial and boundary conditions, and emissions data that were freely available from U.S. EPA.

Table 3 summarizes the CAMx science configurations and options LADCO used for the 2016 and 2020 CAMx modeling for this application. We used the Piecewise Parabolic Method (PPM) advection solver for horizontal transport along with the spatially varying (Smagorinsky) horizontal diffusion approach. We used K-theory for vertical diffusion using the CMAQ-like vertical diffusivities from WRFCAMx. The CB6r4 gas-phase chemical mechanism was selected because it includes the latest chemical kinetic rates and represents improvements over the other alternative CB05 and SAPRC chemical mechanisms as well as active methane chemistry. Additional CAMx inputs were as follows:

> <u>Meteorological Inputs:</u> The U.S. EPA WRF-derived meteorological fields were processed to generate CAMx meteorological inputs using the WRFCAMx processor, as described in Section 3.4.

Initial/Boundary Conditions: LADCO used 2016 chemical boundary conditions for the 12-km continental U.S. modeling domain derived from the U.S. EPA northern hemisphere CMAQ simulations of 2016 (U.S. EPA, 2019c). The EPA 2016 ICBCs are hourly, vertically resolved up to 50 mb, and use the Carbon Bond 6 photochemical mechanism.

Photolysis Rates: LADCO prepared the photolysis rate inputs as well as albedo/haze/ozone/snow inputs for CAMx. Day-specific O_3 column data were based on the Total Ozone Mapping Spectrometer (TOMS) data measured using the satellite-based Ozone Monitoring Instrument (OMI). Albedo were based on land use data. For CAMx there is an ancillary snow cover input that will override the land use based albedo input. LADCO used the TUV photolysis rate processor to prepare clear-sky photolysis rates for CAMx. If there were periods of more than a couple of days where daily TOMS data were unavailable in 2016, the TOMS measurements were interpolated between the days with valid data; in the case where large periods of TOMS data were missing, monthly average TOMS data were used. CAMx was also configured to use the in-line TUV to adjust for cloud cover and account for the effects that modeled aerosol loadings have on photolysis rates; this latter effect on photolysis may be especially important in adjusting the photolysis rates due to the occurrence of PM concentrations associated with emissions from fires.

Landuse: LADCO used landuse/landcover data from the U.S. EPA WRF simulation.

Spin-Up Initialization: A minimum of ten days of model spin up (e.g., December 21-31, 2015) was used for the 12 km modeling domain. LADCO ran monthly CAMx simulations, initializing each month with a 14-day spin-up period.

As the focus of this study is on O_3, LADCO used CAMx to simulate the O_3 season. LADCO simulated April 1 through October 31, 2016 as individual months using 14-day model spin-up periods for each month. LADCO selected a CAMx configuration that was consistent with previous O_3 modeling applications performed by LADCO and U.S. EPA. U.S. EPA (2019) provided completed details their 2016 CAMx simulation, including a performance evaluation.

Table 3. LADCO 2016 CAMx modeling platform configuration

Science Options	Configuration
Model Codes	CAMx V7.0 beta 4
Simulation Period	March 20-October 31, 2016
Horizontal Grid Mesh	12 km, 396 col x 246 rows
Vertical Grid Mesh	36 layers as in WRF outputs
Grid Interaction	None
Initial Conditions	14 day spin-up on 12 km grid
Boundary Conditions	12km from hemispheric CMAQ (U.S. EPA 2016ff)
Emissions	
Baseline Emissions Processing	SMOKE, MOVES and BEIS
Sub-grid-scale Plumes	None
Chemistry	
Gas Phase Chemistry	CB6r4
Aerosol Chemistry	CF + SOAP
Meteorological Processor	WRFCAMx_v4.7
Horizontal Diffusion	Spatially varying
Vertical Diffusion	CMAQ-like in WRF2CAMx
Diffusivity Lower Limit	Kz_min = 0.1 to 1.0 m^2/s or 2.0 m^2/s
Dry Deposition	Zhang dry deposition scheme (CAMx)
Wet Deposition	CAMx-specific formulation
Gas Phase Chemistry Solver	Euler Backward Iterative (EBI) -- Fast Solver
Vertical Advection Scheme	Implicit scheme w/ vertical velocity update (CAMx)
Horizontal Advection Scheme	Piecewise Parabolic Method (PPM) scheme
Integration Time Step	Wind speed dependent

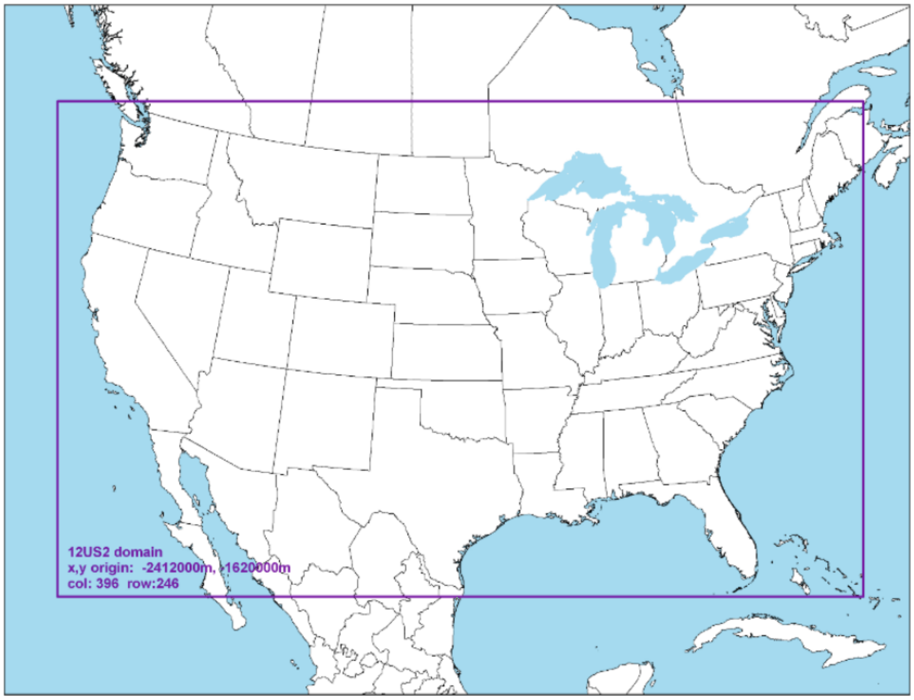

Figure 7. CAMx 12-km modeling domain (12US2)

3.4 Meteorology Data

LADCO used the U.S. EPA 2016 WRF data for this study (US EPA, 2019b). The U.S. EPA used version 3.8 of the WRF model, initialized with the 12-km North American Model (NAM) from the National Climatic Data Center (NCDC) to simulate 2016 meteorology. Complete details of the WRF simulation, including the input data, physics options, and four-dimensional data assimilation (FDDA) configuration are detailed in the Meteorology Model Performance for Annual 2016 Simulation WRFv3.8 report (US EPA, 2019b). LADCO prepared the WRF data for input to CAMx with version 4.6 of the WRFCAMx software.

3.5 Initial and Boundary Conditions

LADCO used 2016 initial and boundary conditions for CAMx generated by the U.S. EPA from a northern hemisphere simulation of the Community Multiscale Air Quality (CMAQ) model (US EPA, 2019c). EPA generated hourly, one-way nested boundary conditions (i.e., hemispheric-scale to regional-scale) from a 2016 108-km x 108-km polar stereographic CMAQ simulation of the northern hemisphere. Following the convention of the U.S. EPA 2016 regional haze modeling (U.S. EPA, 2019), LADCO used year 2016 CMAQ boundary conditions for modeling 2016 and 2020 air quality with CAMx.

3.6 Emissions Data

The 2016 emissions data for this study were based on the U.S. EPA 2016 v1 ("2016fh_16") emissions modeling platform (http://views.cira.colostate.edu/wiki/wiki/10202). U.S. EPA and the 2016 Collaborative generated this platform for use O_3 NAAQS and Regional Haze SIPs. Twelve different workgroups collaborated to construct 2016 and future year emissions estimates. The first version of the 2016 inventories used 2014 inventory data; later versions of the inventory fully integrated 2016 estimates of emissions activities, growth and controls, and the latest emissions factors. Table 4 lists the 2016 base year inventory components that LADCO used to simulate 2016 air quality for this application.

Table 4. LADCO 2016 emissions modeling platform inventory components

Sector	Abbreviation	Data Source	Year
Agriculture	ag	U.S. EPA 2016fh	2016
Airports	airports	U.S. EPA 2016fh	2016
Biogenic	beis	U.S. EPA 2016fh	2016 meteorology
C1/C2 Commercial Marine	cmv_c1c2	U.S. EPA 2016fh	2016
C3 Commercial Marine	cmv_c2	U.S. EPA 2016fh	2016
Nonpoint	nonpt	U.S. EPA 2016fh	2016
Offroad Mobile	nonroad	U.S. EPA 2016fh	2016
Nonpoint Oil & Gas	np_oilgas	U.S. EPA 2016fh	2016
Onroad Mobile	onroad	U.S. EPA 2016fh	2016
Point Oil & Gas	pt_oilgas	U.S. EPA 2016fh	2016
Agricultural Fires	ptagfire	U.S. EPA 2016fh	2016
Electricity Generation	ptertac	ERTAC 16.1 + Hourly CEMs	2016
Wild and Prescribed Fires	ptfire	U.S. EPA 2016fh	2016
Industrial Point	ptnonertac	U.S. EPA 2016fh	2016
Rail	rail	U.S. EPA 2016fh	2016

Residential Wood Combustion	rwc	U.S. EPA 2016fh	2016
Mexico Anthropogenic	othar/othpt/	U.S. EPA 2016fh	2016
Canada Anthropogenic	othar/othpt	U.S. EPA 2016fh	2015

3.6.1 LADCO 2016 Emissions Summary

The tables in this section summarize the emissions used in the LADCO 2016 CAMx simulation. Figure 8 and Figure 9 are tile plots of the 12-km gridded, daily total NOx and VOC emissions, respectively, for a summer weekday (June 7, 2016). The NOx plot illustrates that the highest emissions occur in proximity urban areas and roadways. The VOC plot shows that biogenic sources dominate VOCs in the southern U.S and along the coasts. Table 5 shows the 2016 O_3 season (May-September) weekday NOx and volatile organic compound (VOC) emissions totals by LADCO member state. Table 6 and Table 7 include inventory sector level O_3 season weekday NOx and VOC emissions by state for 2016.

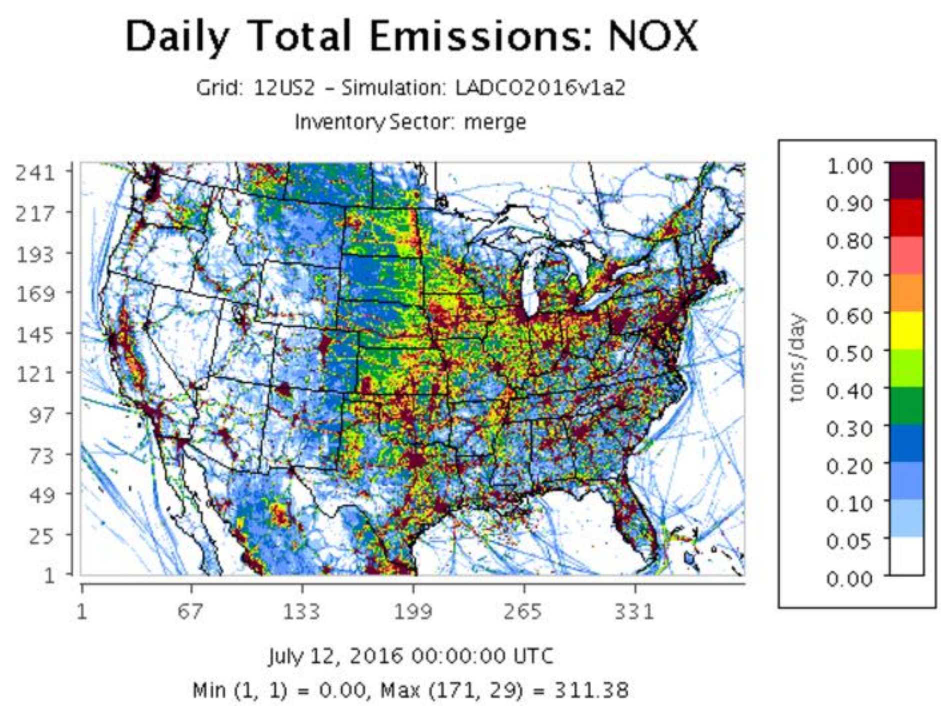

Figure 8. Daily total gridded 2016 NOx emissions for a summer weekday (tons/day)

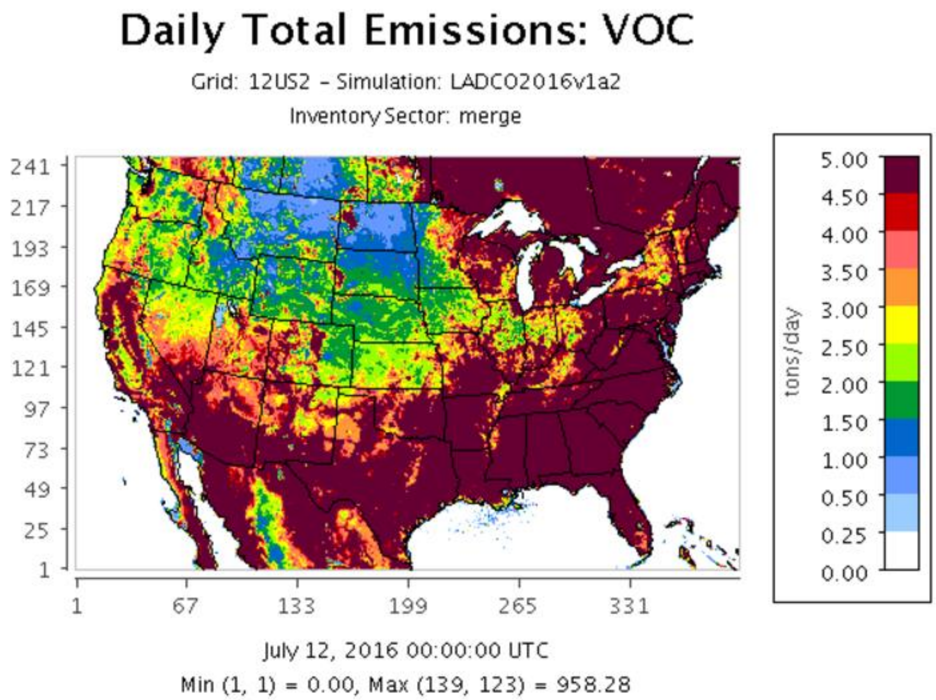

Figure 9. Daily total gridded 2016 VOC emissions for a summer weekday (tons/day)

Table 5. 2016 ozone season weekday emissions total, excluding biogenics (tons/day)

State	NOX	VOC
Illinois	975.2	994.1
Indiana	894.6	660.3
Michigan	807.3	820.1
Minnesota	634.0	711.5
Ohio	956.5	928.4
Wisconsin	510.3	471.8

Table 6. 2016 ozone season weekday NOx emissions by inventory sector (tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0
Airports	14.3	2.5	5.4	4.9	3.2	1.7
Biogenic	167.0	89.6	61.7	126.9	74.0	71.8
C1/C2 Commercial Marine	17.1	4.5	15.8	3.0	7.6	5.7
C3 Commercial Marine	0.6	1.0	19.7	1.9	2.0	2.3
Nonpoint	83.8	18.9	66.4	41.6	62.0	36.5
Offroad Mobile	177.6	140.7	85.9	159.6	151.0	80.3
Nonpoint Oil & Gas	38.4	9.8	35.5	0.0	4.4	0.0
Onroad Mobile	347.9	312.6	291.6	193.9	368.1	241.2
Point Oil & Gas	23.5	14.1	29.3	7.8	31.2	1.5
Agricultural Fires	0.0	0.0	0.0	0.2	0.0	0.0
Electricity Generation	84.6	217.5	120.1	55.9	157.4	46.7
Wild and Prescribed Fires	1.0	0.4	1.6	6.5	0.8	1.9
Industrial Point	94.8	129.3	121.78	116.6	102.7	63.2
Rail	91.3	43.0	12.8	38.2	65.5	28.2
Residential Wood Combustion	0.3	0.3	1.5	3.9	0.5	1.1

Table 7. 2016 ozone season weekday VOC emissions by inventory sector (tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	28.3	27.4	12.6	51.5	25.5	20.4
Airports	4.7	1.2	2.1	1.8	1.6	0.9
Biogenic	2386.1	1575.4	3390.2	2894.4	2038.9	2770.9
C1/C2 Commercial Marine	0.9	0.2	0.5	0.1	0.4	0.2
C3 Commercial Marine	0.0	0.1	0.9	0.1	0.1	0.1
Nonpoint	372.5	261.8	351.7	183.9	417.8	177.5
Offroad Mobile	110.6	64.7	91.8	85.6	116.0	71.4
Nonpoint Oil & Gas	162.5	41.3	61.2		42.2	
Onroad Mobile	182.4	160.9	177.7	108.3	217.0	93.6
Point Oil & Gas	3.7	1.0	3.6	0.5	4.6	0.6
Agricultural Fires	0.0	0.0	0.0	0.5	0.0	0.0
Electricity Generation	3.4	3.6	2.64	1.3	3.1	2.0
Wild and Prescribed Fires	16.2	5.8	38.9	187.2	11.9	35.0
Industrial Point	101.0	85.7	62.5	50.8	78.9	59.1
Rail	4.3	2.0	0.6	1.8	3.1	1.3
Residential Wood Combustion	3.6	4.6	13.4	38.1	6.2	9.7

3.7 LADCO Modeling Platform Summary

Table 8 summarizes the LADCO 2016 air quality modeling platform elements.

Table 8. Listing of the LADCO 2016 air quality modeling platform components

Platform Element	Configuration	Reference	Data source
Meteorology Data	WRFv3.8	U.S. EPA, 2019b	U.S. EPA
Initial and Boundary Conditions	2016 Hemispheric CMAQ	U.S. EPA, 2019c	U.S. EPA
2016 Emissions Data	Inventory Collaborative 2016v1 ERTAC16.1 EGU Point and hourly CEMs		Inventory Collaborative and ERTAC
2020 Emissions Data	Inventory Collaborative 2016v1 ERTAC16.1 EGU Point		LADCO and ERTACT
Emissions Modeling Platform	U.S. EPA 2016fh_16j		U.S. EPA
Photochemical Grid Model	CAMxv7.0 beta4	Ramboll, 2018	LADCO

3.8 2016 CAMx Model Performance Evaluation

LADCO simulated 2016 air quality with CAMx using data derived from the U.S. EPA "2016fg" and "2016fh" modeling platforms. The only input data difference between the EPA and LADCO CAMx modeling was the emissions inventories. For their regional haze modeling platform, U.S. EPA used a modified version of the National Inventory Collaborative 2016beta inventory (U.S. EPA, 2019). LADCO used corrected the Inventory Collaborative 2016v1 inventories for this application, and prepared these emissions for CAMx using the U.S. EPA 2016fh_16j SMOKE modeling platform.

The differences between the LADCO and U.S. EPA 2016 modeling configurations are significant enough to warrant a new performance evaluation of LADCO's CAMx simulation. The CAMx model performance evaluation (MPE) presented here focuses on ozone at surface monitors in the states of Illinois (IL), Indiana (IN), and Wisconsin (WI) as these are the states for which this TSD will be used to support NAA SIPs. LADCO used the Atmospheric Model Evaluation Tool (AMET) version 1.3 to pair the model results and surface observations in space and time, generate bi-variate statistics of model performance, and to produce MPE plots.

LADCO evaluated the CAMx 2016 modeled O_3 concentrations against concurrent measured surface ambient O_3 concentrations using graphical displays of model performance and statistical model performance measures. The statistical measures were compared against established model performance goals and criteria following the procedures recommended in EPA's photochemical modeling guidance documents (e.g., EPA, 1991; 2018).

3.8.1 Available Aerometric Data for the Model Evaluation

The following routine air quality measurement data networks operating in in 2016 were used by LADCO in assessing CAMx O_3 model performance:

> **EPA AQS Surface Air Quality Data**: Data files containing hourly-averaged concentration measurements at a wide variety of state and EPA monitoring networks are available in the Air Quality System (AQS) database throughout the U.S. The AQS consists of many sites that tend to be mainly located in and near major cities. There are several types of networks within AQS that measure different species. The standard hourly AQS AIRS monitoring stations typically measure hourly O_3, NO_2, NO_X and CO concentration and there are thousands of sites across the U.S. Figure 10 shows the locations of AQS surface monitors in the region around Lake Michigan.

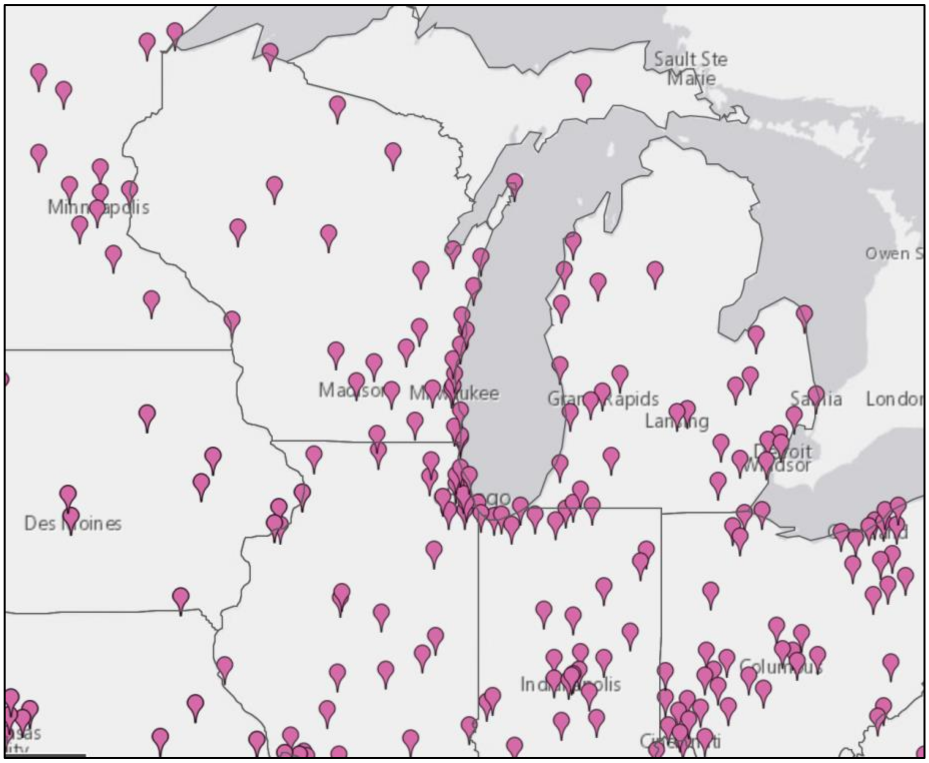

Figure 10. Locations of AQS monitors around Lake Michigan; source: U.S. EPA AirData

CASTNet Monitoring Network: The Clean Air Status and Trends Network (CASTNet) operates approximately 80 monitoring sites in mainly rural areas across the U.S. CASTNet sites typically collect hourly O_3, temperature, wind speed and direction, sigma theta, solar radiation, relative humidity, precipitation and surface wetness. CASTNet also collects weekly (Tuesday to Tuesday) samples of speciated $PM_{2.5}$ sulfate, nitrate, ammonium and other relevant ions and weekly gaseous SO_2 and nitric acid (HNO_3). Figure 11 displays the locations of the approximately 80 CASTNet sites across the U.S.

Figure 11. Locations of CASTNet monitoring sites; source: https://www.epa.gov/castnet

3.8.2 Model Performance Statistics, Goals and Criteria

As recommended by EPA (2018), LADCO used a 60 ppb observed O_3 cut-off threshold when calculating O_3 model performance statistics for this application.

Table 9. Ozone model performance goals

Fractional Bias (FB)	Fractional Error (FE)	Comment
≤±15%	≤35%	Ozone model performance goal that would be considered very good

It should be pointed out that these model performance goals are not used to assign passing or failing grades to model performance, but rather to help interpret the model performance and intercompare across locations, species, time periods and model applications. The model inputs to CAMx vary hourly, but tend to represent average conditions that do not account for unusual or extreme conditions. For example, an accident or large event could cause significant increases in congestion and motor vehicle emissions that are not accounted for in the average emissions inputs used in the model.

EPA compiled and interpreted the model performance from 69 air quality modeling studies in the peer-reviewed literature between 2006 and March 2012 and developed recommendations on what should be reported in a model performance evaluation (Simon, Baker and Phillips, 2012). Included in the most recent EPA guidance (U.S. EPA, 2018), they are useful and were used by LADCO in our model performance evaluation:

- Photochemical modeling MPE studies should at a minimum report the Mean Bias (MB) and Error (ME or RMSE), and Normalized Mean Bias (NMB) and Error (NME) and/or Fractional Bias (FB) and Error (FE). Both the MNB and FB are symmetric around zero with the FB bounded by -200% to +200%.
- The model evaluation statistics should be calculated for the highest temporal resolution available and for important regulatory averaging times (e.g., daily maximum 8-hour O_3).
- It is important to report processing steps in the model evaluation and how the predicted and observed data were paired and whether data are spatially/temporally averaged before the statistics are calculated.
- Predicted values should be taken from the grid cell that contains the monitoring site, although bilinear interpolation to the monitoring site point can be used for higher resolution modeling (< 12 km).
- Evaluation should be performed for subsets of the data including, high observed concentrations (e.g., O_3 > 60 ppb), by subregions and by season or month.
- Evaluation should include more than just O_3 and $PM_{2.5}$, such as SO_2, NO_2 and CO.
- Spatial displays should be used in the model evaluation to evaluate model predictions away from the monitoring sites. Time series of predicted and observed concentrations at a monitoring site should also be used.
- It is necessary to understand measurement artifacts in order to make meaningful interpretation of the model performance evaluation.

We incorporated the recommendations of U.S. EPA (2018) into the LADCO CAMx model performance evaluation. The LADCO evaluation products include qualitative and quantitative evaluation for maximum daily 1-hour and maximum daily 8-hour average (MDA8) O_3, including MDA8 with a 60 ppb threshold.

Table 10. Definition of model performance evaluation statistical measures used to evaluate the CTMs.

Statistical Measure	Mathematical Expression	Notes
Accuracy of paired peak (Ap)	$\frac{P - O_{peak}}{O_{peak}}$	Comparison of the peak observed value (O_{peak}) with the predicted value at same time and location
Normalized Mean Error (NME)	$\frac{\sum_{i=1}^{N} \lvert P_i - O_i \rvert}{\sum_{i=1}^{N} O_i}$	Reported as %
Root Mean Square Error (RMSE)	$\left[\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2\right]^{1/2}$	Reported as %
Fractional Gross Error (FE)	$\frac{2}{N}\sum_{i=1}^{N}\left\lvert\frac{P_i - O_i}{P_i + O_i}\right\rvert$	Reported as % and bounded by 0% to 200%
Mean Absolute Gross Error (MAGE)	$\frac{1}{N}\sum_{i=1}^{N}\lvert P_i - O_i \rvert$	Reported as concentration (e.g., µg/m^3)
Mean Normalized Gross Error (MNGE)	$\frac{1}{N}\sum_{i=1}^{N}\frac{\lvert P_i - O_i \rvert}{O_i}$	Reported as %
Mean Bias (MB)	$\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)$	Reported as concentration (e.g., µg/m^3)
Mean Normalized Bias (MNB)	$\frac{1}{N}\sum_{i=1}^{N}\frac{(P_i - O_i)}{O_i}$	Reported as %
Mean Fractionalized Bias (Fractional Bias, FB)	$\frac{2}{N}\sum_{i=1}^{N}\left(\frac{P_i - O_i}{P_i + O_i}\right)$	Reported as %, bounded by -200% to +200%
Normalized Mean Bias (NMB)	$\frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N} O_i}$	Reported as %
Bias Factor (BF)	$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_i}{O_i}\right)$	Reported as BF:1 or 1: BF or in fractional notation (BF/1 or 1/BF).

3.8.3 Subregional Evaluation of Model Performance

The evaluation of the LADCO 2016 CAMx 12-km simulation focuses on monthly and O_3 season model performance at monitors in IL, IN, and WI. We also examined summer season high O_3 episodes in different parts of the region to determine how well the model performs on O_3 exceedance days and locations.

4 2020 Air Quality Projections

LADCO based our 2020 O_3 air quality and NAA attainment forecasts on the CAMx modeling platform released by the U.S. EPA in September 2019 to support regional haze progress assessments. The U.S. EPA 2016fh_16j emissions modeling platform included an emissions projection to 2023. Given the absence of available emissions data for 2020 at the time that this application initiated in late 2019, LADCO used linear interpolation between 2016 and 2023 for most of the emissions sectors to estimate 2020 emissions. An exception was for the stationary point source electricity generating unit (EGU) sector, which used a 2020 forecast estimated by the ERTAC EGU 16.1 model (MARAMA, 2012).

For the CAMx modeling used to predict 2020 air quality, LADCO kept all of the CAMx inputs other than the emissions the same as the 2016 simulation.

4.1 2020 Emissions Data

LADCO estimated 2020 emissions for most of the anthropogenic inventory sectors by interpolating between the 2016 and 2023 Inventory Collaborative 2016v1 inventories. We used linear interpolation for the emissions because 2020 inventories were not readily available for all of the sectors at the time that this application initiated. While LADCO recognizes that emissions do not change linearly, given the relatively short period between the base and future years (4 years), LADCO considers that linear interpolation was justified for this application. LADCO also considers linear interpolation of the emissions to 2020 better than the alternatives of either holding the emissions constant at 2016 levels or using 2023 emissions estimates to simulate air quality in 2020.

LADCO applied two distinct linear interpolation techniques to estimate 2020 emissions. The first method was applied to gridded non-point, low-level emissions (e.g., area and mobile sources) sectors that had already been processed through SMOKE for 2016 and 2023. LADCO calculated the 2020 gridded emissions using Equation 1,

$$E_{2020,i,j,t,p} = 3/7 \times E_{2016,i,j,t,p} + 4/7 \times E_{2023,i,j,t,p} \quad \text{(Equation 1)}$$

where,
E = hourly, gridded emissions
i = column
j = row
t = hour
p = pollutant

As Equation 1 would not work for point sources because new units came online and other units shut down between the 2016 base and 2020 future years, LADCO applied Equation 2 to interpolate the non-EGU industrial point sources to 2020.

For each process and pollutant:

$$E_{2020,s,p} = (2020\text{-}2016)*(E_{2023,s.p} - E_{2016,s,p})/(2023\text{-}2016) + E_{2016,s,p} \quad \text{(Equation 2)}$$

where,
E = annual emissions
s = point source process
p = pollutant

LADCO developed work-arounds to Equation 2 for (1) units that began operations after 2016 and were included in the 2023 inventory; and (2) units that were in the 2016 inventory but were shut down in the 2023 inventory. After reviewing the sources included in the first list, LADCO established that U.S. EPA developed the 2023 inventories for the new units based on state 2017 NEI submittals. We found that most of the new units that were added to the 2023 inventory by U.S. EPA were already operating by 2020. Without being able to identify all of the unit level startup date information, LADCO made the assumption that all new units listed as operating in 2023 would also be operating in 2020. We calculated the 2020 emissions for these sources using Equation 2, which simplifies to $E_{2020,s.p} = 4/7 \times E_{2023,s.p}$ when we consider that these sources had zero emissions in 2016.

For the second list of sources, or those units that were in the 2016 inventory but not in the 2023 inventory, LADCO found that most of the shutdowns were scheduled before 2020. Several units that we reviewed had already shut down by late-2019 when we were developing the 2020 inventory for this modeling application. LADCO assumed that all units identified to be shut down by 2023 would also be shut down in the 2020 inventory that we developed for this modeling application. We validated this assumption using internet searches to confirm that the largest units in this second list of sources were in fact shut down by 2020.

LADCO replaced the EGU emissions in the U.S. EPA “2016fh” emissions modeling platform with 2020 EGU forecasts estimated with the ERTAC EGU Tool version 16.1 (MARAMA, 2012). ERTAC EGU 16.1 integrates state-reported information on EGU operations and forecasts as of December 2019. LADCO considers that the ERTAC EGU Tool provides more accurate estimates of the growth and control forecasts for EGUs in the Midwest and Northeast states than the U.S. EPA approach used in U.S. EPA’s “2016fh” modeling platform.

4.1.1 LADCO 2020 Electricity Generating Unit Emissions

The ERTAC EGU model for growth was developed around activity pattern matching algorithms designed to provide hourly EGU emissions data for air quality planning. The original goal of the model was to create low-cost software that air quality planning agencies could use for developing EGU emissions projections. States needed a transparent model that was numerically stable and did not produce dramatic changes to the emissions forecasts with small changes in inputs. A key feature of the model includes data transparency; all of the inputs to the model are publicly available. The code is also operationally transparent and includes extensive documentation, open source code, and a diverse user community to support new users of the software.

Operation of the model is straightforward given the complexity of the projection calculations and inputs. The model imports base year Continuous Emissions Monitoring (CEM) data from U.S. EPA and sorts the data from the peak to the lowest generation hour. It applies hour specific growth rates that include peak and off peak rates. The model then balances the system for all units and hours that exceed physical or regulatory limits. ERTAC EGU applies future year controls to the emissions estimates and tests for reserve capacity, generates quality assurance reports, and converts the outputs to SMOKE-ready modeling files.

ERTAC EGU has distinct advantages over other growth methodologies because it is capable of generating hourly future year estimates which are key to understanding O_3 episodes. The model does not shutdown or mothball existing units because economics algorithms suggest they are not economically viable. Additionally, alternate control scenarios are easy to simulate with the model. In recent years significant effort has been put into the model to help users to prevent the generation of new coal plants to fit demand. The model now allows portability of generation to different fuels like renewables and natural gas to prevent this.

Differences between the U.S. EPA and ERTAC EGU emissions forecasts arise from alternative forecast algorithms and from the data used to inform the model predictions. The U.S. EPA based the EGU emissions forecast in their “2016fh” modeling platform on comments from states and stakeholders received through April 2019. ERTAC EGU 16.1 used CEM data from 2016 and state-reported changes to EGUs received through December 2019. The ERTAC EGU 16.1 emissions used for this modeling application represent the best available information on EGU forecasts for the Midwest and Eastern U.S. available December 2019.

Figure 12 through Figure 17 show gridded daily total 2020 NOx and VOC emissions for a summer weekday (June 7). The spatial patterns seen in these figures match with the patterns in the 2016 emissions figures shown previously. The difference and ratio plots illustrate the locations where emissions are projected to change in 2020 relative to 2016. The NOx ratio plot (Figure 14) shows that the largest NOx emissions reductions will occur along roadways and in urban areas; emissions increases are projected in oil and gas development regions, in Mexico, and in Canadian offshore sources in the Great Lakes. The VOC ratio plot (Figure 17) illustrates small emissions reductions in urban areas and emissions increases in oil and gas development areas.

Table 11 shows the LADCO state total 2020 O_3 season (May 1 – September 30) weekday NOx and VOC emissions. Table 12 and Table 13 show the total 2020 O_3 season weekday NOx and VOC emissions for each LADCO state by emissions inventory sector. Table 14 and Table 15 compare 2020 and 2016 ozone season day NOx and VOC emissions, respectively, by inventory sector for each LADCO state. Negative numbers in these tables indicate emissions reductions in 2020 relative to 2016. Comparisons of the EGU and industrial point source emissions changes between 2016 and 2020 is confounded by the different methods used by the U.S EPA and ERTAC EGU projection models for distinguishing EGU from non-EGU industrial point sources. Some of the decreases in EGU emissions in 2020 are due to sources being reclassified from the EGU to the non-EGU industrial point sector by ERTAC EGU.

Overall both the NOx and VOC ozone season day emissions are projected to decrease in 2020 relative to 2016 in all of the LADCO states. The NOx reductions range from 10%-17%, driven primarily by reductions in mobile source emissions. The VOC reductions are more modest, at around 1% in each of the LADCO states and also driven by reductions in mobile sources. For the anthropogenic sectors only (i.e., excluding biogenics), the ozone season day VOC emissions reductions are closer to 5% in each of the LADCO states.

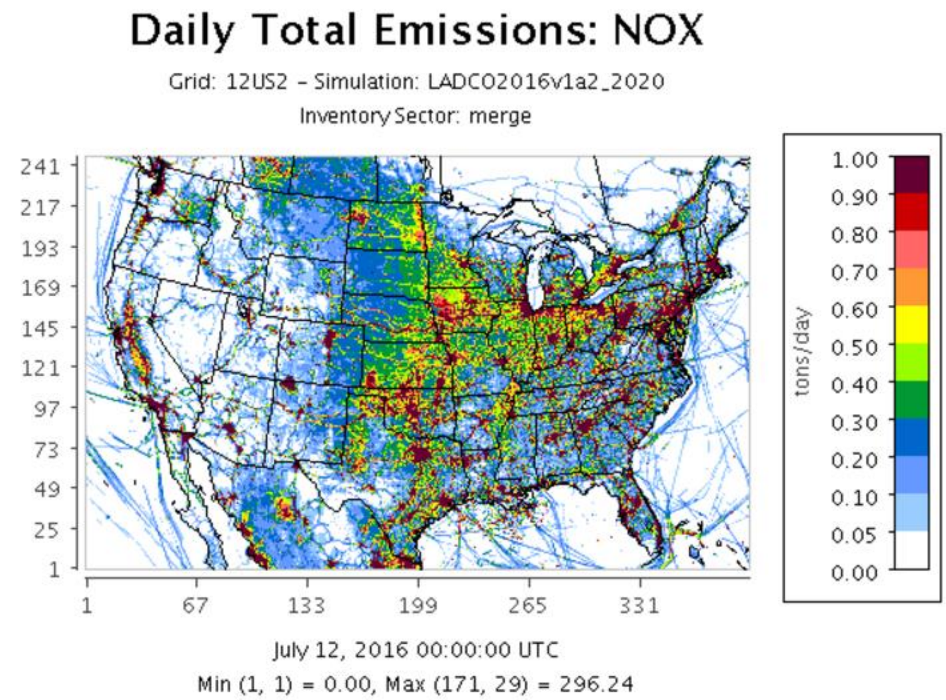

Figure 12. Daily total gridded 2020 NOx emissions for a summer weekday (tons/day)

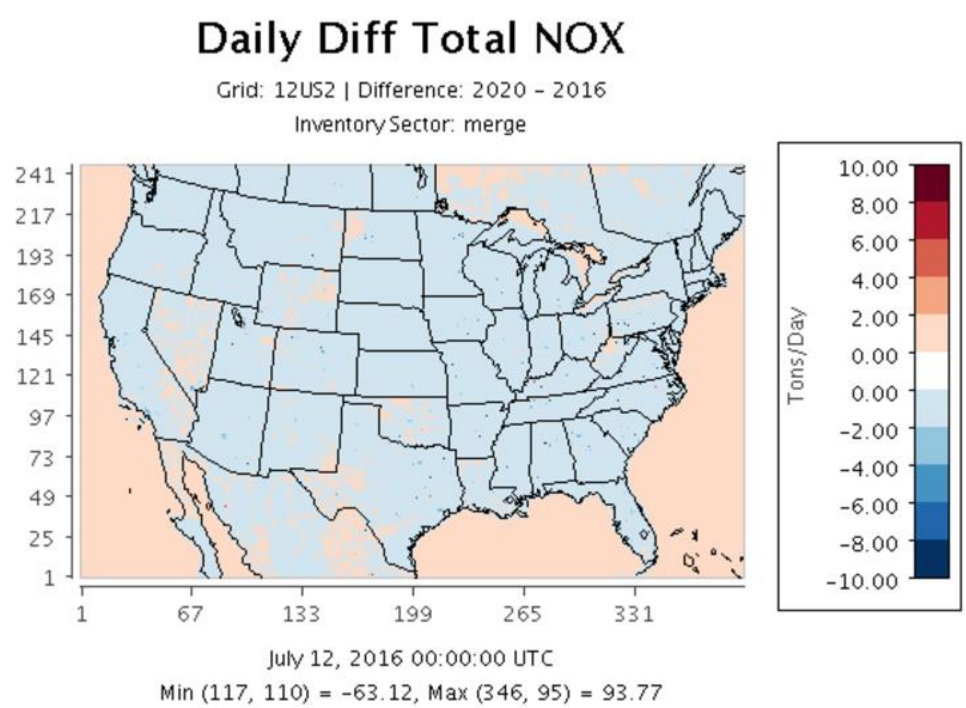

Figure 13. Difference (2020-2016) in daily total gridded NOx emissions for a summer weekday (tons/day)

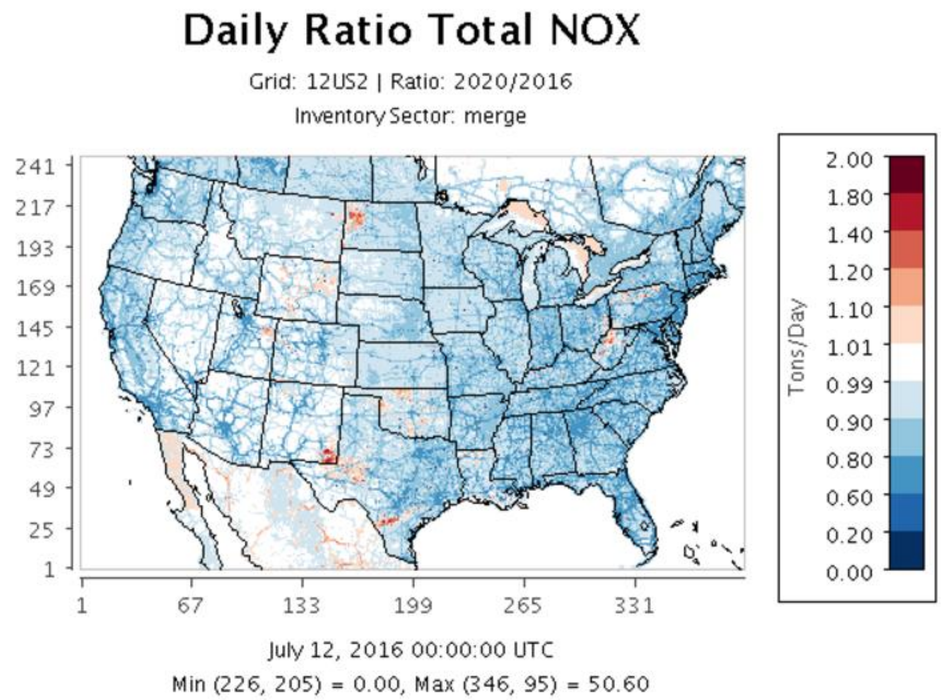

Figure 14. Ratio (2020/2016) of daily total gridded NOx emissions for a summer weekday (unitless)

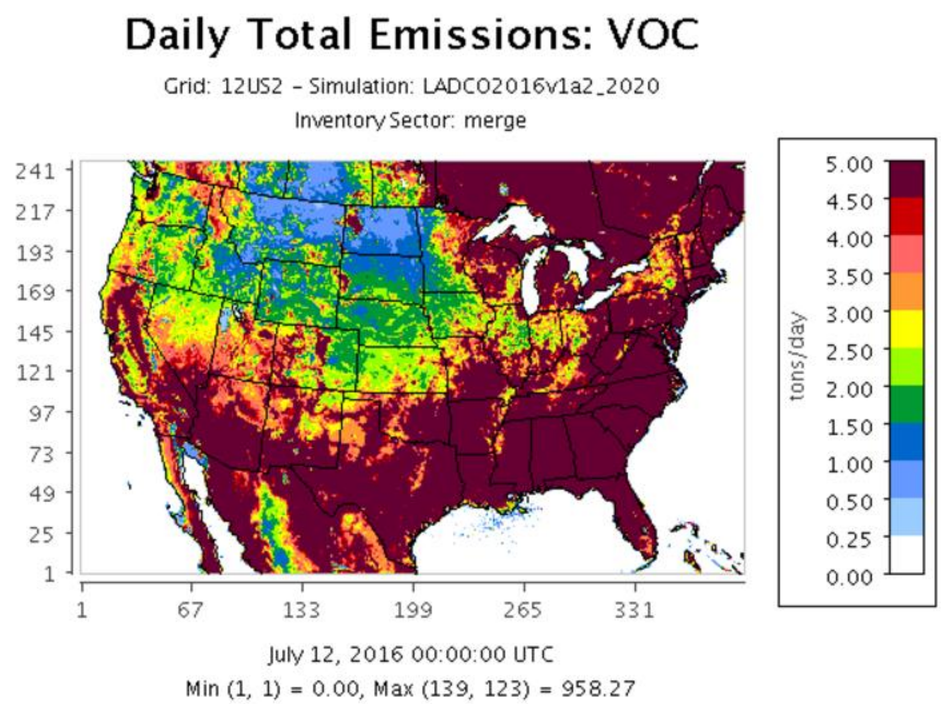

Figure 15. Daily total gridded 2020 VOC emissions for a summer weekday (tons/day)

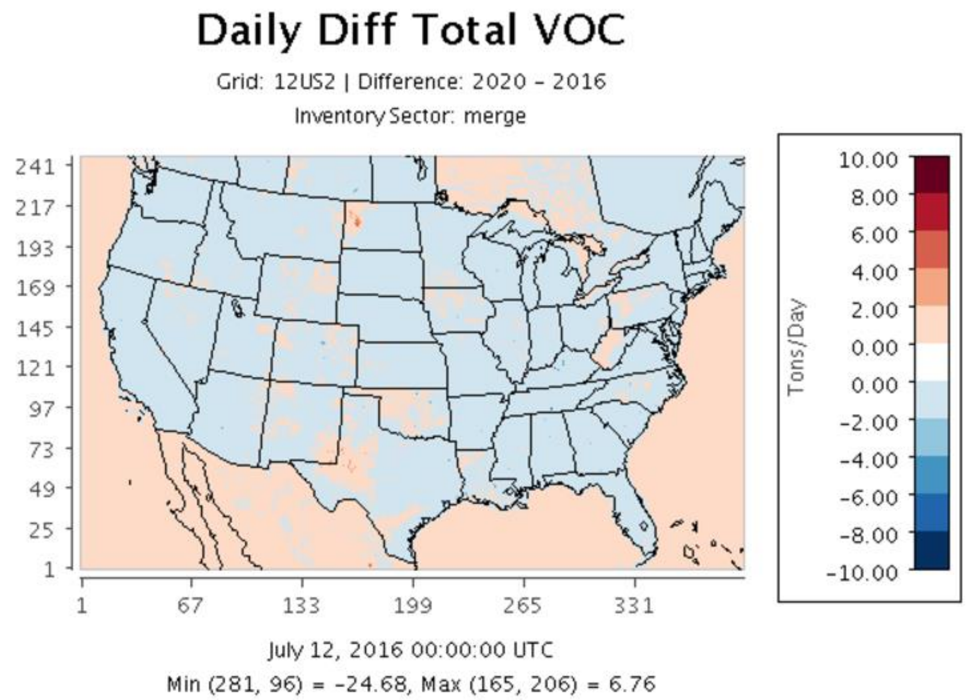

Figure 16. Difference (2020-2016) in daily total gridded VOC emissions for a summer weekday (tons/day)

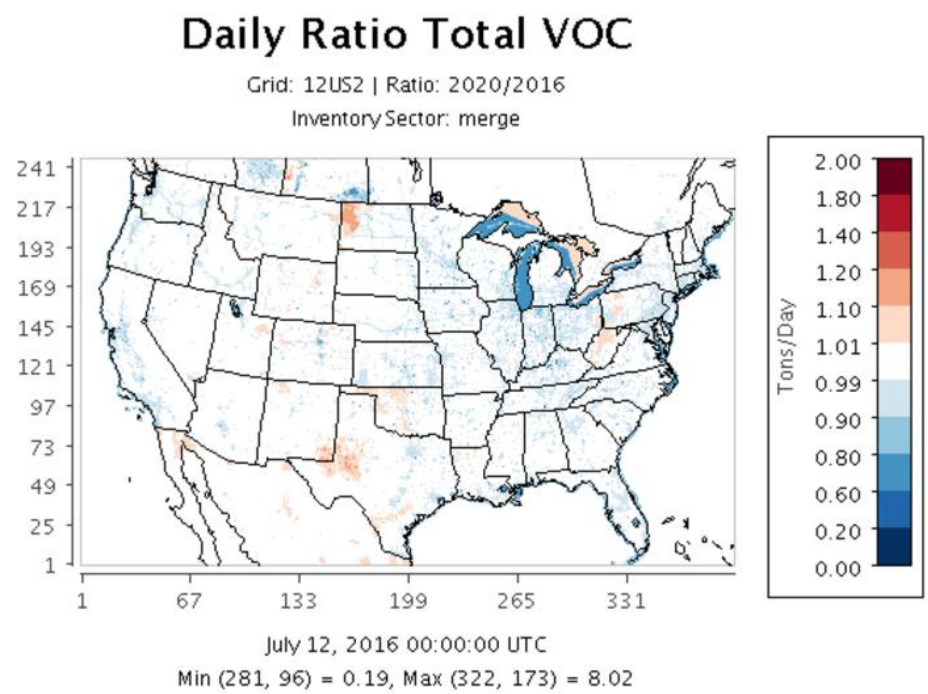

Figure 17. Ratio (2020/2016) of daily total gridded VOC emissions for a summer weekday (unitless)

Table 11. 2020 ozone season weekday emissions total by LADCO state, excluding biogenics (tons/day)

State	NOX	VOC
Illinois	860.8	954.5
Indiana	734.2	627.6
Michigan	693.0	776.5
Minnesota	525.8	685.9
Ohio	806.3	883.6
Wisconsin	429.2	447.9

Table 12. 2020 ozone season weekday NOx emissions by inventory sectors (tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0
Airports	15.9	2.8	6.0	5.3	3.3	1.9
Biogenic	167.0	89.6	61.7	126.9	74.0	71.8
C1/C2 Commercial Marine	14.8	3.9	13.8	2.5	6.6	5.0
C3 Commercial Marine	0.6	1.1	20.5	2.0	2.2	2.5
Nonpoint	82.0	18.8	65.2	41.4	61.0	36.1
Offroad Mobile	148.6	117.2	74.0	135.9	127.8	67.8
Nonpoint Oil & Gas	37.7	9.6	33.6	0.0	5.7	0.0
Onroad Mobile	271.7	244.1	224.2	149.8	283.3	183.6
Point Oil & Gas	25.1	17.5	30.1	7.6	34	1.6
Agricultural Fires	0.0	0.0	0.0	0.2	0.0	0.0
Electricity Generation	84.7	150.6	92.3	48.1	118.1	39.7
Wild and Prescribed Fires	1.0	0.4	1.6	6.5	0.8	1.9
Industrial Point	93.1	127.6	118.0	87	101.9	61.8
Rail	85.4	40.3	12.2	35.6	61.0	26.2
Residential Wood Combustion	0.3	0.3	1.5	3.9	0.5	1.1

Table 13. 2020 ozone season weekday VOC emissions by inventory sectors (tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	29.4	28.4	13.0	53.2	26.3	20.6
Airports	4.9	1.2	2.1	1.8	1.5	0.95

Biogenic	2386.1	1575.4	3390.2	2894.4	2038.8	2770.9
C1/C2 Commercial Marine	0.7	0.2	0.4	0.1	0.3	0.2
C3 Commercial Marine	0.0	0.1	1.0	0.1	0.2	0.2
Nonpoint	374.3	262.8	351.6	185.3	420.0	178.3
Offroad Mobile	101.9	60.2	82.4	76.0	105.9	63.2
Nonpoint Oil & Gas	161.0	41.3	59.9	0.0	44.3	0.0
Onroad Mobile	150.3	131.4	145.6	89.0	177.8	77.1
Point Oil & Gas	4.5	1.2	3.7	0.5	6.1	0.8
Agricultural Fires	0.0	0.0	0.0	0.5	0.0	0.0
Electricity Generation	3.5	3.9	2.4	1.1	2.4	2.1
Wild and Prescribe Fires	16.2	5.8	38.9	187.2	11.8	35.0
Industrial Point	100.3	84.7	61.3	50.8	77.9	58.4
Rail	3.9	1.8	0.6	1.6	2.8	1.2
Residential Wood Combustion	3.6	4.6	13.6	38.7	6.2	9.8

Table 14. Difference between base and future year ozone season weekday NOx emissions (2020-2016; tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Airports	1.6 (10%)	0.3 (10%)	0.6 (11%)	0.4 (7%)	0.1 (6%)	0.2 (12%)
Biogenic	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
C1/C2 Commercial Marine	-2.3 (-13%)	-0.6 (-13%)	-2.0 (-13%)	-0.5 (-17%)	-1.0 (-13%)	-0.7 (-12%)
C3 Commercial Marine	0.0 (0%)	0.1 (10%)	0.8 (4%)	0.1 (5%)	0.2 (10%)	0.2 (9%)
Nonpoint	-1.8 (-2%)	-0.1 (-1%)	-1.2 (-2%)	-0.2 (0%)	-1.0 (-2%)	-0.4 (-1%)
Offroad Mobile	-29.0 (-16%)	-23.5 (-17%)	-11.9 (-14%)	-23.7 (-15%)	-23.2 (-15%)	-12.5 (-16%)
Nonpoint Oil & Gas	-0.7 (-2%)	-0.2 (-2%)	-1.9 (-5%)	0.0 (0%)	1.3 (30%)	0.0 (0%)
Onroad Mobile	-76.2 (-22%)	-68.5 (-22%)	-67.4 (-23%)	-44.1 (-23%)	-84.8 (-23%)	-57.6 (-24%)
Point Oil & Gas	1.6 (7%)	3.4 (24%)	0.8 (3%)	-0.2 (-3%)	2.8 (9%)	0.1 (7%)
Agricultural Fires	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Electricity Generation	0.1 (-0%)	-66.9 (-30%)	-27.8 (-23%)	-7.8 (-13%)	-39.3 (-24%)	-7.0 (-15%)

Wild and Prescribe Fires	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Industrial Point	-1.7 (-2%)	-1.7 (-1%)	-3.78 (-3%)	-29.6 (-25%)	-0.8 (-1%)	-1.4 (-1%)
Rail	-5.9 (-6%)	-2.7 (-6%)	-0.6 (-5%)	-2.6 (-7%)	-4.5 (-7%)	-2.0 (-7%)
Residential Wood Combustion	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Total	-121.5 (-10%)	-164.3 (-17%)	-104.1 (-12%)	-89.4 (-12%)	-161.2 (-15%)	-88.7 (-15%)

Table 15. Difference between base and future year ozone season weekday VOC emissions comparison (2020-2016; tons/day)

Sector	Illinois	Indiana	Michigan	Minnesota	Ohio	Wisconsin
Agriculture	1.1 (4%)	1.0 (4%)	0.4 (3%)	1.7 (3%)	0.8 (3%)	0.2 (1%)
Airports	0.2 (4%)	0.0 (0%)	0.0 (2%)	0.0 (0%)	-0.1 (-3%)	0.0 (0%)
Biogenic	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	-0.1 (0%)	0.0 (0%)
C1/C2 Commercial Marine	-0.2 (-22%)	0.0 (0%)	-0.1 (-20%)	0.0 (0%)	-0.1 (-25%)	0.0 (0%)
C3 Commercial Marine	0.0 (0%)	0.0 (0%)	0.1 (11%)	0.0 (0%)	0.1 (100%)	0.1 (100%)
Nonpoint	1.8 (0%)	1.0 (0%)	-0.1 (0%)	1.4 (1%)	2.2 (1%)	0.8 (0%)
Offroad Mobile	-8.7 (-8%)	-4.5 (-7%)	-9.4 (-10%)	-9.6 (-11%)	-10.1 (-9%)	-8.2 (-11%)
Nonpoint Oil & Gas	-1.5 (-1%)	0.0 (0%)	-1.3 (-2%)	0.0 (0%)	2.1 (5%)	0.0 (0%)
Onroad Mobile	-32.1 (-18%)	-29.5 (-18%)	-32.1 (-18%)	-19.3 (-18%)	-39.2 (-18%)	-16.5 (-18%)
Point Oil & Gas	0.8 (22%)	0.2 (20%)	0.1 (3%)	0.0 (0%)	1.5 (33%)	0.2 (33%)
Agricultural Fires	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Electricity Generation	-.1 (-3%)	0.3 (8%)	-0.24 (-9%)	-0.2 (-15%)	-0.7 (-23%)	-0.1 (-5%)
Wild and Prescribe Fires	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	-0.1 (-1%)	0.0 (0%)
Industrial Point	-0.7 (-1%)	-1.0 (-1%)	-1.2 (-2%)	0.0 (0%)	-1.0 (-1%)	-0.7 (-1%)
Rail	-0.4 (-9%)	-0.2 (-10%)	0.0 (0%)	-0.2 (-11%)	-0.3 (-10%)	-0.1 (-8%)
Residential Wood Combustion	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
Total	-40.5 (-1%)	-33.0 (-1%)	-44.3 (-1%)	-26.4 (-1%)	-45.0 (-2%)	-24.4 (-1%)

4.2 Evaluation of the LADCO 2020 CAMx Simulation

As future year air quality forecasts cannot be compared to observations for evaluation, LADCO relied on our 2016 MPE results to establish validity in the modeling platform. In addition to the MPE for the base year CAMx simulation, the U.S. EPA reported full MPE results for the 2016 WRF modeling (US EPA, 2019b) and for the 2016 hemispheric CMAQ modeling (U.S. EPA, 2019c) used to drive the LADCO 2020 CAMx simulation.

5 Future Year Ozone Design Values

LADCO followed the U.S. EPA Modeling Guidance for Demonstrating Air Quality Goals for Ozone, PM2.5, and Regional Haze (US EPA, 2018) to calculate design values in 2020 (DV_{2020}) for monitors in IL, IN, and WI. As we used a base year of 2016, we estimated the base year design values using surface observations for the years 2014-2018 ($DV_{2014-2018}$). LADCO estimated the DV_{2020} with version 1.6- of the Software for Modeled Attainment Test Community Edition (SMAT-CE)[5]. SMAT-CE was configured to use the daily max average 8-hr (MDA8) O_3 concentration above 60 ppb in a 3x3 matrix around each monitor across for the 10 highest modeled days, per the U.S. EPA Guidance. If there are less than 10 days with MDA8 O3 greater than 60 ppb, SMAT-CE uses all days, as long as there are at least 5 days that meet the minimum threshold criteria.

SMAT-CE uses a four step process to estimate DVs_{2020}:

1. Calculate $DV_{2014-2018}$ for each monitor
    - The O_3 design value is a three-year average of the 4^{th} highest daily maximum 8 hour average O_3 ($MDA8_4$):

    $$DV_{2016} = (MDA8_{4,2014} + MDA8_{4,2015} + MDA8_{4,2016})/3$$

    - Weighted 5-year average of design values centered on the base model year (2016):

    $$DV_{2014\text{-}2018} = (DV_{2016} + DV_{2017} + DV_{2018})/3$$

2. Find highest base year modeled days surrounding each monitor
    - Find ten days with the highest base year modeled MDA8 from within a 3x3 matrix of grid cells surrounding each monitor
    - At least 5 days with modeled MDA8 >= 60 ppb are needed to retain the monitor for the future year DV calculation

3. Calculate relative response factor (RRF) for each monitor

[5] https://www.epa.gov/scram/photochemical-modeling-tools

- Calculate multi-day average MDA8 for the base and future years from the maximum paired in space values in the 3x3 matrix
- Calculate the RRF as the ratio of the multi-day average future to multi-day average base year MDA8:

$$RRF = MDA8_{2020,avg}/MDA8_{2016,avg}$$

4. Calculate DV_{2020} for each monitor

$$DV_{2020} = RRF * DV_{2014\text{-}2018}$$

LADCO used the DV_{2020} to identify nonattainment and maintenance sites in 2020 using the 5-year weighted average baseline design values (2014-2018) per U.S. EPA (2018). Under this methodology, sites with average DVs_{2020} that exceed the 2008 O_3 NAAQS (76 ppb or greater) and that are currently measuring nonattainment would be considered nonattainment receptors in projected year of 2020.

6 Results and Discussion

6.1 2016 CAMx Model Performance Evaluation Results

LADCO simulated the entire O_3 season (April 1 – October 31, 2016) with CAMx using the 2016 CAMx modeling platform described previously. Figure 18 is a spatial plot of the O_3 season average normalized mean bias (NMB) of daily maximum 8-hour average (MDA8) O_3 concentrations. Each colored symbol on the figure is an AQS or CASTNet monitoring location. Cool colors represent monitors at which the observed MDA O_3 concentrations were underestimated by the CAMx simulation; warm colors represent where CAMx overestimated the observations. Grey and lighter shades represent low bias, or acceptable model performance, relative to the model performance goals discussed in Section 3.8.2. Averaged across the entire O_3 season, there is a low negative bias (i.e., underprediction bias) in the CAMx MDA8 O_3 predictions for sites in the LADCO region. Overall, the model estimates O_3 concentrations in the southeast and mid-Atlantic areas well. CAMx underpredicted observed O_3 in inland California, the west, across the Great Lakes, and the northeastern U.S. The CAMx MDA O_3 overestimates are within the EPA (2018) model performance benchmarks.

The CAMx average monthly NMB for MDA8 O_3 shown in Figure 19 reveals a seasonal trend in the bias. Early in the O_3 season (April – June) CAMx underpredicted O_3 throughout the LADCO region. For many of the northern and near-shore monitors in the LADCO region the monthly averaged NMB values miss the model performance goal for O_3 (+/- 15%) in April and May. In the latter part of the season (July – October), CAMx overpredicted O_3 at most of the monitors in the region. Despite the overpredictions in the later months of the season, the absolute model biases are not as high as they are in April and May.

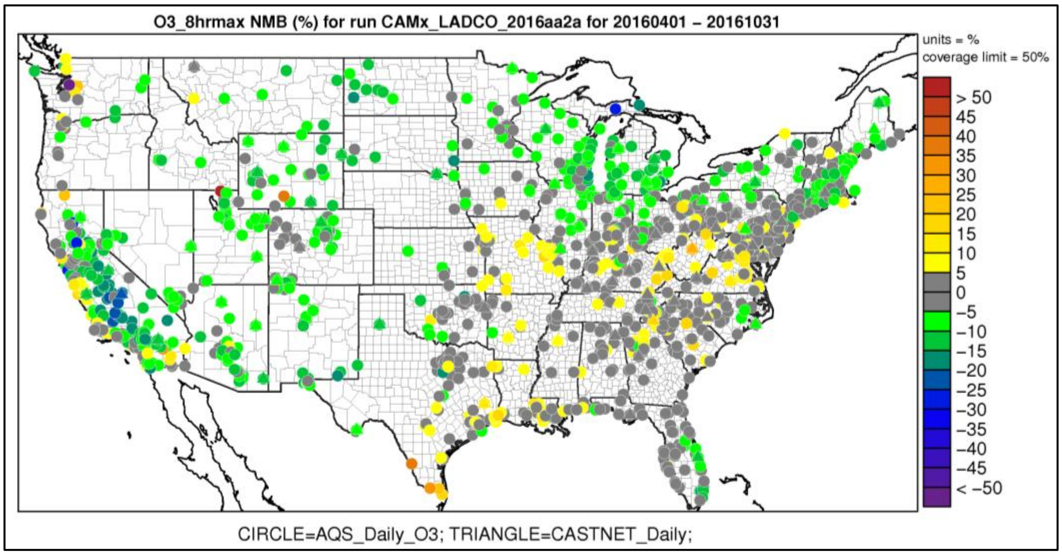

Figure 18. 2016 O_3 season MDA8 O_3 normalized mean bias spatial plot

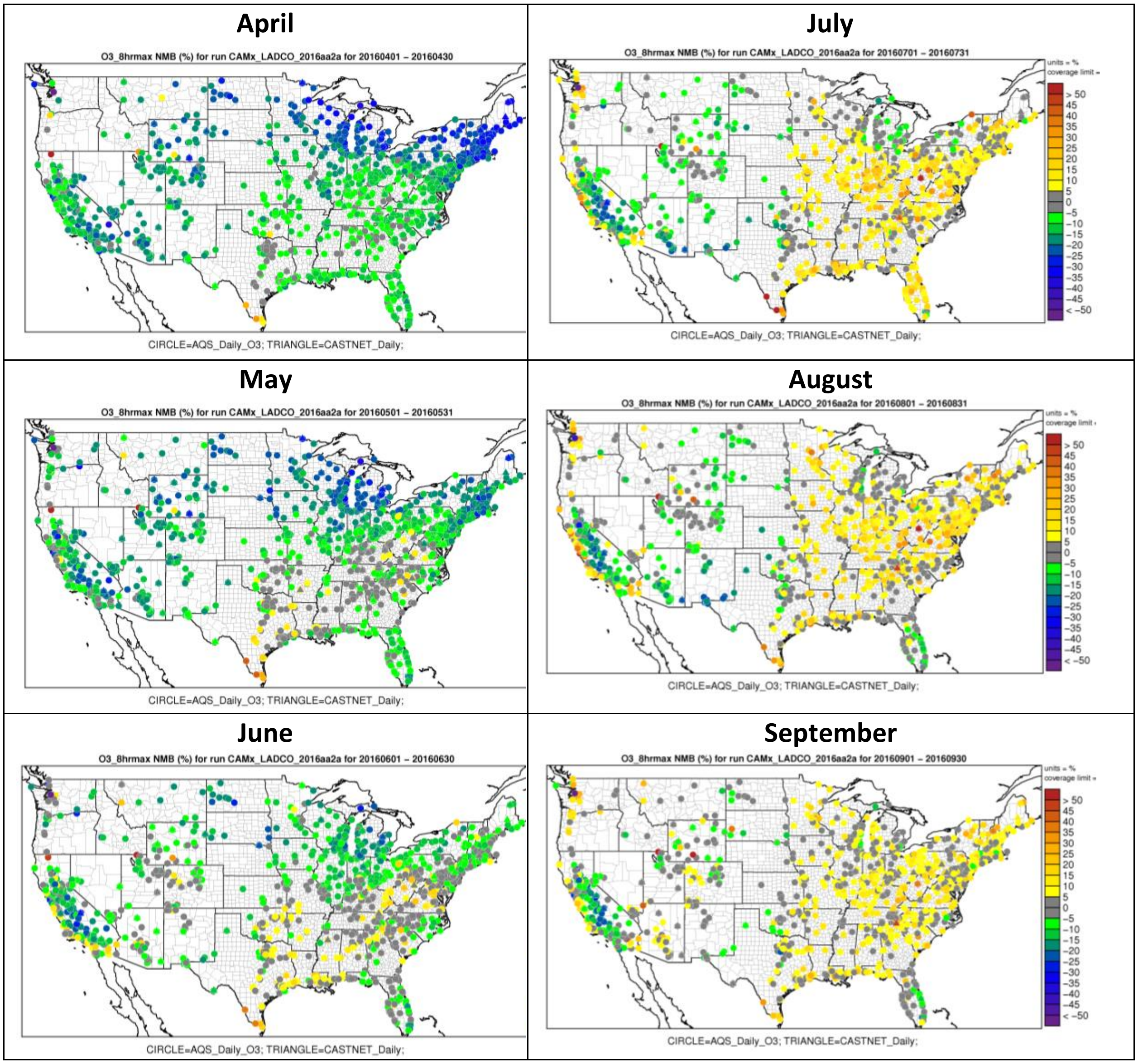

Figure 19. Monthly 2016 MDA8 O_3 normalized mean bias spatial plots

Table 16 and Table 17 show the daily maximum 8-hour average O_3 (MDA8) and daily maximum 1-hour O_3 performance statistics, respectively, for the LADCO 2016 CAMx simulation. Model biases and errors in these tables are O_3 season (April – October) averages across all Air Quality System (AQS) sites in the 12-km modeling domain, and across all AQS sites in each of the states of IL, IN, and WI. Each statistic is calculated for all observations and for only observations > 60 ppb. The latter is used to determine the performance of the model estimates for high observed concentrations.

As described above, these statistics quantify the LADCO CAMx simulation O_3 underpredictions and also show that the model performance degrades at higher observed concentrations. Despite the model performance deficits shown in these statistics, the O_3 model performance goal for bias (<=15%) is missed only for high concentrations at the WI AQS monitors. The performance goal for error (<=35%) is met across all of the locations and O_3 levels presented here.

Figure 20 through Figure 23 are scatter plots of O_3 season MDA8 O_3 concentrations for AQS and CASTNet monitoring sites in the 12-km domain, and for sites in the states of IL, IN, and WI. Plots are shown both with and without a 60 ppb observed O_3 cutoff. These plots indicate that the CAMx predictions are slightly better for observations of 60 ppb and greater MDA8 O_3 concentrations at the urban and suburban AQS sites than at the more rural CASTNet monitoring locations. The model errors at CASTnet monitors are smaller for MDA8 O_3 concentrations when the 60 ppb threshold wasn't applied.

Table 16. CAMx ozone season MDA8 O_3 performance at AQS monitoring locations

Region	FB (ppb)		FE (ppb)		NMB (%)		NME (%)	
	All	> 60ppb	All	> 60ppb	All	> 60ppb	All	> 60ppb
12-km	-2.9	-13.5	15.4	15.6	-3.1	-12.1	14.5	14.3
IL	-5.5	-15.1	16.0	17.1	-5.0	-13.3	15.1	15.5
IN	-1.2	-12.0	15.5	14.0	-0.9	-10.7	14.8	12.8
WI	-9.5	-21.9	17.5	22.5	-9.5	-19.1	16.7	19.7

Table 17. CAMx ozone season daily maximum 1-hour O_3 performance at AQS monitoring locations

Region	FB (ppb)		FE (ppb)		NMB (%)		NME (%)	
	All	> 60ppb	All	> 60ppb	All	> 60ppb	All	> 60ppb
12-km	-4.7	-13.5	15.5	16.3	-4.8	-12.2	14.6	15.0
IL	-5.8	-13.0	15.4	16.1	-5.3	-11.5	14.6	14.7
IN	-1.9	-10.4	14.9	13.3	-1.6	-9.4	14.3	12.4
WI	-11.1	-22.3	17.9	23.5	-10.8	-19.6	17.0	20.7

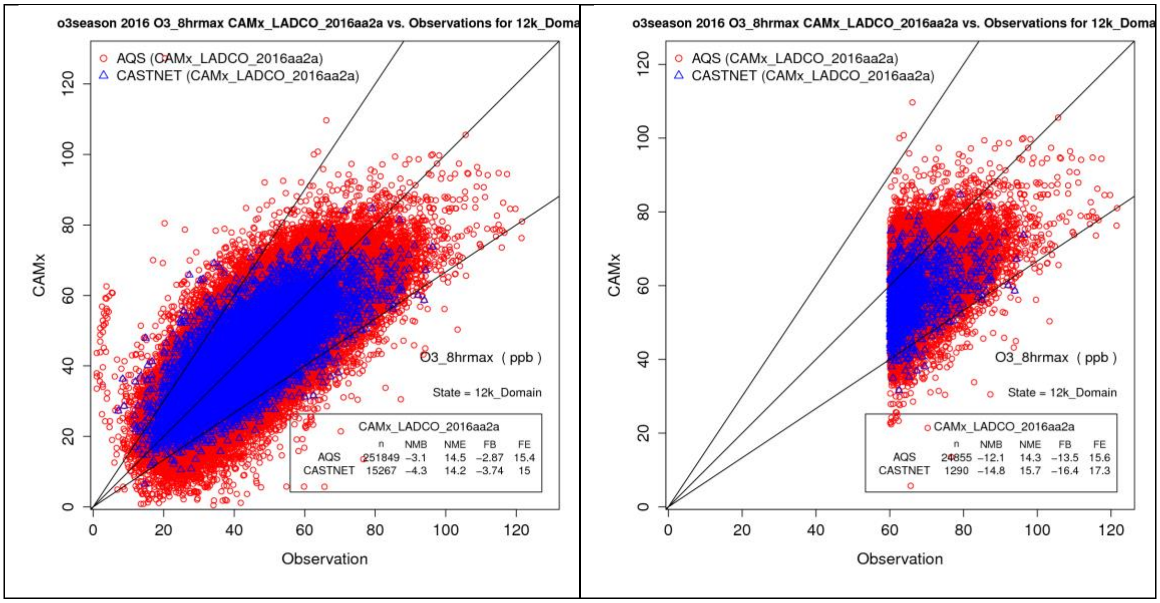

Figure 20. 2016 O_3 season MDA8 O_3 scatter plots for all sites in the 12-km modeling domain; all days (left), days > 60 ppb (right)

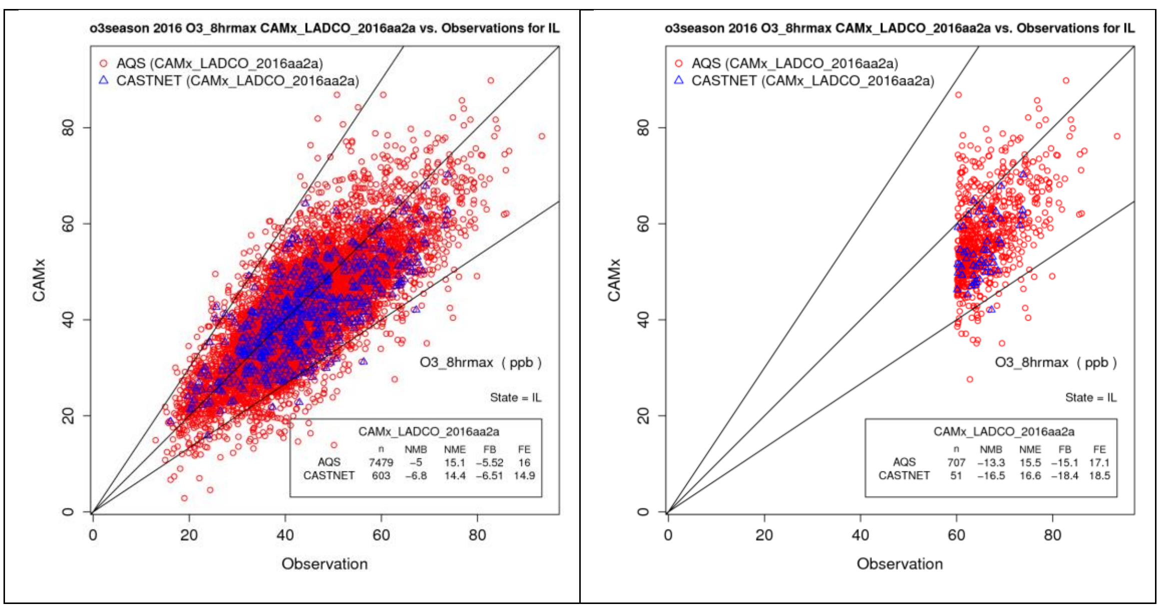

Figure 21. 2016 O_3 season MDA8 O_3 scatter plots for sites in IL; all days (left), days > 60 ppb (right)

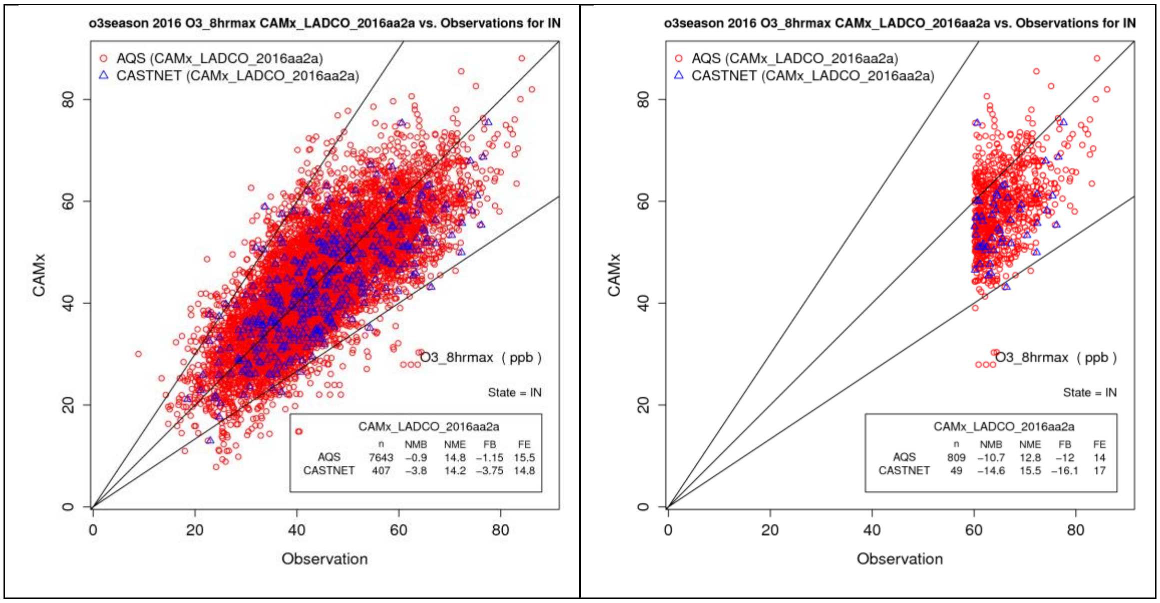

Figure 22. 2016 O_3 season MDA8 O_3 scatter plots for sites in IN; all days (left), days > 60 ppb (right)

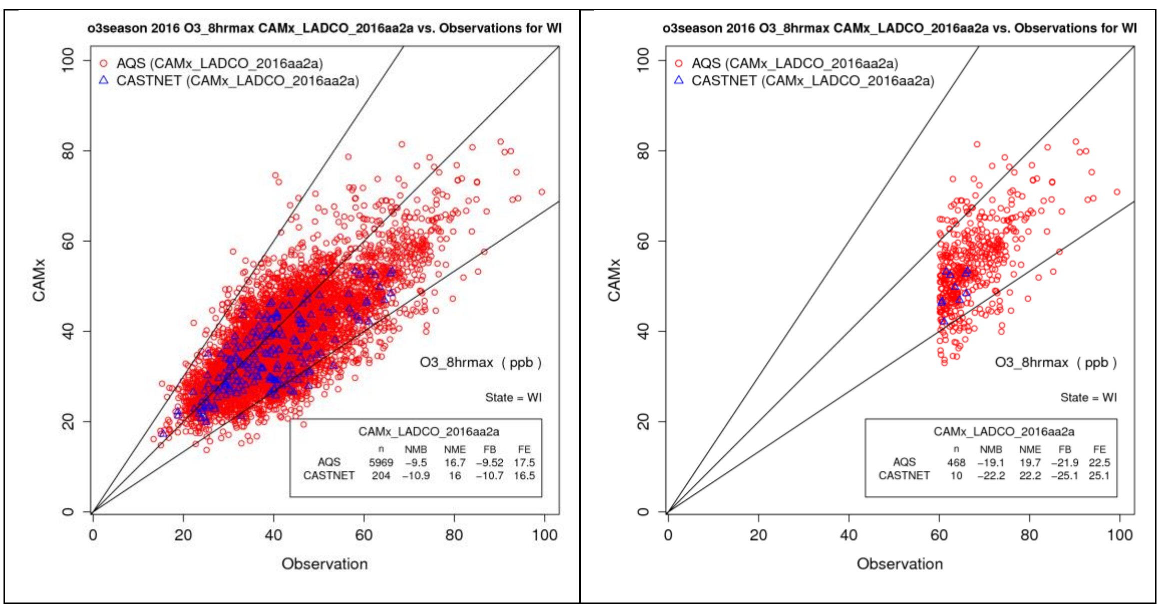

Figure 23. 2016 O_3 season MDA8 O_3 scatter plots for sites in WI; all days (left), days > 60 ppb (right)

Figure 24 through Figure 26 are monthly box and whisker plots of CAMx and observed MDA8 O_3 concentrations for AQS and CASTNet sites in IL, IN, and WI, respectively. The box and whisker plots show the observed and model median concentrations as symbols connected by lines (blue for CAMx and black for observations), the 25th and 75th

percentile concentrations as the bottom and top of each box, and the 5th and 95th percentile concentrations as the bottom and top of each whisker. These plots further highlight the underpredictions during April – June, as seen by the lower median values for CAMx relative to the observations across in all three states during this period. The skill of CAMx to simulate the distribution of observed O_3 concentrations incrementally improves in July – October as seen by the closer correspondence of the median, 75th and 95th percentile observed and predicted concentrations for most of the months in the three states. In general, CAMx has an underprediction in the fourth quantile (i.e., highest concentration) end of the observed O_3 distribution.

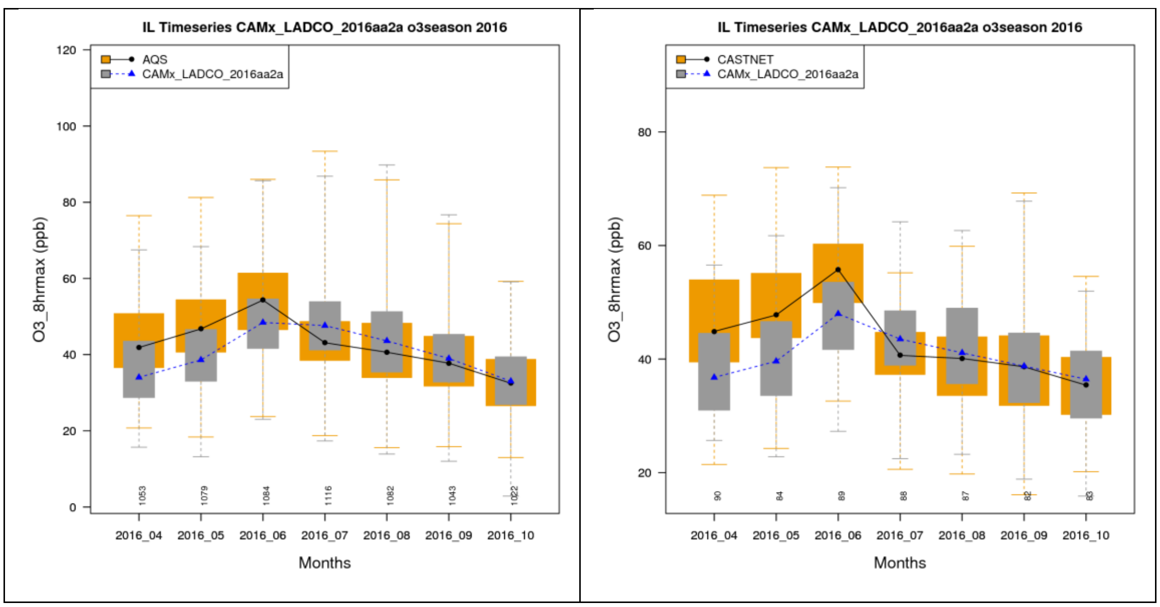

Figure 24. 2016 monthly MDA8 O_3 box and whisker plots comparing CAMx with AQS (left) and CASTNet (right) monitors for sites in IL.

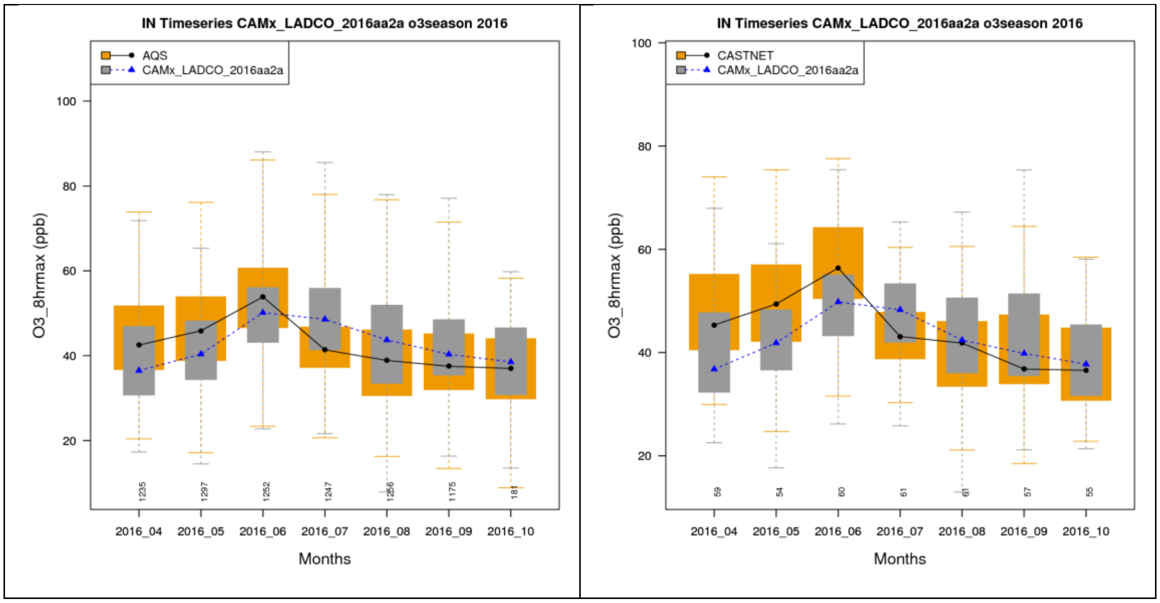

Figure 25. 2016 monthly MDA8 O_3 box and whisker plots comparing CAMx with AQS (left) and CASTNet (right) monitors for sites in IN.

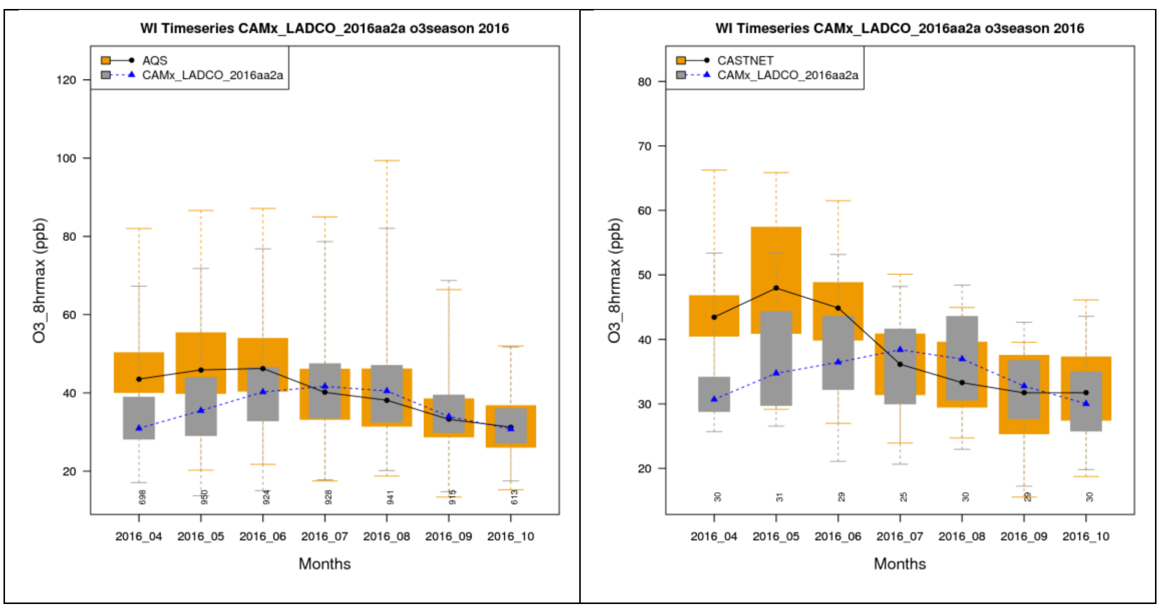

Figure 26. 2016 monthly MDA8 O_3 box and whisker plots comparing CAMx with AQS (left) and CASTNet (right) monitors for sites in WI.

Figure 27 through Figure 29 are monthly concentration-bias plots for MDA8 O_3 in IL, IN, and WI, respectively. These plots superimpose lines of the monthly average MDA8 O_3 CAMx predictions and AQS observed concentrations (right axis) on a bar plot of the monthly average NMB (left axis). The green and red horizontal lines on these plots are the +/- 15% and 35% bias and error goals and criteria for O_3 modeling. The value that these plots provide is a clear image of the switch in the bias signal from negative to positive biases moving from June to July in all three states, and also a reduction in the absolute bias of the model in the later months of the season, particularly for the AQS sites in IL and WI.

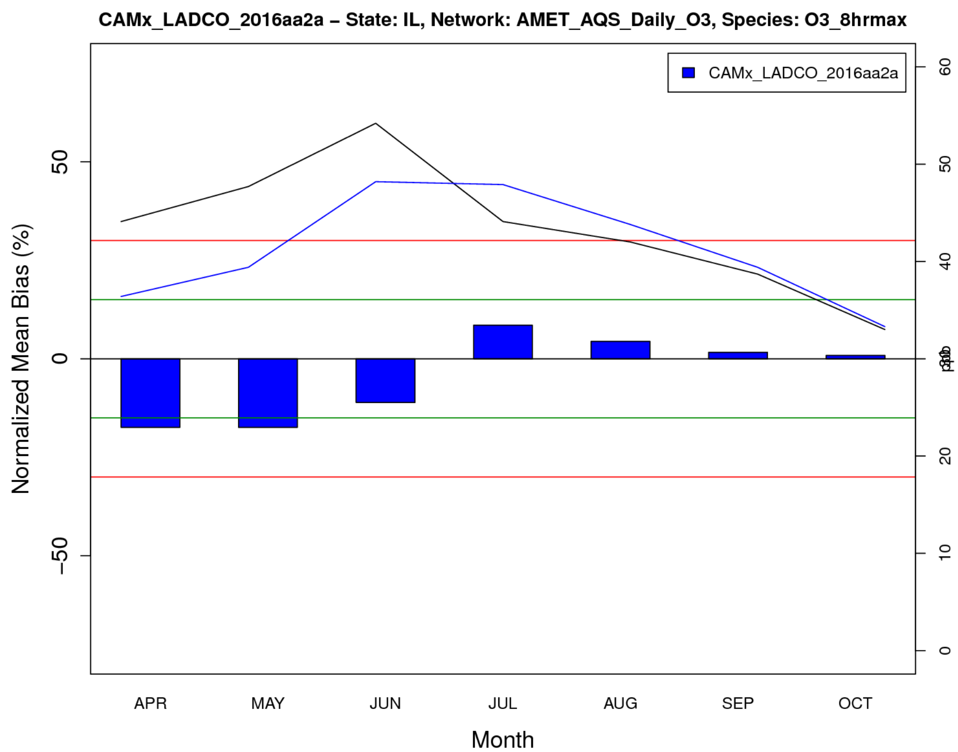

Figure 27. MDA8 O_3 monthly concentration-bias plot for IL AQS sites. Bars plot the average monthly normalized mean bias (left axis), lines are observed (black) and modeled (blue) monthly mean MDA8 O3 concentrations (right axis).

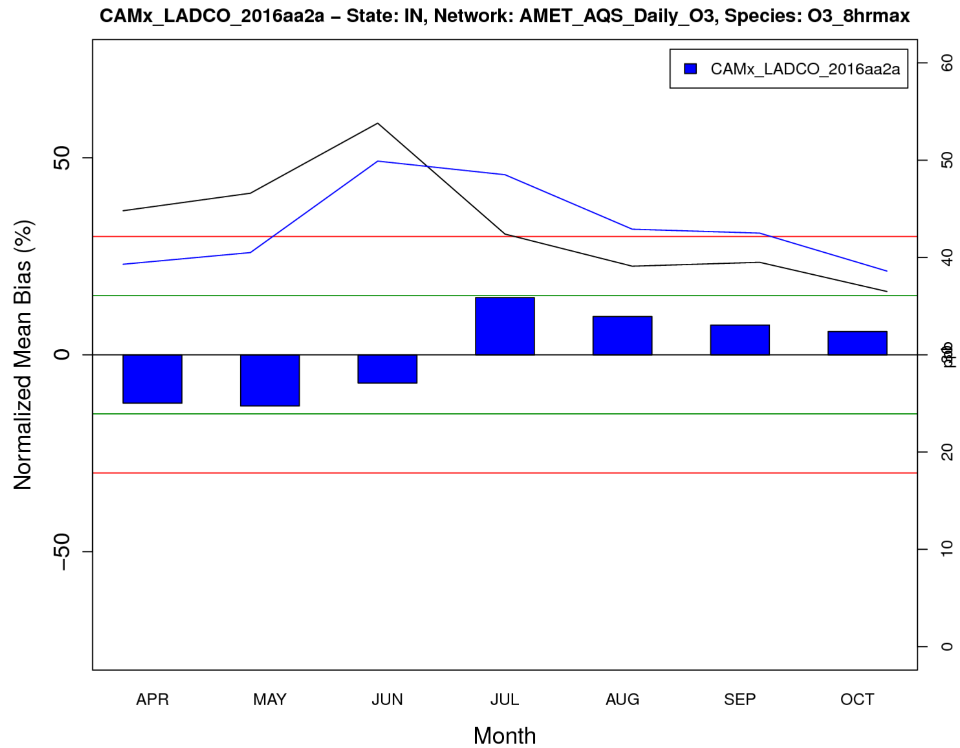

Figure 28. MDA8 O_3 monthly concentration-bias plot for IN AQS sites. Bars plot the average monthly normalized mean bias (left axis), lines are observed (black) and modeled (blue) monthly mean MDA8 O3 concentrations (right axis).

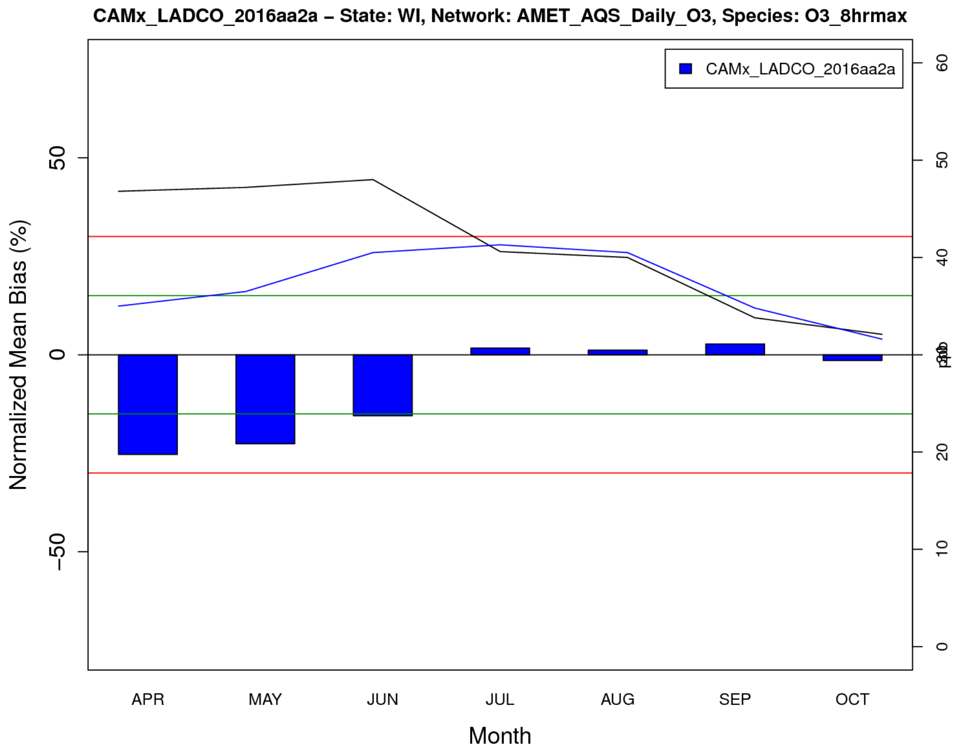

Figure 29. MDA8 O_3 monthly concentration-bias plot for WI AQS sites. Bars plot the average monthly normalized mean bias (left axis), lines are observed (black) and modeled (blue) monthly mean MDA8 O3 concentrations (right axis).

The Appendix of this document includes MDA8 O_3 time series plots for key sites in the Chicago O_3 NAAs. These plots compare the observed MDA8 O_3 concentrations to the CAMx 2016 and 2020 MDA8 O_3 predictions. Each plot shows the concentration comparisons in a top panel and a time series of the model bias in a bottom panel. The green line on the plots is the 2008 O_3 NAAQS (75 ppb).

6.2 CAMx Model Performance Discussion

U.S. EPA (2019) reported model performance for the 2016 CAMx modeling platform upon which we based the LADCO 2016 modeling platform. The U.S. EPA evaluated the model by comparing CAMx-predicted MDA8 O_3 to observations at the U.S. EPA AQS and CASTNet networks. They performed statistical evaluations using modeled and observed data that were paired in space and time. U.S. EPA developed statistics across spatial and temporal scales and in aggregate across multiple sites by climate region.

The results provided by U.S. EPA from their operational model performance evaluation (MPE) of their 2016 simulation are very similar to the results of the LADCO MPE. U.S. EPA and LADCO both found that the 2016 CAMx modeling platform on average underpredicts April – June MDA8 O_3 and overpredicts July – October MDA8 O_3. The biases in the April – June period are more severe than in the later months. In July – October the mean bias is within +/- 5 ppb at many sites in the LADCO region.

Investigation of the diurnal variability at key monitors demonstrated that CAMx generally captured day to day fluctuations in observed MDA8 O_3 but missed the peaks on many of the highest observed days, particularly during April – June. Figure 36 through Figure 40 compare daily AQS observations of MDA8 O_3 to the LADCO 2016 and 2020 CAMx simulations at monitors in the Chicago NAA.

Despite persistent deficiencies in model performance on days when the observed MDA8 $O_3 \geqq$ 60 ppb, the statistics in

Table 18 shows that CAMx performance was still within acceptable model performance criteria at key controlling sites within the Chicago NAA.

Table 18. LADCO CAMx April – September 2016 MDA8 O_3 model performance statistics at key monitors where observations >= 60 ppb

Site_ID	County, ST	Mean Obs	Mean Mod	MB (ppb)	ME (ppb)
170314201	Cook, IL	68.9	59.0	-9.8	11.4
170317002	Cook, IL	68.2	60.9	-7.3	9.2
170971007	Lake, IL	68.2	61.3	-6.9	11.1
550590019	Kenosha, WI	70.3	54.3	-16.0	17.2
550590025	Kenosha, WI	68.0	58.6	-9.4	10.4

6.3 LADCO 2020 Air Quality Projections

LADCO modified the emissions in the U.S. EPA 2016fg CAMx modeling platform to create a LADCO 2016 modeling platform with a projection year to 2020 (see Section 4.1). The LADCO 2020 simulation forecasted air quality for the continental U.S. using the best available information for North American emissions, including EGU emissions forecasts from the ERTAC v16.1 model. Figure 30 and Figure 31 show the O_3 season (April through October) maximum of MDA8 O_3 for the LADCO 2016 and 2020 CAMx simulations, respectively on the CONUS12 modeling domain. Figure 32 shows the difference in O_3 season maximum (2020-2016) between the two simulations. Cool colors indicate that the 2020 simulation forecasted lower O_3 than the 2016 simulation; warm colors indicate higher O_3 in the 2020 forecast. The 2020 CAMx simulation predicted lower seasonal maximum O_3 concentrations across the majority of the modeling domain with the largest reductions occurring in the southeast U.S., east Texas, and the Central Valley of California. Figure 33 zooms into the Lake Michigan area of the difference plot to highlight the predicted changes in O_3 season maximum MDA8 O_3 concentrations. This figure shows that in 2020 CAMx predicts that the seasonal maximum MDA8 O_3 concentrations will decline along the western Lake Michigan shoreline in the range of 1-5 ppb. Note that the trends shown in these figures mask finer temporal resolution features (i.e., hourly and daily) that also exist between the base and future year simulations.

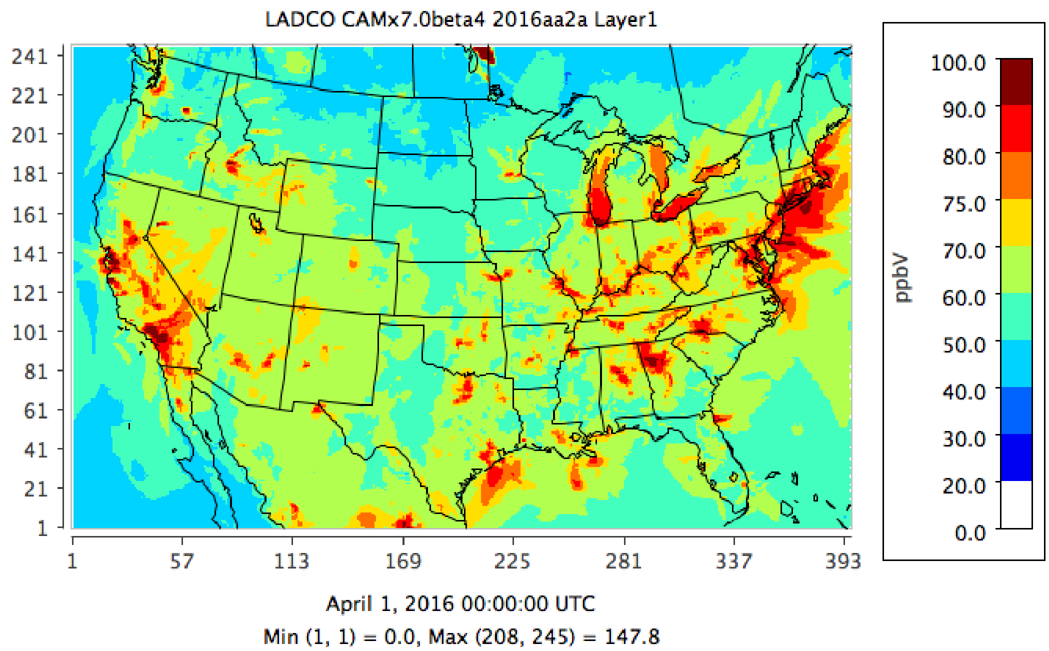

Figure 30. LADCO CAMx 2016aa2a O_3 season maximum MDA8 O_3 concentrations

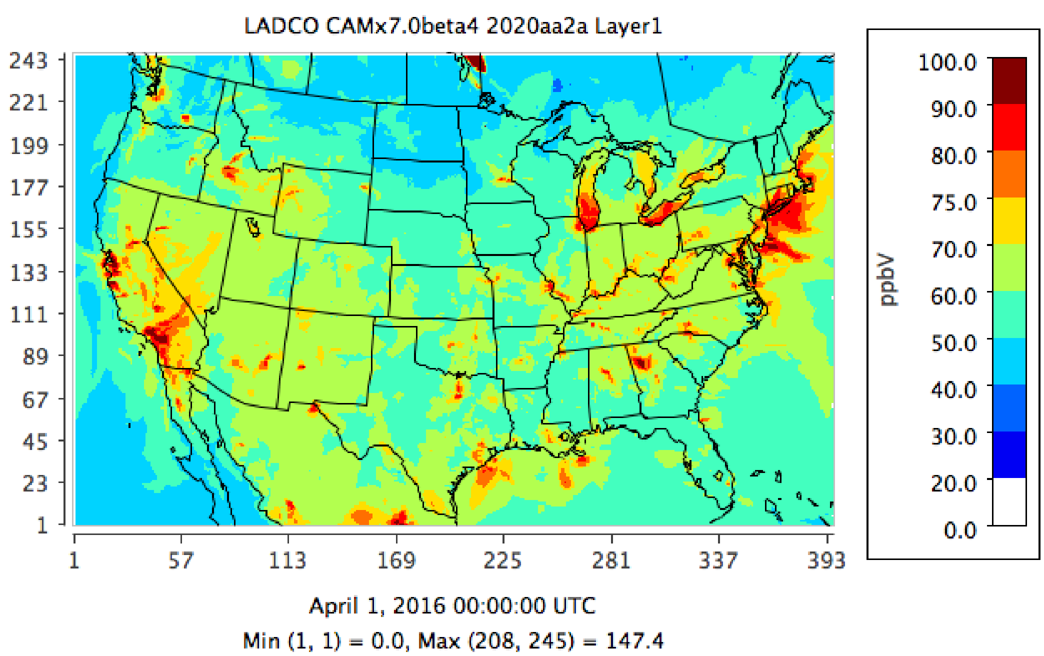

Figure 31. LADCO CAMx 2020aa2a O_3 season maximum MDA8 O_3 concentrations

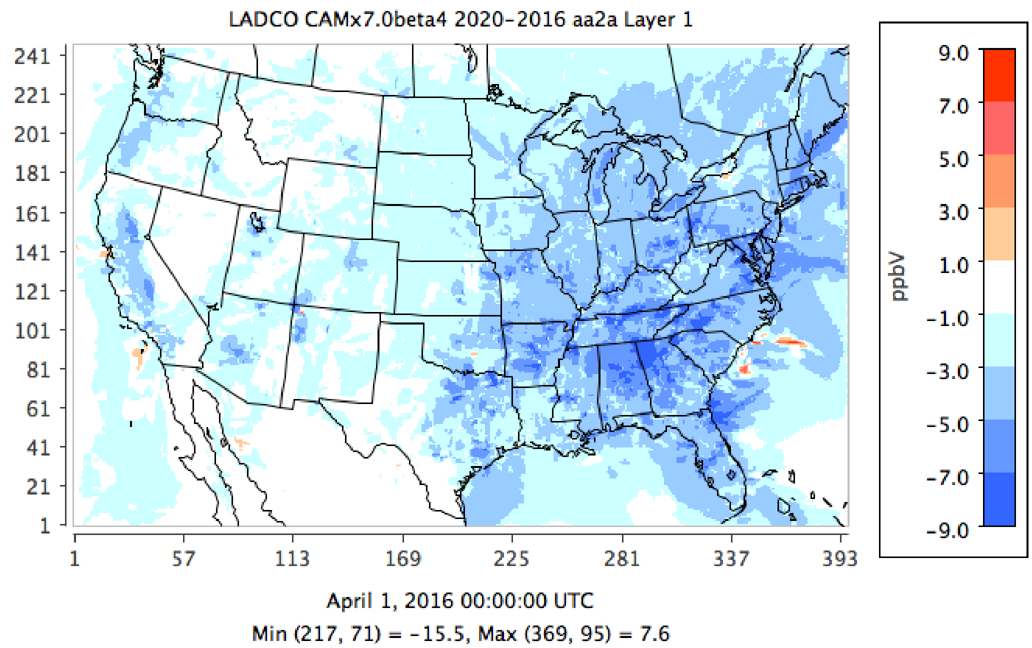

Figure 32. LADCO CAMx difference (2020-2016) in O_3 season maximum MDA8 O_3 concentrations

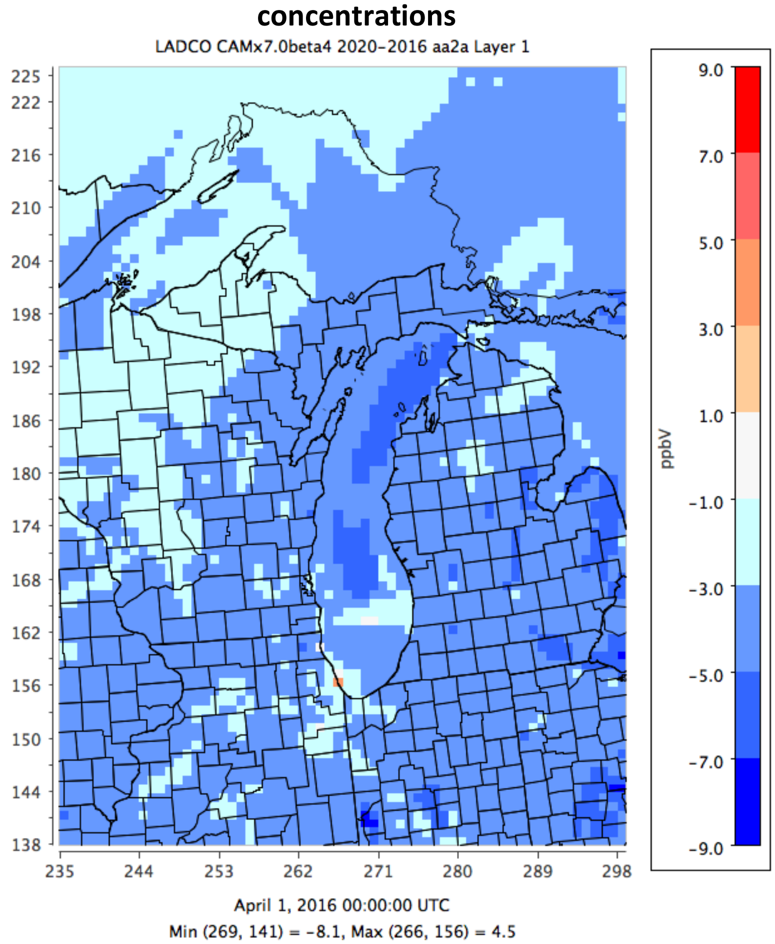

Figure 33. LADCO CAMx difference (2020-2016) in O_3 season maximum MDA8 O_3

concentrations; zoom to the Lake Michigan area

Figure 34 and Figure 35 show the O_3 DV_{2020} and RRFs from the LADCO 2020 simulation, respectively. LADCO generated these results with SMAT-CE using the standard U.S. EPA attainment test configuration (top 10 modeled days, 3x3 cell matrix around the monitor, including water cells). The LADCO O_3 DVs_{2020} presented here used observational data completeness criteria based on the 2015 O_3 NAAQS. The completeness criteria are tied to the level of the standard in cases in which the number of valid observations falls below a statutory threshold but when at least one of the valid observations is greater than the NAAQS (see 40 CFR Part 50 Appendix U, Section 3(d)). By using the 2015 O_3 NAAQS for determining completeness, LADCO includes more available data points in the DV calculations than if we had used the 2008 O_3 NAAQS completeness criteria because the lower standard is more inclusive of the available monitoring data (i.e., there are more MDA8 O_3 observations >= 70 ppb than there are observations >= 75 ppb).

The LADCO 2020 CAMx simulation predicts that no monitor in the region will have an average DV_{2020} that exceeds the 2008 O_3 NAAQS. The RRF plot indicates that most of the DV decreases in the Chicago NAA are in the range of 4-5%. The modest changes to the DVs in 2020 are due primarily to the short time period between the base and future years.

Table 19 presents the average and maximum DVs_{2020} for key monitors in the Chicago 2008 O_3 NAAQS NAA. As of September 30, 2020, the Chicago NAA has one monitor (Northbrook, IL) that is in violation of the 2008 O3 NAAQS.

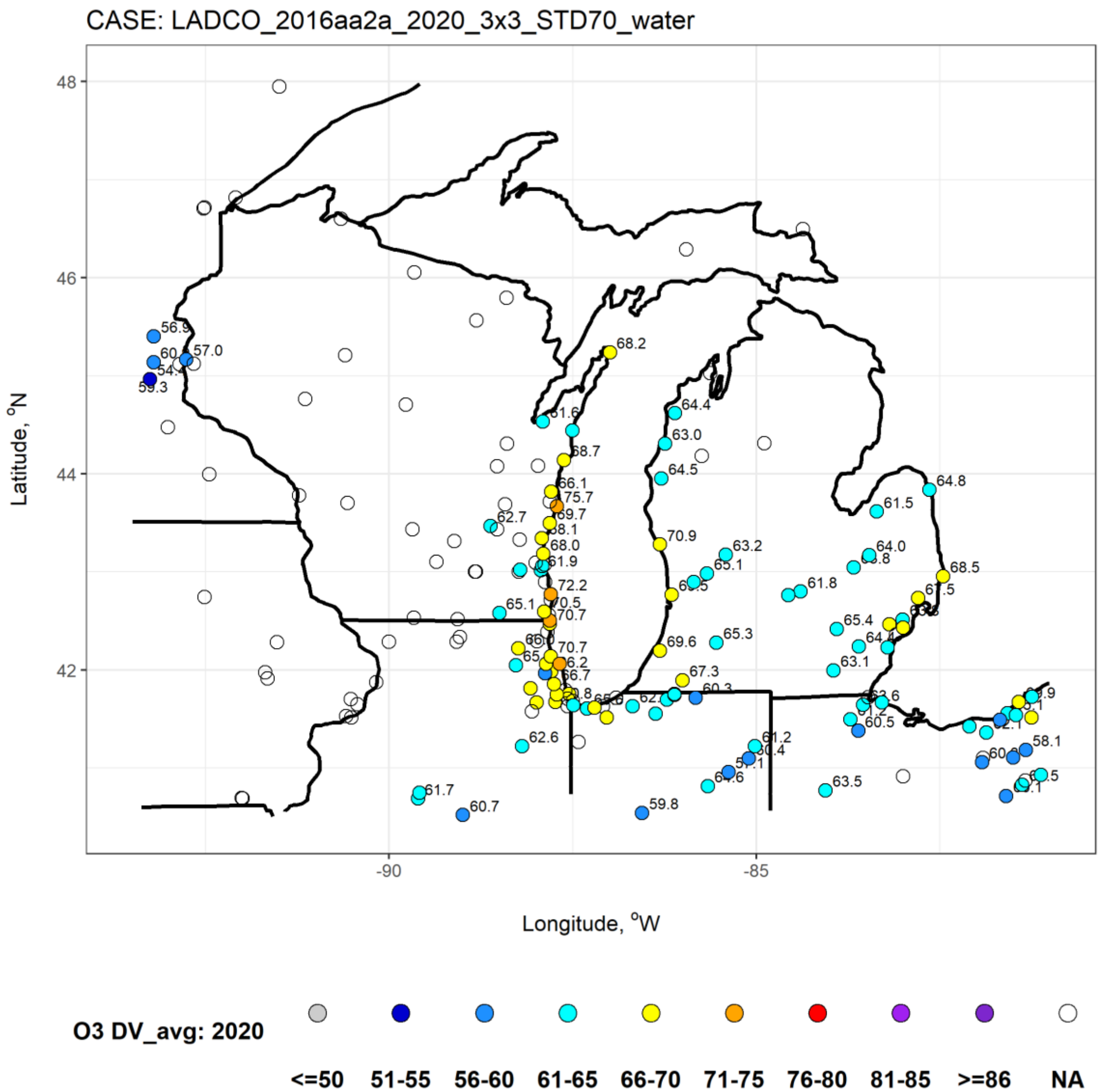

Figure 34. Future year O_3 design values calculated with WATER from the LADCO 2020 CAMx simulation.

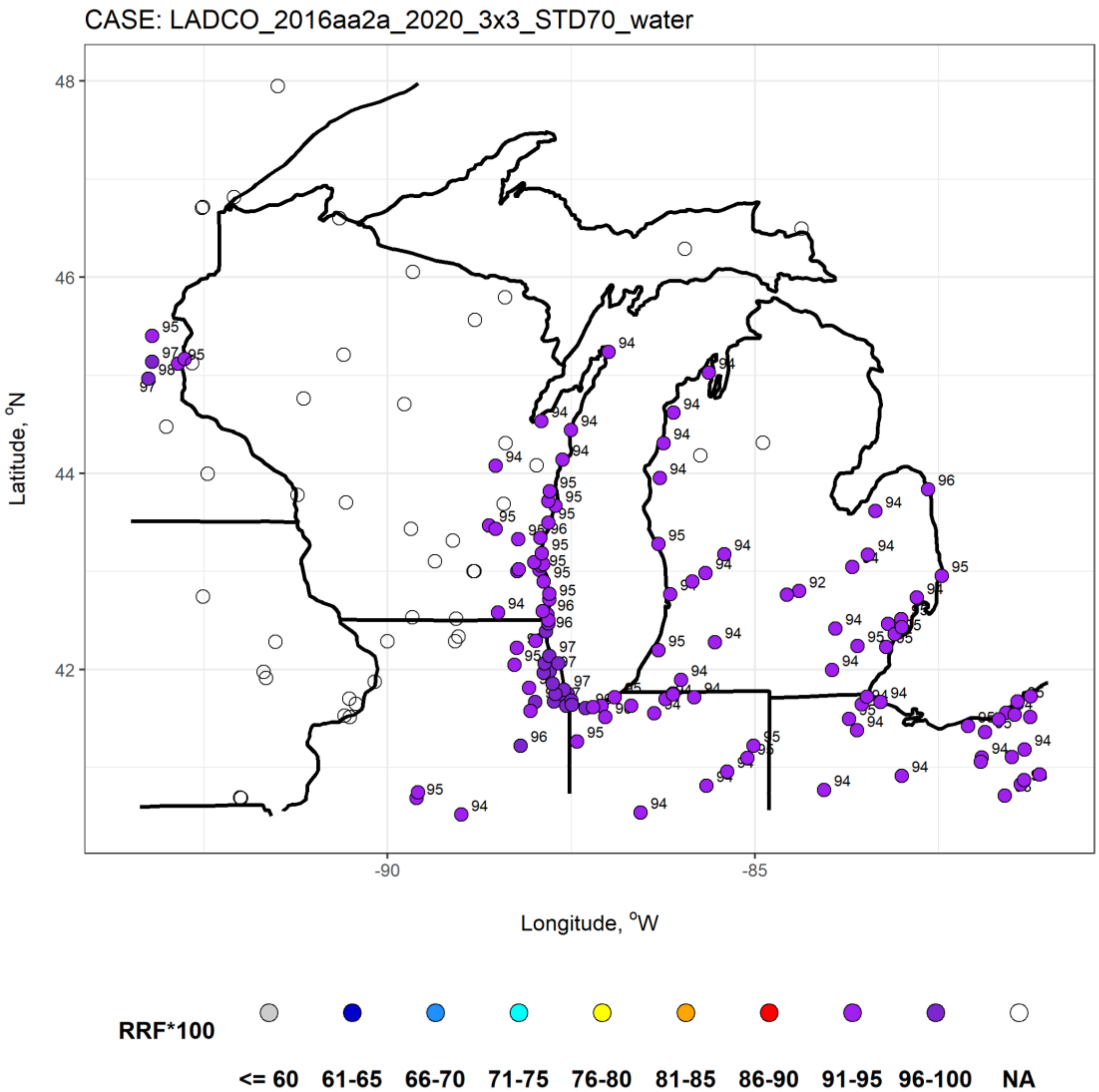

Figure 35. Future year O_3 relative response factors calculated with WATER from the LADCO 2020 CAMx simulation.

Table 19. LADCO 2020 O_3 design values with WATER at key monitors in the Chicago 2008 O_3 NAAQS NAAs

AQS ID	Monitor	ST	LADCO 2020 DVF	2014-2018 DVB	2017-2019 DV	2018-2020 DV*
170310001	Alsip	IL	70	73	75	75
170310032	South Water Filtration Plan	IL	70	72	73	74
170314201	Northbrook	IL	70	73	75	77
170317002	Evanston	IL	71	74	75	75
170971007	Zion	IL	70	73	71	72
550590019	Chiwaukee Prairie	WI	74	78	75	74
550590025	Kenosha Water Tower	WI	70	73	74	74

* Unofficial, as of September 30, 2020

6.3.1 Impacts of Water Cells on Design Values

Confidence in the ability of photochemical models to accurately estimate O_3 over water is a persistent concern with the use of the models for air quality planning. This concern recently prompted measurement campaigns in the Eastern U.S. to address the issue (see the 2017 Lake Michigan Ozone Study, Long Island Sound Tropospheric Ozone Study, and OWLETS). The meteorology and chemistry processes in model grid cells that are dominated by water (> 50% landuse area) are a challenge to simulate because the conventional technical formulations of the models were not optimized for water cells. Even with the introduction of new algorithms to simulate the dynamical and chemical features of water cells, a lack of over-water observations hinders our ability to verify the accuracy of the models in simulating these conditions.

In consideration that the models may not perform well in simulating water cells, EPA and others have presented alterative DVF calculation approaches that exclude water cells. Although not explicitly listed in Attachment A of the EPA's March 2018 memo on O_3 Transport Modeling as a flexibility to consider in developing a Good Neighbor SIP, the EPA used the exclusion of water cells in their own DVF calculations (US EPA, 2017; US

EPA, 2018b). Per EPA (2018, pg. 109), when appropriate there may be cases where certain cells along the periphery of the 3 x 3 array have different modeled responses than what would be expected at the monitor location at the center of array due to a specific local topographic or geographical feature (e.g., a large water body or a significant elevation change). A potential example of this situation would be an array where several cells are over water and where the meteorological conditions and relevant emissions sources differ substantially from the land-based monitor location. Again, in these types of cases and with appropriate justification, air agencies could consider removing the unrepresentative cells from the calculation.

Factoring in the impact of water cells on the DV calculation does not require additional CAMx simulations. It is implemented through a postprocessing sequence per U.S. EPA (2018b) in which model grid cells that are dominated by water (> 50% landuse area) are removed from the 3x3 matrix in the RRF and DVF calculation. One important modification to this process is to override the exclusion condition for cells that contain monitors; in other words, grid cells that contain monitors will be included in the 3x3 matrix regardless of the amount of water coverage in the cell.

Table 20 and Table 21 present the impacts of excluding water cells from the DV_{2020} calculations for the LADCO 2020 CAMx simulation. Excluding water cells in the attainment test calculation has different impacts on the DVs_{2020} for the lakeshore monitors in the LADCO region. The South Water Filtration Plant and Northbrook DVs_{2020} increase if water cells are excluded from the attainment test; the Evanston, Zion, Chiwaukee Prairie, and Kenosha Water Tower DVs_{2020} decrease.

Table 20. LADCO 2020 O_3 design values with NO WATER at key monitors in the Chicago 2008 O_3 NAAQS NAA

AQS ID	Monitor ID	ST	LADCO 2020		2014-2018		2017-2019 DV
			3x3 avrg	3x3 max	3x3 avrg	3x3 max	
170310001	Alsip	IL	70.8	74.6	73.0	77.0	75
170310032	South Water Filtration Plant	IL	70.7	73.4	72.3	75.0	73
170314201	Northbrook	IL	70.8	74.3	73.3	77.0	75
170317002	Evanston	IL	71.3	74.2	74.0	77.0	75
170971007	Zion	IL	70.1	71.4	73.7	75.0	71
550590019	Chiwaukee Prairie	WI	74.2	75.2	78.0	79.0	75
550590025	Kenosha Water Tower	WI	70.1	73.2	73.7	77.0	74

Table 21. Comparison of LADCO 2020 O_3 design values at key monitors in the Chicago 2008 O_3 NAAQS NAA with and without water cells included in the DV calculation

AQS ID	Monitor ID	ST	Water	No Water
			2020 DV	2020DV
170310001	Alsip	IL	70.8	70.8
170310032	South Water Filtration Plant	IL	70.1	70.7
170314201	Northbrook	IL	70.7	70.8
170317002	Evanston	IL	71.6	71.3
170971007	Zion	IL	70.7	70.1
550590019	Chiwaukee Prairie	WI	74.5	74.2
550590025	Kenosha Water Tower	WI	70.5	70.1

7 Conclusions and Significant Findings

LADCO presents in this TSD a regional air quality modeling platform for quantifying and evaluating future year O_3 concentrations pursuant to testing attainment of the 2008 O_3 NAAQS serious designations for the Chicago NAA. After establishing that the LADCO 2016-based modeling platform is an acceptable tool for simulating regional O_3 concentrations, we presented the results from projections of future O_3 concentrations and for calculating DVs_{2020}. A summary of the significant findings from the LADCO modeling follows.

- Finding 1: While the 2016 CAMx modeling platform has an underprediction bias for high O_3 concentrations, the platform skill is consistent with the U.S. EPA 2016 modeling platform used to support recent air quality analyses.
- Finding 2: The LADCO 2020 CAMx simulation predicts that no monitor in the LADCO region will have an average DV_{2020} that exceeds the 2008 O_3 NAAQS.
- Finding 3: Excluding water cells in the attainment test calculation results in both higher and lower DVs_{2020} for the lakeshore monitors in the LADCO region.

As with all regional air quality modeling applications, there are uncertainties in the model inputs and in the model formulation that produce biases in the results presented here. LADCO determined that as of the writing of this TSD the EPA 2016fh emissions modeling platform and the ERTAC EGU 16.1 emissions were the best available data for forecasting air quality in 2020.

References

Breiman, L., J. Friedman, R. Olshen, and C. Stone, Classification and Regression Trees, Pacific Grove, CA: Wadsworth (1984).

MARAMA, 2012. Software Technical Documentation For Software to Estimate Future Activity and Air Emissions from Electric Generating Units (EGUs). https://s3.amazonaws.com/marama.org/wp-content/uploads/2020/02/13091911/SystemDocumentation_EGUProjectionModel.pdf

Ramboll. 2020. User's Guide: Comprehensive Air Quality Model with Extensions version 7.00. Novato, CA. http://www.camx.com/files/camxusersguide_v7-00.pdf

US EPA. 2019. Technical Support Document for EPA's Updated Regional Haze Modeling. Research Triangle Park, NC. https://www.epa.gov/sites/production/files/2019-10/documents/updated_2028_regional_haze_modeling-tsd-2019_0.pdf

US EPA, 2019b. Meteorological Model Performance for Annual 2016 Simulation WRF v3.8. (EPA-454/R-19-010) United States Environmental Protection Agency.

U.S. EPA, 2019c. 2016 Hemispheric Modeling Platform Version 1: Implementation, Evaluation, and Attribution. Research Triangle Park, NC. U.S. Environmental Protection Agency. U.S. EPA

US EPA. 2018. Modeling Guidance for Demonstrating 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

US EPA. 2018b. Memorandum: Information on the Interstate Transport State Implementation Plan Submissions for the 2015 Ozone National Ambient Air Quality Standards under Clean Air Act Section 110(a)(2)(D)(i)(I), Research Triangle Park, NC. https://www.epa.gov/sites/production/files/2018-03/documents/transport_memo_03_27_18_1.pdf

US EPA. 2017. Memorandum: Supplemental Information on the Interstate Transport SIP Submissions for the 2008 Ozone NAAQS under Clean Air Act Section 110(a)(2)(D)(i)(I), Research Triangle Park, NC. https://www.epa.gov/sites/production/files/2017-10/documents/final_2008_o3_naaqs_transport_memo_10-27-17b.pdf.

US EPA. 1991. Guideline on the Regulatory Application of the Urban Airshed Model. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, technical Support Division, Source Receptor Analysis Branch, Research Triangle Park, North Carolina. July. (http://www.epa.gov/ttn/scram/guidance/guide/uamreg.pdf)

Appendix A: Monitor-specific Ozone Timeseries

Additional LADCO CAMx 2016aa2a simulation MPE plots are available on the LADCO website:

https://www.ladco.org/technical/modeling-results/ladco-2016-modeling/#Air_Quality/CAMx_LADCO_2016aa2a_2020

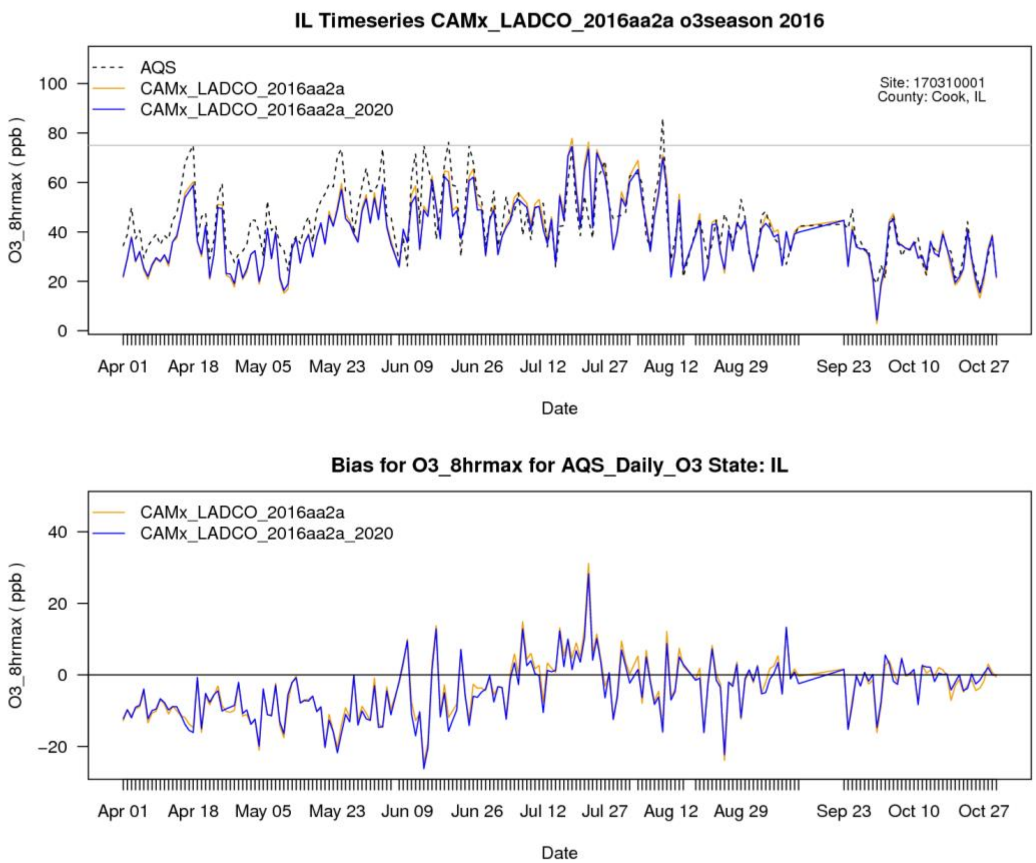

Figure 36. MDA8 O_3 observed, CAMx 2016, and CAMx 2020 concentrations (top) and bias (bottom) at the Alsip, IL monitor

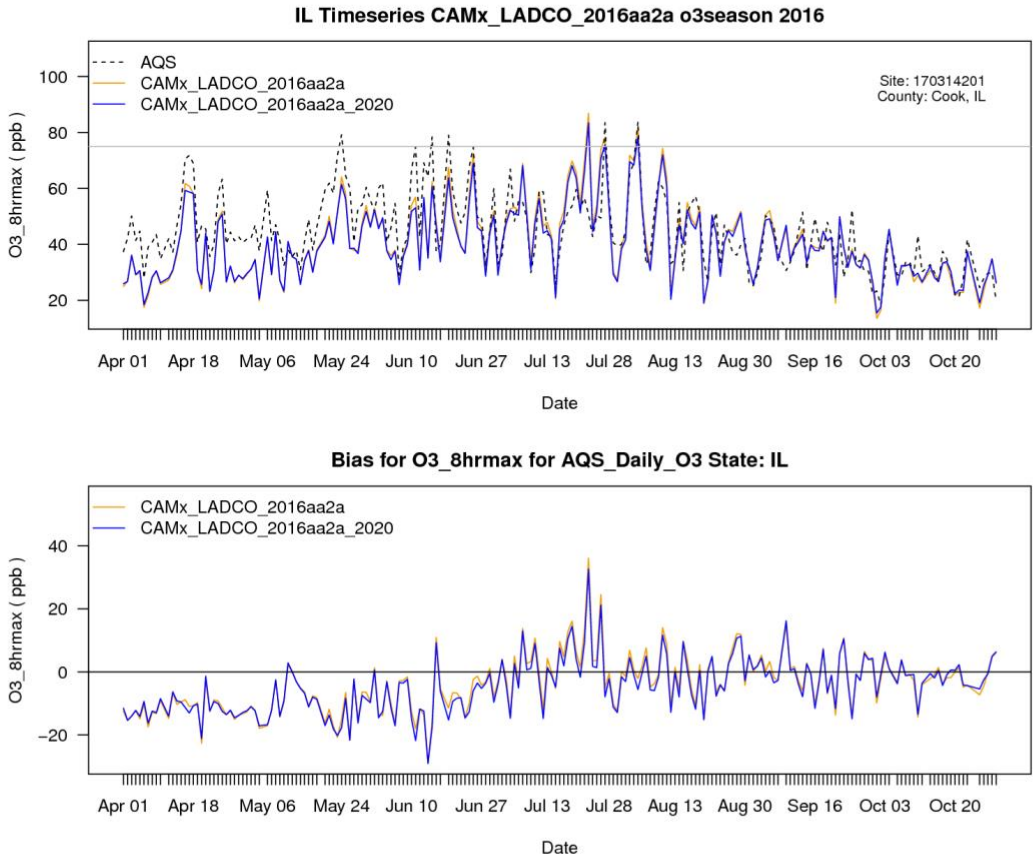

Figure 37. MDA8 O_3 observed, CAMx 2016, and CAMx 2020 concentrations (top) and bias (bottom) at the Northbrook, IL monitor

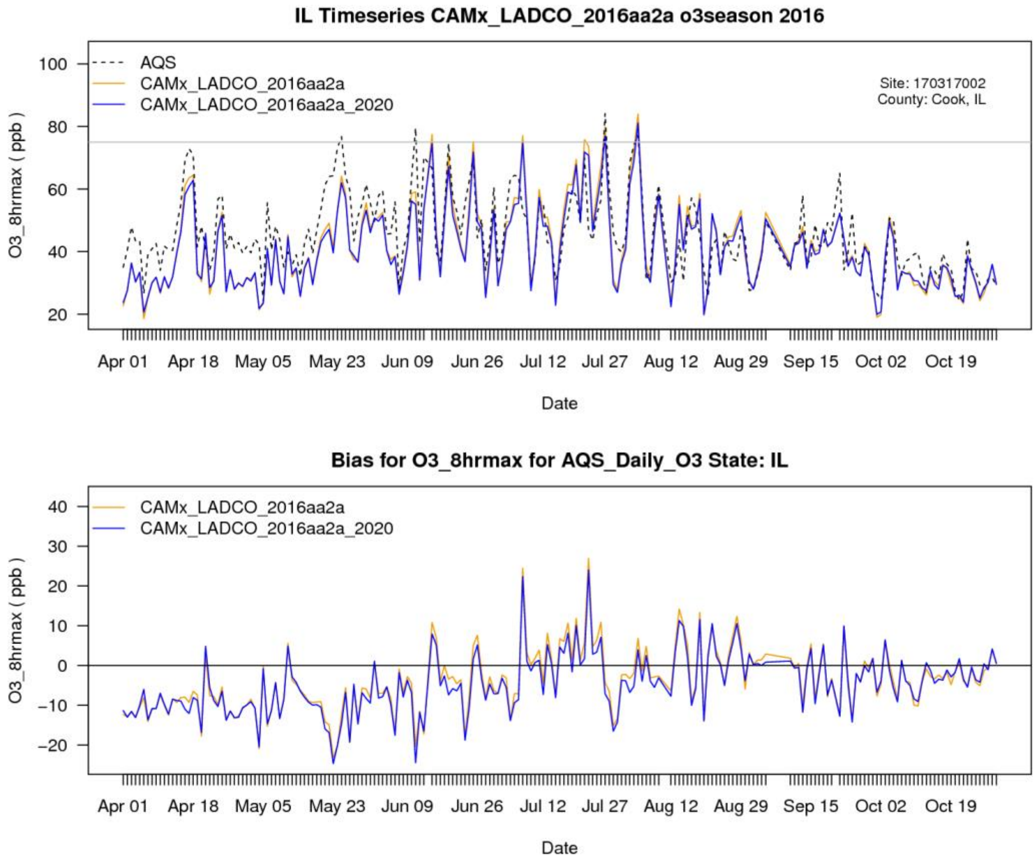

Figure 38. MDA8 O_3 observed, CAMx 2016, and CAMx 2020 concentrations (top) and bias (bottom) at the Evanston, IL monitor

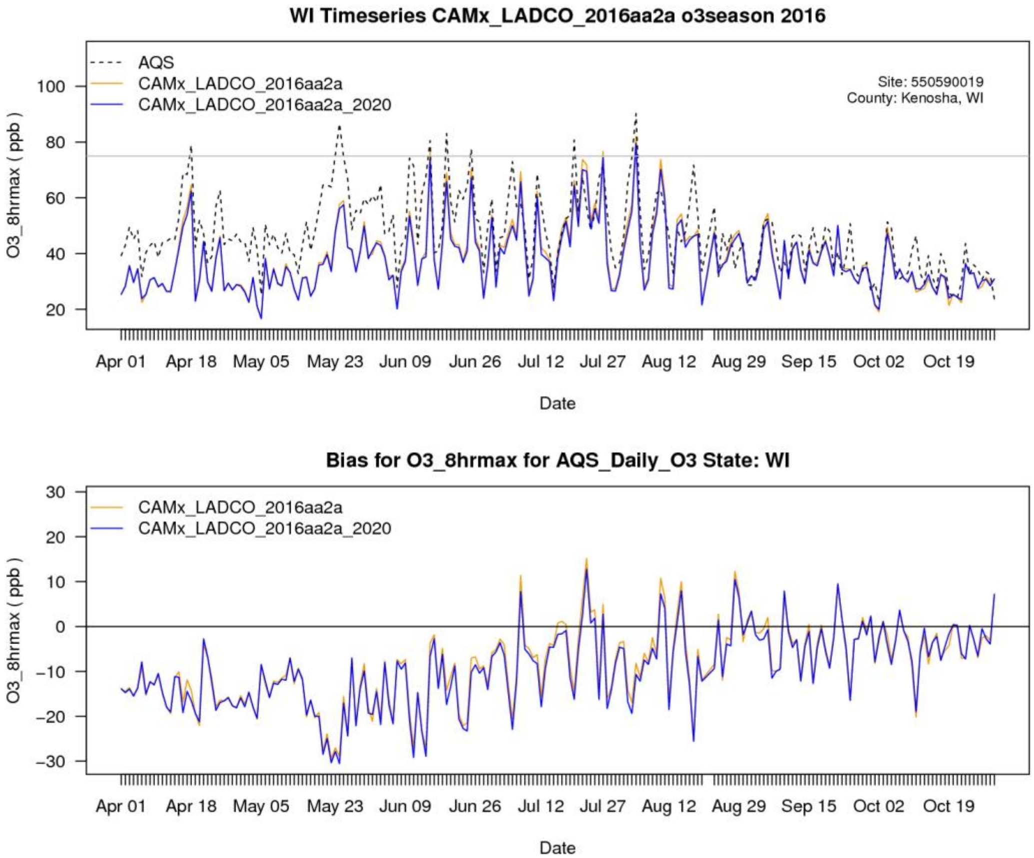

Figure 39. MDA8 O_3 observed, CAMx 2016, and CAMx 2020 concentrations (top) and bias (bottom) at the Chiwaukee Prairie, WI monitor

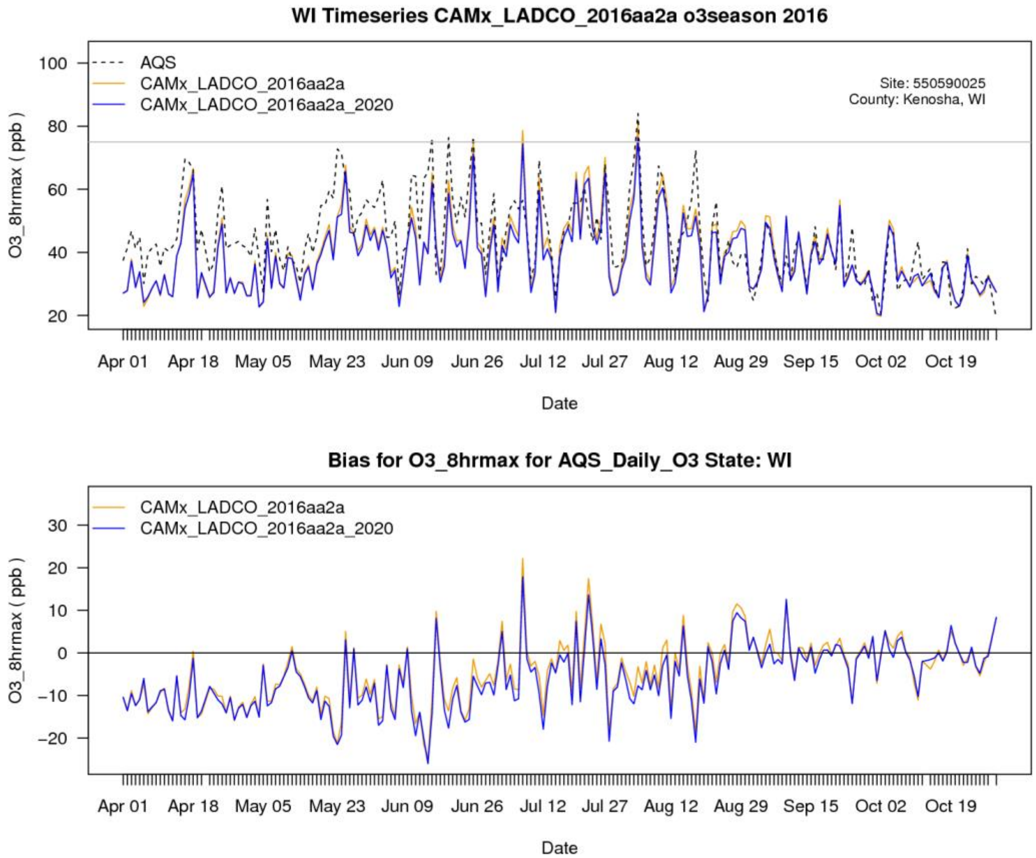

Figure 40. MDA8 O_3 observed, CAMx 2016, and CAMx 2020 concentrations (top) and bias (bottom) at the Kenosha Water Tower, WI monitor

Appendix B: Description of CART Analysis

What is CART Analysis?

- Classification and Regression Tree, aka binary recursive partitioning, decision tree
- Classifies data by yes/no questions -- is temp. < 75, is RH < 80; easy to interpret
- Nonparametric, so insensitive to distributions of variables
- Insensitive to transformations of variables
- Insensitive to outliers and missing data
- Frequently more accurate than parametric models

The beauty of CART is the ease of interpreting the results--you get back a natural-language sequence of questions that anyone can use to classify a new data set You can also adjust the sensitivity of the model to various parameters or outcomes, if it's more important to accurately classify group 1 rather than group 2, for example.

How do you use CART for ozone analysis?

- CART is used to categorize each day by ozone concentration and associated met conditions
- Incorporates 30+ meteorological variables
- Results in a decision tree with 10-15 branches, each describing the meteorological conditions associated with a particular ozone concentration
- Trends are then developed for meteorologically similar days to minimize the effects of meteorological variability on ozone trends

Which meteorology variables were used in the LADCO CART analysis?

These variables were selected from previous model runs that had many more variables included; these are just those that had any influence in previous models:

- Daily precipitation
- Cloud cover

- 850 and 700 mb temperatures at 6 am
- Maximum daily temperature, dew point, relative humidity, pressure
- Average daily wind spee
- Average daily, morning, and afternoon wind direction as N/S and E/W vectors
- Morning, afternoon and evening dewpoint and pressure
- Day of week
- Previous day's average temperature, pressure, wind speed, wind direction
- Change in temperature and pressure from previous day
- 2- and 3-day average wind speed and temperature
- 24-hour transport direction and distance (from Hysplit trajectories)
- Deviation from 10-year averages of 850 and 700 mb temperature and height

Where did the meteorology data come from that were used in the LADCO CART analysis?

- Hourly surface observations from 2379 sites around the US collected from National Climatic Data Center's (NCDC) Integrated Surface Database (mostly airports)
- Upper air observations from 85 sites collected from NCDC's Integrated Global Radiosonde Archive
- Each surface site is paired with closest upper air site (upper air data can be less spatially representative than surface obs)
- Hysplit back trajectories calculated for each site at noon every day to provide transport distance and u,v,w vectors
- Data for each year/site is acquired from NCDC, processed to calculated derived values (daily max/min, mixing heights, e.g.)
- QA flags assigned based on completeness, upper air site proximity
- Lags and deviation from long term means are calculated
- Data are combined and formatted into ASCII and SAS datasets

- Years 2005 to 2018