PGA 582-16-60851-01, Amendment 3
Task 6.2, Deliverable 6.2.2

2012 Modeling Platform Performance Evaluation for the CAPCOG Region

Prepared by the Capital Area Council of Governments
November 28, 2017

PREPARED UNDER A GRANT FROM THE TEXAS COMMISSION ON ENVIRONMENTAL QUALITY

The preparation of this report was financed through grants from the State of Texas through the Texas Commission on Environmental Quality. The content, findings, opinions, and conclusions are the work of the author(s) and do not necessarily represent findings, opinions, or conclusions of the TCEQ.

Table of Contents

1 Introduction

The Capital Area Council of Governments (CAPCOG) covers 10 counties in Central Texas – Bastrop, Blanco, Burnet, Caldwell, Fayette, Lee, Llano, Travis, and Williamson Counties. Five of these counties – Bastrop, Caldwell, Hays, Travis, and Williamson Counties – constitute the Austin-Round Rock Metropolitan Statistical Area (MSA). The CAPCOG region's Federal Reference Method (FRM) air quality monitors show that the region's ground-level ozone (O_3) levels are narrowly attaining the 2015 O_3 National Ambient Air Quality Standard (NAAQS) of 70 parts per billion (ppb). CAPCOG's air quality program seeks to keep the area in attainment of all NAAQS and continue improving air quality within the region. One of the tools CAPCOG uses to do this planning is photochemical modeling.

The Texas Commission on Environmental Quality (TCEQ) has developed a photochemical modeling platform covering May 1, 2012 – September 30, 2012, that it is now using for air quality planning efforts within the state. This model is now in its 2nd release, which CAPCOG has used recently for a control strategy and sensitivity analysis completed in the summer of 2017.[1] CAPCOG used "release 0" of the June segment of this modeling platform for its source apportionment modeling project completed in early 2017[2]. The purpose of this report is to conduct a performance evaluation of this modeling platform for the CAPCOG region. The report provides analysis of how well this model simulates air pollution concentrations within the CAPCOG region by comparing modeled air pollution concentrations to ambient monitoring data collected in the region during the same period. The report will also compare performance statistics for this model to the performance statistics of other models that are available for the region. These analyses should help improve the understanding of the quality of the data produced by this model.

Since CAPCOG's primary regional air quality concern is O_3, CAPCOG focuses on performance statistics for maximum daily 8-hour average (MDA8) O_3 concentrations. In addition to release 2 of the May 1, 2012 – September 30, 2012 model, CAPCOG also used performance statistics for release 4 of TCEQ's May 31, 2006 – July 2, 2006, and its August 13, 2006 – September 15, 2006, modeling platforms, as well as EPA's May 1, 2011 – September 30, 2011 modeling platform as points of comparison. Due to the importance of nitrogen oxides (NO_X) in regional O_3 formation, CAPCOG also analyzed 1-hour NO_X concentrations for this report. CAPCOG is including an analysis of measurement uncertainty in an appendix to this report.

1.1 Modeling Platforms

TCEQ's modeling data are available online at https://www.tceq.texas.gov/airquality/airmod/data. CAPCOG obtained the modeling and monitoring data from TCEQ's "Texas Photochemical Modeling Results – Interactive Time-Series Plot" tool.[3] This includes both the 2006 and 2012 models, including multiple release versions of each. CAPCOG downloaded O_3 hourly and 8-hour data and NO_X hourly data for each monitor in the region for release 2 of the 2012 model and release 4 of the May 31 – July 2, 2006 and August 13 – September 15, 2006 platforms, using the "cell value" spatial option. Both of these

[1] http://www.capcog.org/documents/airquality/reports/2017/6.3.2a-AACOG_Sensitivity_and_Control_Strategy_Modeling_Report.pdf
[2] http://www.capcog.org/documents/airquality/reports/2017/6.1.2b_AACOG_APCA_Report_for_CAPCOG_2017-04-04.pdf
[3] https://www.tceq.texas.gov/airquality/airmod/data/ts?eps=20120501-20120531

models use a 4 km x 4 km fine grid system covering Eastern Texas, a 12 km x 12 km grid system covering all of Texas and most of the area in surrounding states, and a wider "Regional Planning Organization" 36 km x 36 km grid system covering the continental United States. Info on these grid systems can be found here: https://www.tceq.texas.gov/airquality/airmod/data/domain.

EPA currently uses a photochemical modeling platform based on May 1, 2011 – September 30, 2011. EPA refers to this model as its "2011 Version 6" model or "2011v6." This model is now on its third iteration, denoted as "2011v6.3." Info on these platforms can be found here: https://www.epa.gov/air-emissions-modeling/2011-version-6-air-emissions-modeling-platforms. EPA used the 2011v6.3 data as the basis for its Transport Modeling for the 2015 O_3 NAAQS. The technical support document (TSD) for this modeling includes some summary performance statistics for this version of the platform for the State of Texas, but does not include data specific to the CAPCOG region. The 2015 O_3 NAAQS Transport Modeling also does not have projected O_3 data for the CAPCOG region due to too few days with projected MDA8 $O_3 \geq 60$ ppb in 2023. However, monitor-specific performance data is available for 2011v6.2, which EPA used for its 2008 O_3 NAAQS Transport Modeling, which included 2017 projections of O_3 levels at both of CAPCOG's regulatory monitoring stations. Therefore, CAPCOG used this version of the model as the basis for comparisons in this report. These EPA modeling platforms use a 12 km x 12 km grid system for the entire continental U.S., which means that it has a higher spatial resolution outside of Texas and its adjacent states, but it has a lower spatial resolution for the part of Eastern Texas where CAPCOG's 10 counties are located.

1.2 Monitoring Stations Within the Region

A number of FRM ("regulatory") and research ("non-regulatory") air quality monitoring stations were collecting air pollution concentration data between May 1, 2012 – September 30, 2012. Several other sites were in operation during an earlier 2006 episode that TCEQ used and that will be used as a point of comparison in this report. The following table summarizes key information for each monitoring station that collected either O_3 or NO_X data in 2006 or 2012 with data stored in TCEQ's Leading Environmental Analysis Display System (LEADS) online pages.

Table 1-1. O_3 and NO_X Air Quality Monitors with Data in LEADS in the CAPCOG Region in 2006 and 2012

Station Name	Short Name[4]	CAMS #	EPA #	County	O_3 2006	O_3 2012	NO_X 2006	NO_X 2012
Austin Northwest	ANWC	3	484530014	Travis	FRM	FRM	n/a	FRM
Austin Audubon Society	AUDU	38	484530020	Travis	FRM	FRM	FRM	n/a
Fayette County	FAYT	601	481490001	Fayette	Research	Research	Research	n/a
Pflugerville Wastewater	PFLU	613	484530613	Travis	FRM	n/a	FRM	n/a

[4] Used on TCEQ's modeling pages

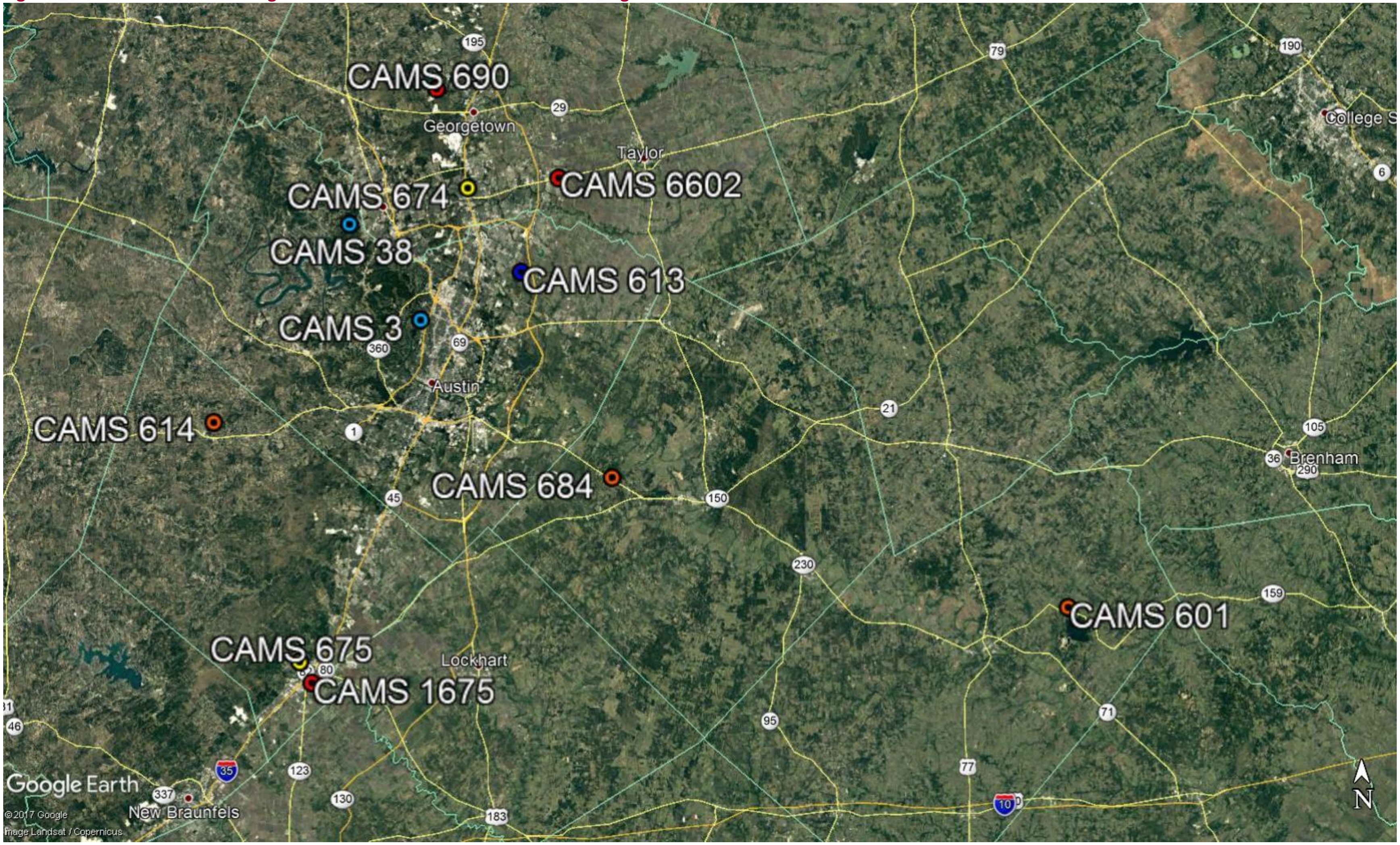
Station Name	Short Name[4]	CAMS #	EPA #	County	O_3 2006	O_3 2012	NO_X 2006	NO_X 2012
Dripping Springs School	DRIP	614	482090614	Hays	Research	Research	Research	n/a
CAPCOG Round Rock	RRCG	674	484910674	Williamson	Research	n/a	n/a	n/a
CAPCOG San Marcos	SMCG	675	482090675	Hays	Research	Research	n/a	n/a
CAPCOG McKinney Roughs	MRCG	684	480210684	Bastrop	Research	Research	n/a	n/a
CAPCOG Lake Georgetown	LGTN	690	484910690	Williamson	Research	Research	n/a	n/a
CAPCOG San Marcos Staples Road	SMSR	1675	482091675	Hays	n/a	Research	n/a	n/a
CAPCOG Hutto College St.	HUTO	6602	484916602	Williamson	n/a	Research	n/a	n/a

Other notes on these sites:

- CAMS 674 in Round Rock and CAMS 675 in San Marcos both were activated on June 2, 2006
- CAMS 675 in San Marcos was shut down on 9/14/2011 and CAMS 1675 established nearby on 9/20/2012. While these monitors are not close enough to keep the same designation, they are close enough for modeling purposes to be considered measuring the same neighborhood-scale and urban-scale O_3 concentrations

The following map shows the locations of the air monitoring sites within the region that were in service during May 31, 2006 – July 2, 2006; August 13, 2006 – September 15, 2006; or May 1, 2012 – September 30, 2012. Blue dots represent regulatory monitors operated by TCEQ. Red dots represent research monitors operated by CAPCOG.

Figure 1-1. Location of Monitoring Stations Included in 2012 or 2006 Modeling Platform



Light blue: TCEQ, 2006 & 2012. Dark blue: TCEQ, 2006 only. Yellow: CAPCOG 2006 only. Red: CAPCOG, 2012 only. Orange: CAPCOG 2006 & 2012.

1.3 Performance Evaluation Statistics

EPA's draft modeling guidance released in 2014 describes a number of different metrics that can be used to analyze the suitability or performance of a particular modeling platform for air quality planning purposes. Performance metrics include:

- Mean bias;
- Mean (gross) error;
- Root mean square error;
- Normalized mean bias;
- Normalized mean error;
- Mean fractional bias;
- Mean fractional error; and
- Correlation coefficient.

CAPCOG focused on a subset of these performance metrics – mean bias, mean error, normalized mean bias, normalized mean error, and root mean square error. CAPCOG also analyzed the maximum and minimum bias and error.

For O_3, EPA recommends comparing observed MDA8 O_3 to modeled MDA8 O_3 when MDA8 $O_3 \geq 60$ ppb. EPA's modeling guidance does not specify whether this means that the observed MDA8 O_3 should be the reference point (i.e., analyze the data for when observed MDA8 $O_3 \geq 60$ ppb) or the modeled MDA8 O_3 should be the reference point (i.e., analyze the data for when modeled MDA8 $O_3 \geq 60$ ppb). However, subsequent modeling analyses conducted by both TCEQ and EPA have used observed MDA8 $O_3 \geq 60$ ppb as the reference point for calculating performance statistics.[5] Therefore, CAPCOG also used observed MDA8 $O_3 \geq 60$ ppb as the key reference point for this analysis. CAPCOG included one analysis of root square mean error for all days as a point of reference.

2 MDA8 O_3 Performance in Release 2 of the 2012 Model for all Sites

This section includes performance statics for MDA8 O_3 levels for release 2 of the 2012 model. This includes statistics for each monitoring station by month and statistic type. Evaluating these data by month allows for a better understanding of how the June segment in particular compares to each other month in this episode, as well as the episode as a whole.

Table 2-1. Days with Observed MDA8 $O_3 \geq 60$ ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	6	8	1	5	4	**24**

[5] See TCEQ's photochemical modeling appendix for the Houston area's most recent attainment demonstration SIP revision (https://www.tceq.texas.gov/assets/public/implementation/air/sip/hgb/HGB_2016_AD_RFP/AD_Adoption/HGB_AD_SIP_Appendix_C_Adoption.pdf) and EPA's TSD for the 2015 O_3 NAAQS modeling (https://www.epa.gov/sites/production/files/2017-01/documents/aq_modeling_tsd_2015_o3_naaqs_preliminary_interstate_transport_assessmen.pdf)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
38	3	8	0	9	5	**25**
601	2	5	0	1	1	**9**
614	6	6	0	7	6	**25**
684	5	3	0	6	4	**18**
690	6	6	0	8	5	**25**
1675	7	6	0	7	5	**25**
6602	6	4	0	3	2	**15**
Sum	41	46	1	46	32	**166**
Monitors ≥ 5	6	6	0	6	4	**8**

Table 2-2. Days with Modeled MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	9	11	6	6	7	**39**
38	3	8	0	9	5	**25**
601	6	4	0	3	1	**14**
614	8	7	1	8	3	**27**
684	5	5	3	4	2	**19**
690	6	10	4	10	8	**38**
1675	5	7	1	6	4	**23**
6602	6	8	3	7	4	**28**
Sum	48	60	18	53	34	**213**
Monitors ≥ 5	7	7	1	6	3	**8**

Table 2-3. Mean Obs. MDA8 O_3 when Obs. MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	66.98	70.58	61.11	72.41	64.91	**68.72**
38	70.72	69.37	n/a	65.45	63.84	**67.02**
601	62.13	67.76	n/a	70.18	60.30	**65.95**
614	68.91	70.00	n/a	66.24	63.91	**67.22**
684	62.90	71.36	n/a	67.13	67.20	**66.68**
690	68.95	69.39	n/a	69.45	64.98	**68.42**
1675	65.46	68.14	n/a	68.66	68.15	**67.54**
6602	64.64	65.88	n/a	70.09	61.40	**65.63**

Table 2-4. Mean Modeled MDA8 O_3 when Obs. MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	67.11	69.44	76.17	65.07	62.57	**67.08**
38	58.36	70.63	n/a	63.08	61.90	**64.69**
601	63.65	65.62	n/a	59.10	55.30	**63.31**
614	63.72	68.53	n/a	64.38	59.43	**64.03**
684	58.32	67.94	n/a	62.29	58.04	**61.18**
690	65.67	68.68	n/a	63.37	64.13	**65.35**
1675	60.37	65.19	n/a	61.15	61.07	**61.88**
6602	63.22	66.09	n/a	63.01	60.81	**63.62**

Table 2-5. Mean Bias when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	0.13	-1.13	15.06	-7.34	-2.34	**-1.64**
38	-12.36	1.26	n/a	-2.38	-1.94	**-2.32**
601	1.53	-2.14	n/a	-11.08	-5.00	**-2.63**
614	-5.19	-1.47	n/a	-1.86	-4.48	**-3.19**
684	-4.58	-3.42	n/a	-4.84	-9.17	**-5.49**
690	-3.28	-0.71	n/a	-6.08	-0.85	**-3.07**
1675	-5.09	-2.95	n/a	-7.52	-7.07	**-5.65**
6602	-1.42	0.21	n/a	-7.08	-0.59	**-2.01**

Table 2-6. Mean Error when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	3.21	3.04	15.06	7.56	4.45	**4.76**
38	12.36	3.86	n/a	5.69	3.44	**5.45**
601	1.53	3.76	n/a	11.08	5.00	**4.21**
614	5.19	2.47	n/a	5.73	5.48	**4.76**
684	4.58	5.20	n/a	7.36	9.37	**6.68**
690	3.92	3.14	n/a	7.09	5.80	**5.12**
1675	5.43	4.10	n/a	7.98	7.07	**6.15**
6602	3.07	2.36	n/a	7.08	1.03	**3.41**

Table 2-7. Normalized Mean Bias when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	0.20%	-1.60%	24.65%	-10.14%	-3.60%	**-2.38%**
38	-17.47%	1.81%	n/a	-3.63%	-3.04%	**-3.47%**
601	2.46%	-3.15%	n/a	-15.79%	-8.30%	**-3.99%**
614	-7.53%	-2.09%	n/a	-2.81%	-7.01%	**-4.75%**
684	-7.28%	-4.79%	n/a	-7.21%	-13.64%	**-8.24%**
690	-4.76%	-1.02%	n/a	-8.75%	-1.30%	**-4.49%**
1675	-7.78%	-4.32%	n/a	-10.95%	-10.38%	**-8.37%**
6602	-2.19%	0.31%	n/a	-10.10%	-0.96%	**-3.06%**

Table 2-8. Normalized Mean Error when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	4.79%	4.30%	24.65%	10.44%	6.86%	**6.92%**
38	17.47%	5.56%	n/a	8.69%	5.38%	**8.13%**
601	4.68%	8.86%	n/a	15.79%	8.30%	**15.79%**
614	7.53%	3.53%	n/a	8.65%	8.58%	**7.08%**
684	7.28%	7.29%	n/a	10.96%	13.95%	**10.01%**
690	5.69%	4.53%	n/a	10.21%	8.92%	**7.49%**
1675	8.30%	6.02%	n/a	11.62%	10.38%	**9.11%**
6602	4.76%	3.58%	n/a	10.10%	1.67%	**5.20%**

Table 2-9. Max Error when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	10.96%	9.23%	24.65%	21.07%	7.67%	**24.65%**
38	19.15%	10.00%	n/a	19.46%	7.77%	**19.46%**
601	4.68%	8.86%	n/a	15.79%	8.30%	**15.79%**
614	16.47%	10.22%	n/a	21.99%	14.83%	**21.99%**
684	16.96%	10.88%	n/a	26.19%	18.55%	**26.19%**
690	12.11%	9.18%	n/a	21.31%	14.41%	**21.31%**
1675	15.62%	9.66%	n/a	30.69%	17.41%	**30.69%**
6602	10.57%	4.57%	n/a	15.15%	2.65%	**15.15%**

Table 2-10. Root Mean Square Error when Observed MDA8 O_3 ≥ 60 ppb by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	4.21	3.57	15.06	10.71	4.49	**6.74**
38	12.40	4.46	n/a	7.06	3.79	**6.76**
601	2.06	3.92	n/a	11.08	5.00	**5.09**
614	6.59	3.52	n/a	7.80	6.21	**6.30**
684	5.61	5.53	n/a	9.81	10.73	**8.46**
690	5.25	3.67	n/a	8.90	6.10	**6.53**
1675	4.21	3.57	15.06	10.71	4.49	**6.74**
6602	3.90	2.40	n/a	8.09	1.18	**4.57**

Table 2-11. Root Mean Square Error when Observed MDA8 O_3 all days by Month and Monitor, 2012 Model, Release 2 (ppb)

CAMS	May	Jun.	Jul.	Aug.	Sep.	Total
3	9.97	8.06	11.08	9.88	9.39	**9.73**
38	10.23	7.42	10.72	8.90	9.62	**9.44**
601	7.92	6.88	10.43	12.14	8.24	**9.43**
614	7.19	6.42	9.15	7.76	10.67	**8.39**
684	9.05	11.23	9.42	8.67	8.63	**9.43**
690	9.64	7.08	10.72	8.25	6.98	**8.67**
1675	9.97	8.06	11.08	9.88	9.39	**9.73**
6602	8.79	6.63	12.84	11.69	10.00	**10.29**

Based on these statistics, the June segment appears to be the best month to use for modeling if the full 5-month period is not modeled.

- June is tied with August for the largest number of MDA8 O_3 observations ≥ 60 ppb at 46, followed by May at 41, September at 32, and July at 1
- June has the most number of days with modeled MDA8 O_3 concentrations ≥ 60 ppb at these grid cells at 60, followed by August with 53, May with 48, September with 34, and July with 18
- June has the largest number of both observed and modeled MDA8 O_3 days ≥ 60 ppb for the key regulatory monitor in the region – CAMS 3

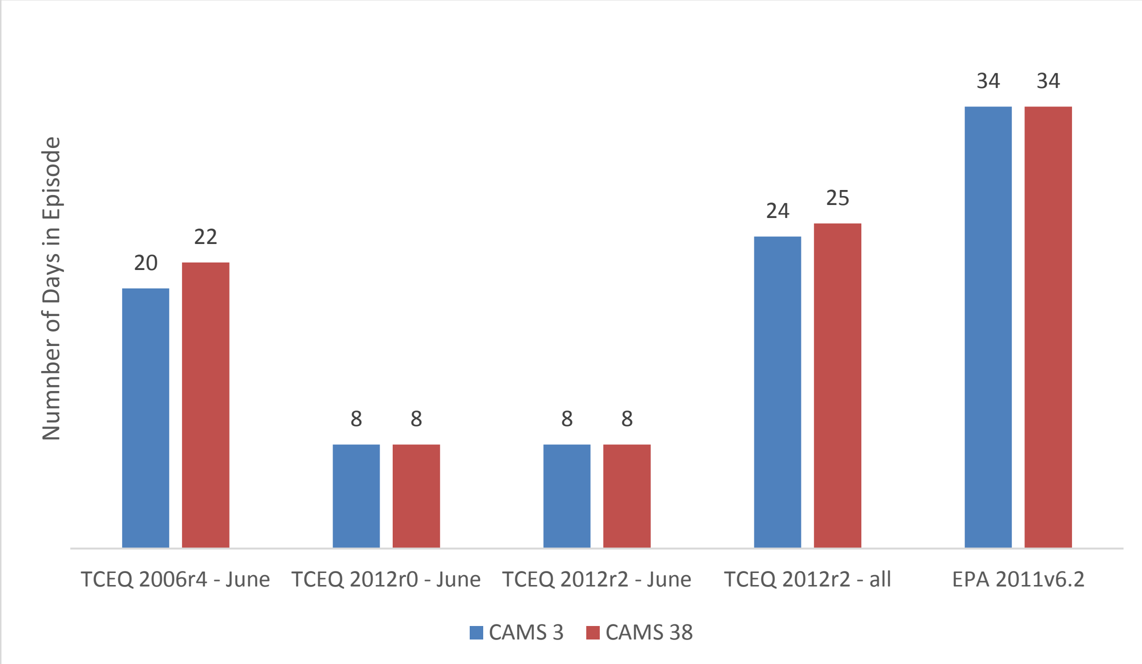
- June is tied with May and August for having the largest number of monitoring stations with at least five days with observed MDA8 $O_3 \geq 60$ ppb in that month at 6 out of 8
- June is tied with May for having the largest number of monitoring stations with at least five days with modeled MDA8 $O_3 \geq 60$ ppb in that month at 7 out of 8, which is significant for two reasons:
    - EPA guidance on performance modeled attainment tests for projected O_3 levels requires at least 5 days with modeled MDA8 $O_3 \geq 60$ ppb
    - EPA's Transport Modeling for the O_3 NAAQS and CAPCOG's recent source apportionment modeling analysis rely on the average contribution on the top 5 modeled MDA8 O_3 days
- June has the smallest normalized mean error for four of the eight monitoring stations, including CAMS 3
- The maximum error for any day with observed MDA8 $O_3 \geq 60$ ppb is smaller in June than any other month for half of the monitoring stations, and the 2nd-smallest for CAMS 3
- June's RSME is smaller than the seasonal RSMEs for all monitors with observed MDA8 $O_3 \geq 60$ ppb and for all monitors on all days except for CAMS 684

3 Comparisons in MDA8 O_3 Performance across Models

This section provides a comparison for some of the key performance model performance statistics across various modeling episodes for the two regulatory monitors in the CAPCOG region. CAPCOG's source apportionment modeling completed in early 2017 used both release 4 of TCEQ's June 2006 platform and release 0 of TCEQ's June 2012 platform. CAPCOG's recent control strategy and sensitivity modeling analysis used the June segment of release 2 of TCEQ's 2012 modeling platform. EPA's 2011v6.2 platform was used for its 2017 projections for the 2008 O_3 NAAQS Transport Modeling. All five months of release 2 of TCEQ's 2012 model were used for TCEQ's most recent Houston-area attainment demonstration.

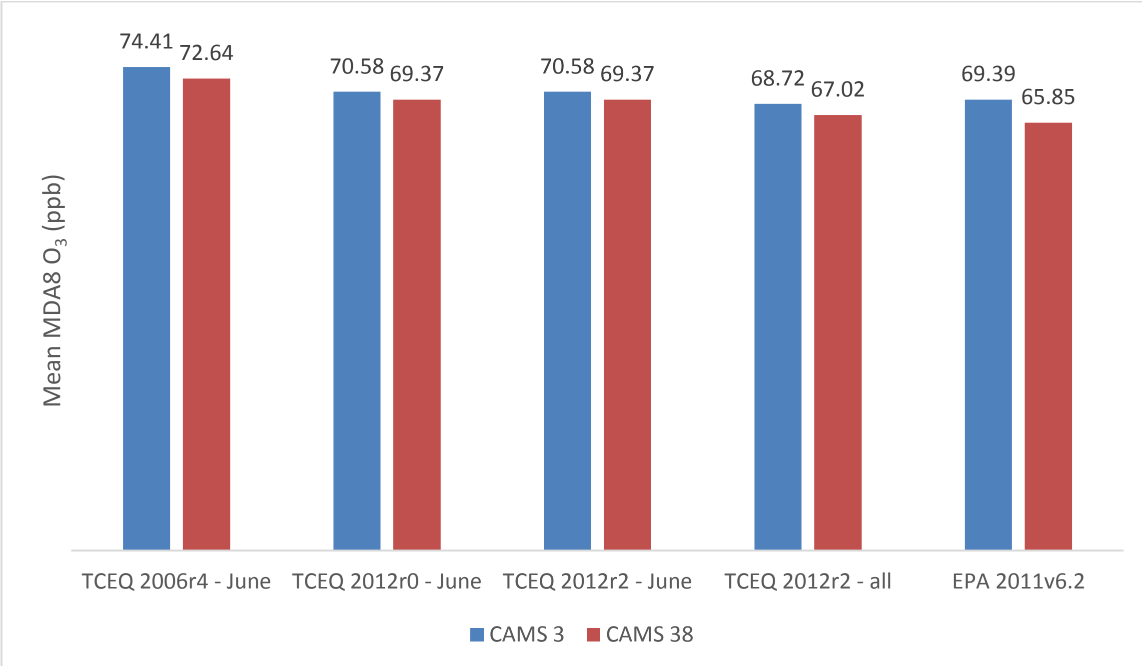
The first figure shows the number of MDA8 O_3 observations by modeling platform. The June 2012 segment has considerably fewer days with observed MDA8 $O_3 \geq 60$ ppb than the other platforms analyzed.

Figure 3-1. Observed MDA8 $O_3 \geq 60$ ppb in TCEQ and EPA Modeling Platforms



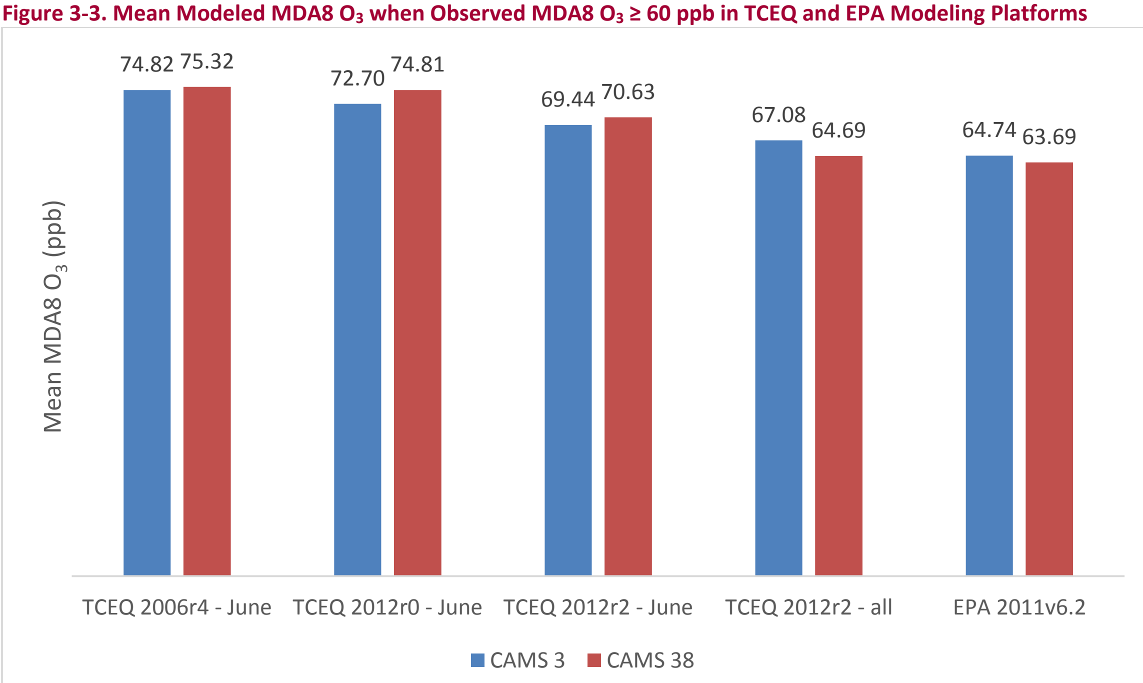
The next figure shows the mean observed MDA8 O_3 concentrations for each modeling platform on days when MDA8 $O_3 \geq 60$ ppb. As would be expected in light of the downward trend in O_3 design values between 2006 and 2012, the June 2006 episode has higher mean observed MDA8 O_3 concentrations than EPA's 2011 seasonal model, TCEQ's 2012 seasonal model, and the June portion of TCEQ's 2012 seasonal model. The 2011 and 2012 models had mean observed MDA8 O_3 near the 70 ppb O_3 NAAQS for CAMS 3.

Figure 3-2. Mean Observed MDA8 O_3 when Observed MDA8 O_3 ≥ 60 ppb in TCEQ and EPA Modeling Platforms



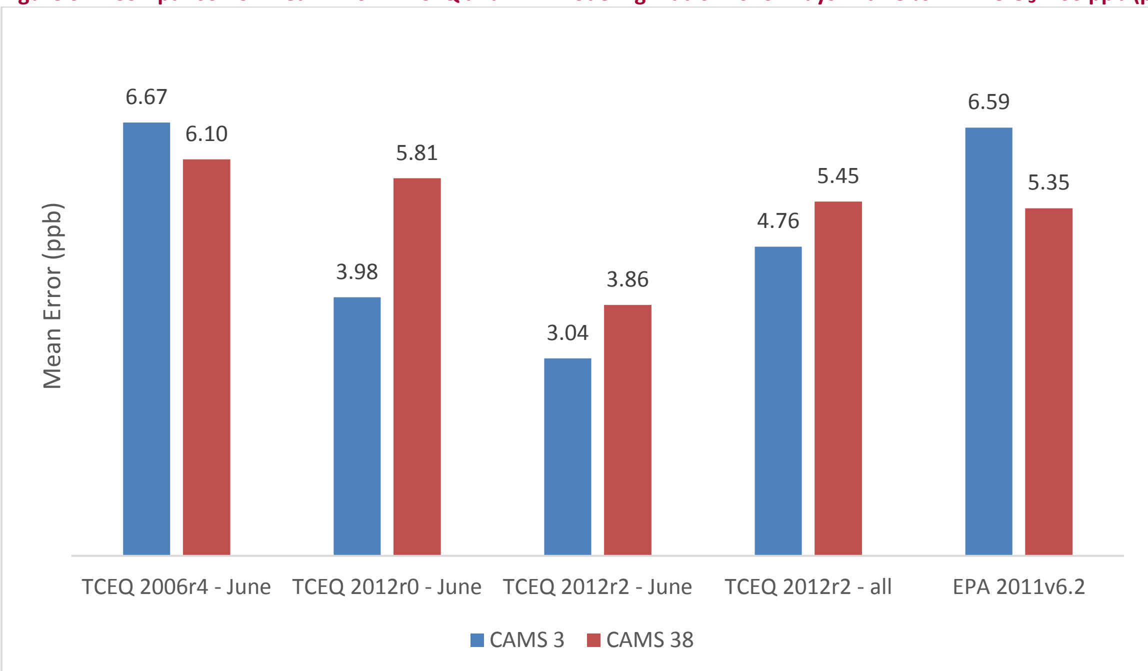
The next figure shows the mean modeled MDA8 O_3 concentrations for each modeling platform on days when observed MDA8 O_3 ≥ 60 ppb. Notably, mean modeled MDA8 O_3 concentrations for release 2 of the June 2012 platform are several ppb lower at both monitoring stations than release 0. Both releases, as well as the entire seasonal model, actually have higher modeled mean MDA8 O_3 than the 2011v6.2 platform.

Figure 3-3. Mean Modeled MDA8 O_3 when Observed MDA8 O_3 ≥ 60 ppb in TCEQ and EPA Modeling Platforms



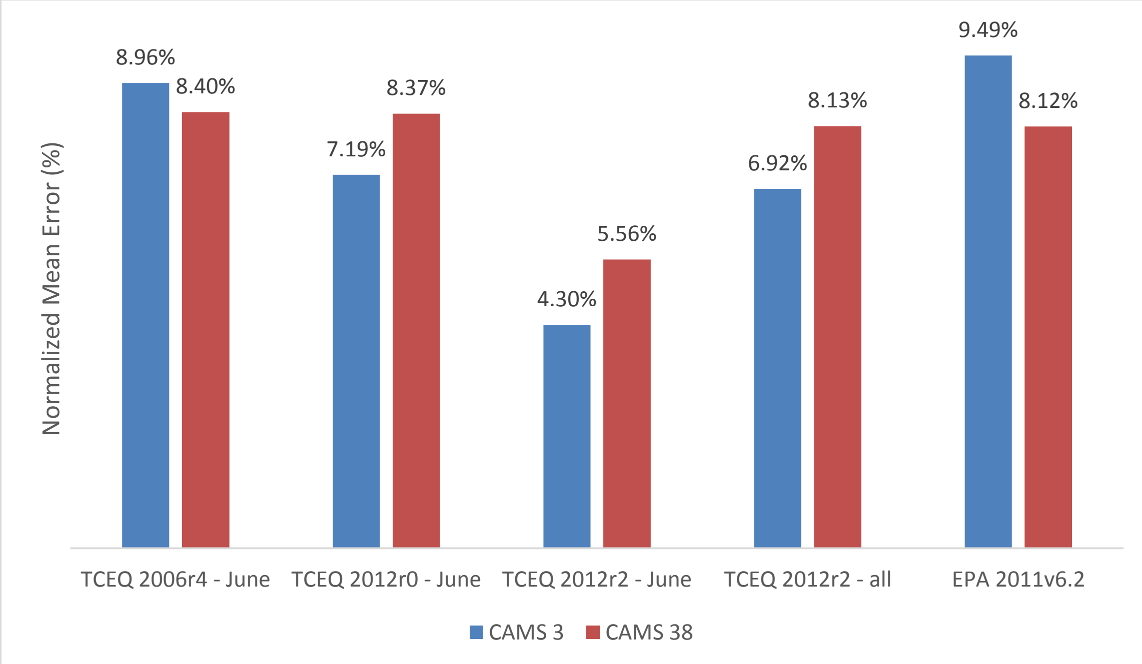
The following figure shows the mean error for each platform when observed MDA8 O_3 ≥ 60 ppb. As the figure shows, the June 2006 episode has the largest mean error for both stations, while release 2 of the June 2012 episode has the smallest mean error for both stations. Release 0 of the June 2012 model actually also performs better than EPA's 2011v6.2 model for these two monitors, as does the full five-month release 2 of the 2012 model.

Figure 3-4. Comparison of Mean Error in TCEQ and EPA Modeling Platforms for Days with Obs. MDA8 O_3 ≥ 60 ppb (ppb)



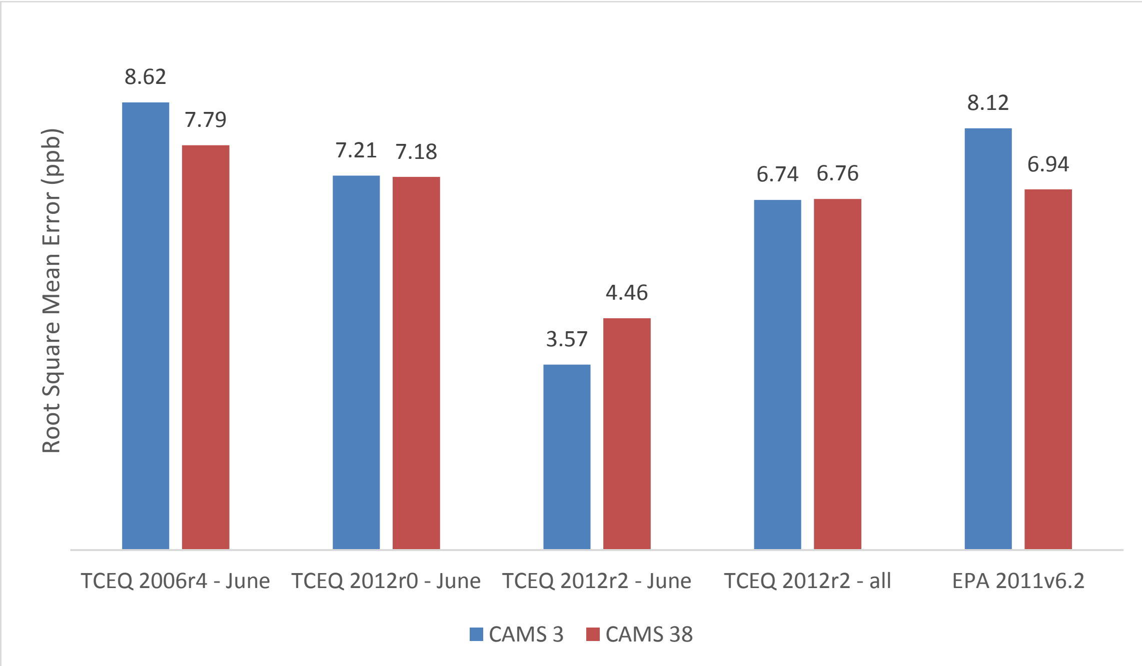
The next comparison graph shows the normalized mean error (NME) for each platform. As the figure shows, release 2 of the June 2012 segment performs substantially better than all of the other options analyzed here, with less than 6% NME for both monitoring stations. The other options all have NME ranging from 6.9% - 9.5%. This suggests that release 4 of the June 2006 platform, release 0 of the June 2012 platform, and the full five-month version of release 2 of the 2012 platform all have NME statistics comparable to EPA's 2011v6.2.

Figure 3-5. Comparison of Normalized Mean Error in TCEQ and EPA Modeling Platforms for Days with Obs. MDA8 O_3 ≥ 60 ppb (%)



The next comparison graph shows the root square mean error (RSME) for each platform. Similar to the statistics above, the figure below shows that the June segment of release 2 of the 2012 model performs substantially better than all of the other options analyzed here.

Figure 3-6. Comparison of Root Mean Square Error in TCEQ and EPA Modeling Platforms for Days with Obs. MDA8 O_3 ≥ 60 ppb (ppb)

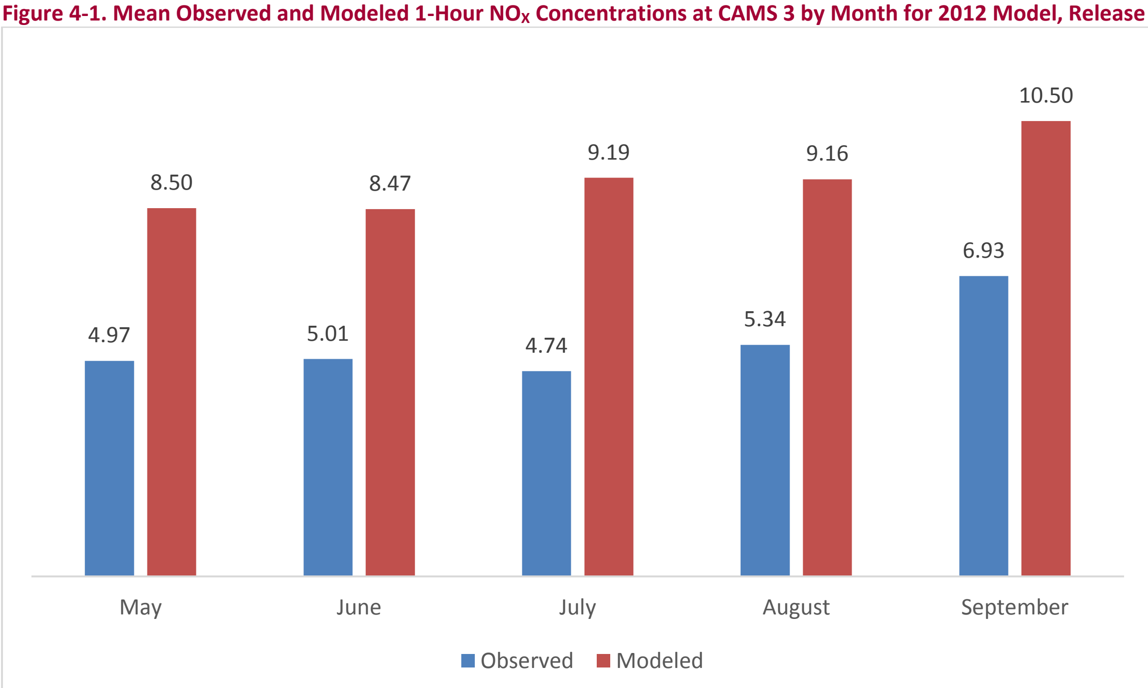


These comparisons all indicate the June segment of release 2 of the 2012 model has a higher degree of accuracy for CAPCOG's two regulatory O_3 monitors than EPA's 2011v6.2 model, and that the June 2006 release 4 and June 2012 release 0, which CAPCOG has also used for modeling analyses recently, have comparable or better performance than EPA's 2011v6.2 model. Since EPA's model has been used for nation-wide O_3 transport analysis for regulatory purposes, these comparisons suggest that each of the TCEQ modeling platforms analyzed above would be at least as good or better at predicting high MDA8 O_3 levels within the region for air quality planning purposes for the region.

4 Performance of Modeled 1-Hour NO_X Concentrations

The following figure shows the mean observed and modeled 1-hour NO_X concentrations at CAMS 3 by month for release 2 of the 2012 model. This includes all days, not just days when MDA8 $O_3 \geq 60$ ppb.

Figure 4-1. Mean Observed and Modeled 1-Hour NO_X Concentrations at CAMS 3 by Month for 2012 Model, Release 2 (ppb)



The average modeled concentrations were 3.5 – 4.5 ppb and 50-100% higher than observed concentrations over the course of each month of this episode. When limiting the analysis to only days with MDA8 $O_3 \geq 60$ ppb, the mean concentrations are higher – 6.67 ppb observed and 10.08 ppb, a bias of +3.42 ppb and 51%.

Despite these substantial differences between the observed and modeled ground-level NO_X concentrations at CAMS 3, it obviously did not mean that there were similarly substantial differences in the MDA8 O_3 concentrations.

5 Conclusion and Future Analysis

TCEQ's 2012 model, including the June 2012 segment in both release 0 and release 2, performs well compared to other modeling platforms used by CAPCOG, TCEQ, and EPA for other studies. The June segment of release 2 of the 2012 model stands out as the most useful single month among the five available months for this release, lending additional creditability to the modeling results from CAPCOG's control strategy and sensitivity modeling completed in 2017. Release 0 of this platform, which was one of the two platforms CAPCOG used for its 2017 source apportionment analysis, did not have quite as good performance as the June segment of release 2, but it had comparable performance to the June 2006 platform, the entire five-month 2012 release 2, and EPA's 2011v6.2. This suggests that the modeling data in both studies is reliable and that future studies using these platforms should be expected to provide high-quality data as well. Modeled NO_X concentrations were about 50% higher than observed NO_X concentrations at CAMS 3 when MDA8 O_3 at that location was ≥ 60 ppb. It isn't clear why there would be such a large difference, but the solid performance in predicting MDA8 O_3 concentrations suggests that this is not a cause for concern.

Under Task 6.4, CAPCOG may perform some additional performance evaluation for sub-sets of days that were important to the source apportionment and control strategy/sensitivity analyses. Specifically, calculating the performance of the five days used to calculate relative contribution factors for the source apportionment analysis and calculating the performance of the top 10 modeled MDA8 O_3 ≥ 60 ppb would add additional value and perspective to the modeling results for these studies. Gaining a better understanding of the potential causes for the large differences in observed and modeled NO_X concentrations could also help CAPCOG better understand how to interpret modeling results for this pollutant.

Appendix A: Monitoring Measurement Uncertainty

One of the factors that is not always considered in evaluating the performance of an air quality model is the quality of the air pollution measurements that will be used as a point of comparison to the model output. If a modeled MDA8 O_3 concentration is 5% higher than a measured MDA8 O_3 concentration, and the instrument reports an O_3 concentration 3% higher than a 90 ppb reference concentration, the difference between modeled O_3 and "actual" O_3 is actually higher than the 5% difference between modeled and observed O_3.

For the 2012 episode, data from two Federal Reference Method (FRM) monitors operated by TCEQ (CAMS 3 and CAMS 38) are available, and data from six non-FRM research monitors operated by CAPCOG (CAMS 601, 614, 684, 690, 1675, and 6602) are available. The FRM sites used a different set of QA/QC procedures from the procedures used at CAPCOG's six non-FRM sites, meaning that direct comparison of performance statistics between the FRM monitors and non-FRM monitors is problematic without accounting for these differences.

EPA's most recent monitoring handbook from January 2017 identifies deviations of +/- 1.5 ppb or +/- 7% from reference concentrations (whichever is greater) as the appropriate acceptance criteria.[6] An earlier version used +/- 7% for one-point QC checks and 3% of full scale for a zero check (equivalent to 15 ppb for most analyzers).[7]

TCEQ FRM Sites

The two FRM sites are equipped with automated calibration equipment that enables automatic checks of the accuracy of the instrument's measurements and adjustments of the data recorded by the instrument in what gets transmitted to TCEQ's database in order to account for the most recent calibrations. Every 28 days, the system conducted a five-point calibration at 0 ppb, 90 ppb, 200 ppb, 300 ppb, and 400 ppb. Based on the values recorded by the instrument in response to these reference concentrations, a linear regression produced slope and intercept values that were used to adjust the data recorded by the instrument and bring it closer in line with what the "ideal" slope (1) and intercept (0) would be.

For example, there was a five-point calibration of CAMS 3's O_3 instrument on August 21, 2012 at 21:45 CST. The following table shows the results of that calibration (note – values reported in millivolts, or mV, with 2,000 mV = 1 ppm O_3).

Table A-1. Example of five-point calibration results at CAMS 3 for August 21, 2012

Reference Concentration	O_3 Instrument Value Recorded (mV)	Difference (mV)	% Difference
0 mV (0.00 ppm)	13.0	+13.0	n/a
180 mV (0.090 ppm)	196.7	+16.7	+9.29%
400 mV (0.200 ppm)	420.7	+20.7	+5.18%
600 mV (0.300 ppm)	624.3	+24.3	+4.05%
800 mV (0.400 ppm)	828.3	+28.3	+3.38%

[6] https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/Final%20Handbook%20Document%201_17.pdf
[7] https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/vol2sec03.pdf

Using these data, the system generates an intercept set at the value recorded at the 0 mV check point (13.0 mV) and a slope based on a regression through these points, using the 0 mV measurement level as a fixed intercept point. For this particular calibration, therefore, the intercept was 0 mV and the slope was 2,038.21 mV/ppm. The full expression of the relationship, with x = the concentration recorded in LEADS (in ppm) and y = the concentration recorded by the instrument (in ppm) is shown below:

$$Ideal\ Slope \times Instrument\ Concentration + Ideal\ Incercept = Instrument\ Slope \times LEADS\ Concentration + Instrument\ Intercept$$

$$Instrument\ Concentration = \frac{Instrument\ Slope \times LEADS\ Concentration + Instrument\ Intercept - Ideal\ Intercept}{Ideal\ Slope}$$

$$LEADS\ Concentration = \frac{Ideal\ Slope \times Instrument\ Concentration + Ideal\ Incercept - Instrument}{Instrument\ Slope}$$

$$2000\frac{mV}{ppm} \times Y\ (Instrument)\ ppm + 0\ mV = 2038.21\frac{mV}{ppm} \times X(LEADS)\ ppm + 13.0\ mV$$

$$X\ (LEADS)\ ppm = \frac{2000\frac{mV}{ppm} \times X(Instrument)\ ppm + 0\ mV - 13.0\ mV}{2038.21\frac{mV}{ppm}}$$

$$Y\ (Instrument)\ ppm = \frac{2038.21\frac{mV}{ppm} \times X(LEADS)\ ppm + 13.0\ mV - 0\ mV}{2000\frac{mV}{ppm}}$$

On August 22, 2017, there was is 1-hour O_3 concentration of 65 ppb recorded in leads for 4:00 pm – 5:00 pm. This value included the adjustments described above. The following equation shows the calculation for the instrument's reading versus the value reported to LEADS.

$$Y\ (Instrument)\ ppm = \frac{2038.21\frac{mV}{ppm} \times 0.065\ ppm + 13.0\ mV - 0\ mV}{2000\frac{mV}{ppm}} = 0.0727\ ppm$$

This means that the instrument's recorded value was actually 11.9% higher than the value recorded in LEADS. Since the LEADS system includes this adjustment, however, 65 ppm would be the best estimate for the actual 1-hour avg. O_3 concentration that CAMS 3 was sampling during this time.

TCEQ's system applies the slope and correction factors to all data starting with the completion of the five-point calibrations (marked as "CAL") that occur every 28 days. Every seven days, the system also conducted a 3-point "SPAN" check at 0 ppb, 90 ppb, and 400 ppb levels. Based on the results of this check, the intercept point in the data system is adjusted. So, for example, the SPAN check conducted on August 28, 2012, at 21:50 CST produced values of 12.33 mV, 195.3 mV, and 825.3 mV at the 0 mV, 180 mV, and 800 mV check points. Since the most recent five-point calibration check was the August 21, 2012, check described above, the slope remained 2038.21 mV/ppm, while the intercept changed from 13.0 mV to 12.33 mV.

In addition to these checks, there are also "SPANZ" checks conducted on all other days at night that test 0 ppb and 400 ppb levels. As long as the values are within accepted ranges, the measurements are accepted with no additional adjustments.

One way to characterize measurement uncertainty for a given O_3 measurement is to compare the value that would be reported in LEADS based on the most recent CAL or SPAN adjustment and the next CAL or SPAN adjustment. For example, the 90 ppb check conducted on August 21, 2012, showed a value of 98.35 ppb. With slope and intercept adjustments based on that day's calibration, the value that would have been reported to LEADS for a 90 ppb concentration would have been 90.1 ppb. This means that the LEADS-recorded value was only off by 0.1 ppb, or 0.1%. Using the same 98.35 ppb concentration recorded for the 90 ppb check on August 21, 2012, and the updated intercept applied based on the SPAN check that occurred on August 28, there would be a LEADS-reported value of 90.5 ppb, a difference of 0.5%. Using the same 98.35 ppb concentration recorded on August 21, 2012, and instead applying the slope and intercept corrections from the next five-point calibration that occurred on September 18, 2012, the value reported in LEADS would instead have been 90.7 ppb – off by 0.7 ppb or 0.8%. In this example, you could assume that "actual" O_3 concentrations were about 0.1 ppb – 0.7 ppb lower than what showed up in LEADS within this time-frame.

The following tables show the five-point calibration check values for CAMS 3 and 38 covering the 2012 episode.

Table A-2. 2012 Five-Point Calibration Reference Concentration Deviations for CAMS 3 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)	Note
4/24/2012	19:00 CST	6.30	13.94%	6.47%	4.02%	6.36%	Passed
5/12/2012	21:45 CST	6.37	14.28%	6.85%	4.08%	6.08%	Passed
6/22/2012	17:45 CST	6.17	9.06%	5.18%	4.00%	3.50%	Passed
6/26/2012	21:45 CST	6.00	9.06%	5.43%	4.17%	3.66%	Passed
7/20/2012	16:10 CST	61.67	-27.22%	-66.75%	-77.55%	-83.71%	Failed
7/23/2012	10:30 CST	6.00	9.61%	4.93%	3.83%	3.29%	Passed
7/24/2012	21:45 CST	6.34	9.44%	5.25%	4.22%	3.54%	Passed
8/21/2012	21:45 CST	6.50	9.28%	5.18%	4.05%	3.54%	Passed
9/18/2012	21:45 CST	6.00	9.06%	4.83%	3.72%	3.38%	Passed
10/16/2012	21:45 CST	6.17	9.83%	4.83%	4.00%	3.50%	Passed

Table A-3. 2012 Five-Point Calibration Reference Concentration Deviations for CAMS 3 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)	Note
4/14/2012	21:45 CST	-0.43	-4.94%	-6.38%	-6.08%	-6.60%	Passed
5/12/2012	21:45 CST	-0.43	-5.22%	-6.28%	-6.10%	-6.46%	Passed
5/21/2012	14:30 CST	N/A	N/A	N/A	N/A	N/A	Incomplete

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)	Note
5/22/2012	14:20 CST	-0.63	-2.67%	0.10%	-2.20%	-0.99%	Passed
6/9/2012	21:45 CST	-0.33	-4.33%	-0.90%	-2.48%	-1.00%	Passed
6/7/2012	21:45 CST	-0.77	-4.56%	-0.80%	-2.80%	-1.24%	Passed
8/4/2012	21:45 CST	-0.20	-3.89%	-0.85%	-2.63%	-1.26%	Passed
9/1/2012	21:45 CST	-0.97	-4.33%	-0.95%	-2.73%	-1.29%	Passed
9/29/2012	21:45 CST	-0.67	-4.17%	-0.70%	-2.45%	-1.19%	Passed
10/27/2012	21:45 CST	0.20	-2.78%	0.58%	-1.58%	-0.22%	Passed

CAPCOG Sites

Unlike TCEQ sites, CAPCOG sites relied on manual calibration checks, and did not include adjustments to the data recorded by the instrument. Therefore, to the extent that there were differences between the values recorded by the instrument and the reference concentrations when a calibration occurred, those differences would be present in the data recorded in LEADS as well.

Where the exact times of a calibration were unknown, CAPCOG assumed the calibration occurred at noon of that day.

Table A-4: 2012 Calibration Reference Concentration Deviations for CAMS 601 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
4/10/2012	Unknown	0	1.11%	1.00%	0.67%	-1.25%
6/30/2012	Unknown	1	2.22%	1.00%	1.00%	-0.75%
7/31/2012	Unknown	2	1.11%	N/A	N/A	-1.00%
9/14/2012	14:27:23	1	0.00%	4.50%	5.00%	5.75%
10/23/2012	12:38:15	3	11.11%	5.50%	6.67%	6.50%

Table A-5: 2012 Calibration Reference Concentration Deviations for CAMS 614 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
4/11/2012	18:22:57	1.86	2.28%	0.30%	1.16%	0.91%
5/31/2012	13:32:07	3.20	3.58%	N/A	N/A	1.04%
6/28/2012	14:18:35	1.50	-0.54%	N/A	N/A	[8]85.61%
7/26/2012	11:00:11	3.00	5.86%	2.44%	3.13%	3.52%
8/29/2012	10:27:06	3.10	2.61%	N/A	N/A	1.26%
10/24/2012	9:32:05	3.00	-14.12%	-16.95%	-13.84%	-7.43%

[8] CAPCOG believes that this reading was recorded in the operator logs incorrectly

Table A-6: 2012 Calibration Reference Concentration Deviations for CAMS 684 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
5/30/2012	11:41:55	0.6	7.89%	N/A	N/A	6.95%
6/13/2012	13:04:28	1.8	9.78%	N/A	N/A	7.00%
6/13/2012	13:04:28	1.0	13.33%	8.45%	7.63%	7.44%
7/25/2012	11:06:54	3.8	17.33%	11.05%	9.73%	10.15%
8/28/2012	12:43:40	4.4	17.33%	N/A	N/A	11.03%
10/23/2012	15:42:40	3.0	16.67%	10.50%	9.33%	8.75%

Table A-7: 2012 Calibration Reference Concentration Deviations for CAMS 690 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
4/14/2012	13:03:46	1.8	-1.33%	-1.15%	0.13%	0.50%
5/2/2012	13:38:56	0.2	-0.28%	N/A	N/A	0.92%
6/12/2012	12:45:00	0.0	-1.11%	N/A	N/A	0.75%
6/14/2012	12:43:11	0.1	-2.11%	N/A	N/A	0.75%
6/14/2012	12:43:11	0.0	22.33%	4.45%	4.33%	2.78%
7/10/2012	13:15:29	1.4	10.22%	4.55%	4.50%	4.05%
7/23/2012	10:23:03	1.8	11.00%	5.10%	5.53%	4.00%
8/20/2012	12:35:02	3.0	-1.67%	N/A	N/A	5.00%
10/30/2012	10:47:27	2.0	6.67%	4.00%	4.67%	4.00%

Table A-8: 2012 Calibration Reference Concentration Deviations for CAMS 1675 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
4/11/2012	13:56:54	2.83	5.22%	4.74%	5.00%	3.38%
4/13/2012	17:31:13	N/A	1.44%	N/A	4.67%	4.10%
4/18/2012	14:50:25	0.66	7.73%	5.63%	6.24%	6.35%
5/31/2012	11:07:50	3.00	6.67%	N/A	N/A	4.63%
6/15/2012	12:43:19	0.30	22.78%	4.30%	6.07%	4.68%
7/24/2012	14:22:20	1.20	8.67%	6.30%	6.10%	6.63%
8/29/2012	12:47:36	3.70	13.67%	N/A	N/A	9.45%
9/26/2017	11:12:31	3.40	9.89%	N/A	N/A	8.53%
10/24/2012	12:42:08	3.00	12.22%	8.50%	7.67%	8.00%

Table A-9: 2012 Calibration Reference Concentration Deviations for CAMS 6602 by Reference Concentration Level

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
4/17/2012	15:42:34	0.42	0.56%	1.61%	1.73%	-0.22%
5/25/2017	11:37:46	0.80	9.11%	N/A	N/A	9.60%
6/26/2012	11:06:55	0.80	10.56%	N/A	N/A	9.35%

Date	Time	0 ppb (ppb)	90 ppb (%)	200 ppb (%)	300 ppb (%)	400 ppb (%)
7/30/2012	10:07:35	2.10	9.78%	7.35%	7.53%	6.43%
8/28/2012	10:35:59	2.90	6.22%	N/A	N/A	6.20%
9/25/2012	10:22:59	1.90	1.11%	N/A	N/A	2.85%
10/29/2012	11:31:29	1.00	3.33%	4.00%	5.33%	4.50%

There were a number of instances when deviations from reference concentrations exceeded +/- 7%, meaning that these data would not meet EPA's data quality criteria. This qualification should be considered when interpreting performance evaluation metrics and the modeling data at CAPCOG's monitoring stations for the 2012 model.