4.2 PERFORMANCE EVALUATION OF A YEAR-LONG RUN OF AN AIR QUALITY FORECASTING SYSTEM FOR SOUTHEAST TEXAS

Xiangshang Li *, Daegyun Lee, Soon-Tae Kim, Hyun-Cheol Kim, Fong Ngan, Fang-Yi Cheng, Daewon Byun
Institute for Multidimensional Air Quality Studies, University of Houston, Houston, TX

1. INTRODUCTION

The Institute for Multidimensional Air Quality Studies (IMAQS) of University of Houston (UH) AQF system has been running real-time air quality forecasting systems (AQFS) since 2005. Its F1 forecast system started issuing forecast in May 2005. The center piece of the forecasting systems is the Community Multiscale Air Quality (CMAQ) model (Byun and Ching, 1999). After over a year's test run, IMAQS began extensive performance evaluation of the system, including the meteorology and chemistry components. Since a detailed evaluation is unlikely to be fit in a short paper, here we only summarize the most important findings in ozone evaluation of the model's first year run. Performance evaluation of the forecasting results and analyses of model output help to understand model behavior, detect model deficiencies and identify possible causes of error.

Most current model evaluation efforts focus on certain meteorological episodes and their results generally reflect model behavior during the episodes (e.g., Zhang et al 2006, Smythe at al 2006). The evaluation in this work, however, spans a year and therefore, capable to capture model's seasonal characteristics, and verify some model settings such as the boundary conditions (BCs). It is also more effective to detect model's overall biases as the longer period included more weather modes. Given that model performance are affected many factors, we will not attempt to compare the model performance by numbers.

The one year evaluation period begins June 2005 and ends May 2006. Our focus domain is a 4-km resolution domain covering Southeastern Texas (SETX).

**Corresponding author address*: Xiangshang Li, Univ. of Houston, Institute for Multidimensional Air Quality Studies, 312 Sci. & Res. Bldg., Houston, TX 77204-5007; E-mail: xli@mail.uh.edu

The evaluations are based on "exact match" comparison in which the observation is compared to the model value at the grid cell exactly matching the observation location.

2. MODEL CONFIGURATIONS AND INPUTS

The IMAQS F1 modeling system has three major components: MM5 v3.6.1, SMOKE v2.1, and CMAQ v4.4. The detailed descriptions of modeling system can be found on IMAQS website (http://imaqs.uh.edu/ModelSetup.html). The key model settings are summarized below. For MM5 science options:

- Initialization – ETA
- Subgrid cloud convection – Grell at 36/12 km domain, no scheme at 4 km domain
- Radiation scheme – RRTM
- PBL scheme – UH modified MRF
- LSM scheme – UH modified NOAH LSM

For CMAQ/SMOKE configurations:

- Chemical mechanism – CB4_aq_ae3
- Emission - Texas Emissions Inventory (TEI) for 2000 + NEI99
- Boundary conditions – downscaled linkage from GEOS-CHEM
- Advection scheme – PPM
- Horizontal diffusion – multiscale
- Cloud scheme – RADM
- Forecasting hours – 48 hours (+ 6 hrs spin-up)

The emission scenario for F1 does not consider the emission reduction in SETX in recent years. The HRVOC emissions from industrial sources in the Houston area may have decreased by a factor of two since 2000 (Cowling et al, 2006). However, recent field study (Mellqvist et al, 2007) suggests that the 2004 VOC emission inventory may be too low. Hence using an old emission inventory may be a good compromise.

Both the MM5 and CMAQ modeling domains consist of three nested domains with resolution of 36-km, 12-km and 4-km, respectively. The 12-km and 4-km domains for MM5 are slightly larger than those of CMAQ. The domains are illustrated in Figure 2.1.

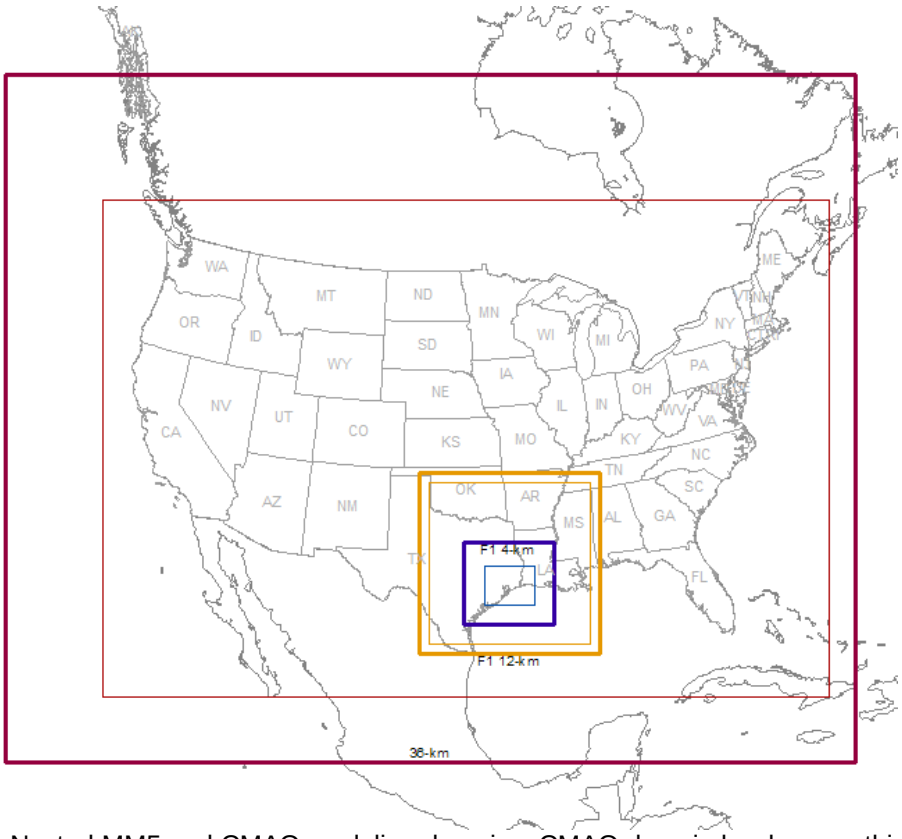

Figure 2.1 Nested MM5 and CMAQ modeling domains. CMAQ domain borders are thin lines.

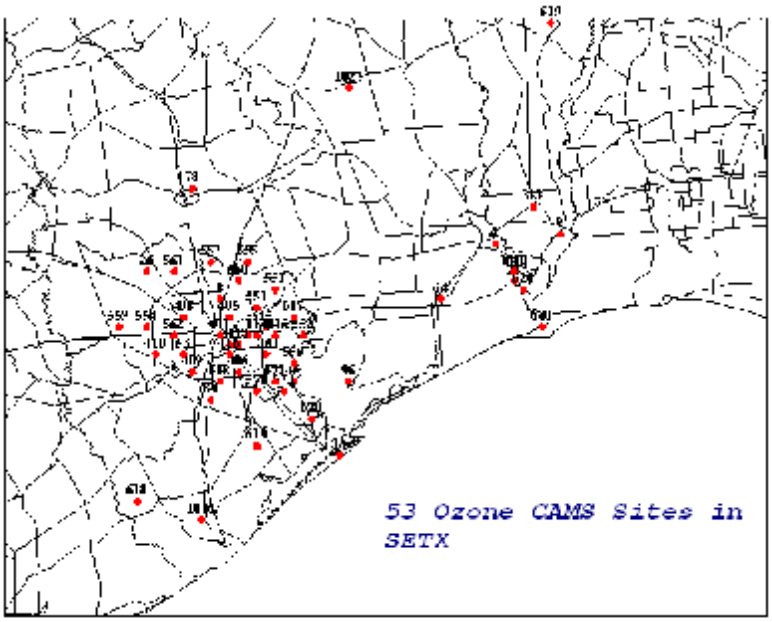

Figure 3.1 53 CAMS sites in the CMAQ 4-km domain (SETX domain), with numbers being CAMS IDs.

3. OBSERVATIONS

We obtained hourly ozone measurements taken by the Continuous Ambient Monitoring Stations (CAMS) network operated by Texas Commission of Environmental Quality (TCEQ). There are 53 CAMS sites (Figure 3.1) reporting ozone data inside the 4-km CMAQ modeling domain during the evaluation period. Most of the CAMS sites are located around Houston and Beaumont. The time-tagging for ozone hourly observation is the beginning of each hour. For example, the 10 AM observation is averaged on measurements taken from 10 to 11 AM.

4. EVALUATION PROTOCOLS

Observed ground level ozone in our modeling domain generally falls in the range of 0 to 200 ppb. The hour to hour changes are usually less than 70 ppb, with most of them less than 30 ppb. Therefore, we believe a simple set of evaluation protocols are sufficient for model evaluation. The set of statistical protocols include: mean bias (MB), mean absolute error (MAE), root mean square error (RMSE), correlation (R), Index of Agreement (IOA), as well as mean and standard deviation (Std Dev or SD) of model and observed data series. All the statistics are commonly used in the modeling community. IOA (Willmott, 1981), popular in the earth science community, is a modified correlation coefficient that accounts for phase errors between model and observed data series. Though there are newly developed statistics such as mean normalized factor bias (MNFB) etc. (Yu et al, 2003), they are not widely adopted and their functionality can be approximated by the simple set of statistics here.

4. EVALUATION RESULTS

The ground level ozone at SETX has a strong seasonal cycle, an evident diurnal variation and noticeable spatial patterns which are driven by the meteorology and the emission distribution.

4.1 *Overall Regional Statistics*

4.1.1 *Ozone of All Available Data Points*

The model performance statistics over the whole region and the whole year can be calculated with all data pairs. Each data pair consists of an hourly ozone observation and its model counterpart. The single point comparison is the most common in episode studies.

Mon	N	R	IOA	RMSE	MAE	MB	O_Mean	M_Mean	O_SD	M_SD
1	37680	0.62	0.78	11.6	8.9	2.3	23.7	26	14.3	11.9
2	33936	0.66	0.8	11.8	9.5	3.9	25.3	29.2	15.1	13.3
3	38688	0.66	0.77	12	11.1	6.9	30.1	37	15.7	12.8
4	37440	0.62	0.75	13.8	12	6.7	31.1	37.7	16.9	14.6
5	39096	0.76	0.84	10.7	10	5.5	35.8	41.2	15.9	14.8
6	35400	0.69	0.79	17.3	14.6	9.3	31.6	40.9	23.3	19.2
7	37296	0.67	0.74	15.4	15.4	11.1	23.1	34.2	20.5	15.6
8	37584	0.66	0.74	17.4	17.5	12.9	27.4	40.3	22.8	18
9	28680	0.77	0.84	14.1	12.8	7.9	27.4	35.3	21.7	18.5
10	34632	0.69	0.8	16	12.8	5.8	31.1	36.9	21.8	17.5
11	36408	0.56	0.74	11.6	9.3	2.3	20.6	22.9	13	11.7
12	37944	0.62	0.78	11.8	8.9	1.3	19.6	20.9	13.9	13
Year	421171	0.68	0.8	14.4	11.9	6.2	26.7	32.9	18.8	16.6

Table 4.1 Statistics for all data pairs. Units for column RMSE to M_SD are ppb.
N – Number of data pairs; R – Correlation; O – Observation; M – Model; SD – Standard Deviation; O_SD - Standard Deviation of observation.

Table 4.1 gives the monthly and the whole year statistics for all the data pairs. It can be seen that the model performance as measured by R and IOA are quite consistent month-by-

month. The overall R and IOA are slightly less than those of past episode studies by IMAQS (Byun et al, 2004). Considering the fact that this study is based on forecast-mode simulations rather than analysis-mode simulations, the performance exceeded our expectation.

The month-by-month trend of model mean ozone is simulated quite well. The monthly mean biases (MB) are quite small during winter and spring and they gradually increase approaching summer months. As shown later, the positive summer time biases are the results of model night-time and early morning overprediction.

The standard deviation of model data series is substantially lower than that of observation. This is likely because the model values are averaged over a 4x4 km area and observations are from a single point. Area averages always respond less than single point measurement.

4.1.2 *Regional Daily Max Ozone*

Regional max ozone is the maximum daily ozone in CMAQ 4-km SETX domain. The observational regional max is the highest measurement value of all the 53 CAMS measurements. There is no extrapolation although it is very likely the actual max ozone may not be recorded by CAMS. The model regional max is the highest value of all model land cells.

The statistics for regional daily max ozone are shown in Table 4.2. The yearly R and IOA are 0.79 and 0.87 respectively, which indicates that the model has good skill in tracking the max ozone.

Mon	N	Corr	IOA	RMSE	MAE	MB	O_Mean	M_Mean	O_SD	M_SD
1	31	0.44	0.6	11.5	8.5	-2.6	51.1	48.5	12.6	6.9
2	28	0.61	0.72	12.3	8.4	-3.8	56.8	53	15.4	10.5
3	31	0.48	0.65	11.6	9	-1.2	65	63.8	13.1	7.9
4	30	0.75	0.82	16.9	13.2	0.7	79.7	80.4	25.3	16.3
5	31	0.47	0.68	20.1	14	0.9	76.9	77.8	21.2	17.4
6	30	0.8	0.87	18.3	13.7	-4.8	96.4	91.5	30.4	22.7
7	31	0.59	0.69	24.3	17.8	-7.4	84.9	77.5	30	17.3
8	31	0.87	0.83	21	16.5	-5.9	102.9	97	35.2	19
9	25	0.81	0.86	15	11.2	-5.1	93.9	88.9	25.4	18.3
10	31	0.8	0.77	20	18.1	-12.6	88.8	76.2	31.4	18.2
11	30	0.14	0.43	19.6	13.7	-8.5	59.2	50.7	18.7	8.7
12	31	0.34	0.47	18.1	12.5	-9.4	52.9	43.5	19	9.1
Year	360	0.79	0.87	17.8	12.6	-4.9	75.5	70.5	29.6	23

Table 4.2 Statistics for regional max ozone. Units for column RMSE to M_SD are ppb.

The mean biases (MB) for regional max are small and negative for most summer months, indicating that the model tends to under-predict max ozone slightly in summer. The -12.6 ppb underprediction in October is likely due to the extra emissions from plant restart after Hurricane Rita. The mean absolute errors (MAE) are in the teens for most of the summer months, showing that the model is doing quite well in capturing the regional max. The observed SD is also markedly higher than the simulated SD, suggesting the observed day-by-day max ozone swings more wildly than the simulated. The wilder swing in observed max ozone may be explained in two ways. First the model values are area averages and secondly, the observed max ozone is often subject to small scale meteorology and other events that may not be captured by the model.

Figure 4.1 shows the yearly time series plot of regional max ozone. The model predictions track observations quite well although there is an evident higher day-to-day variation in observed regional max ozone.

Figure 4.2 displays the monthly mean max ozone which is essentially a smoothed version of Figure 4.1. The simulated monthly mean matches observation very well.

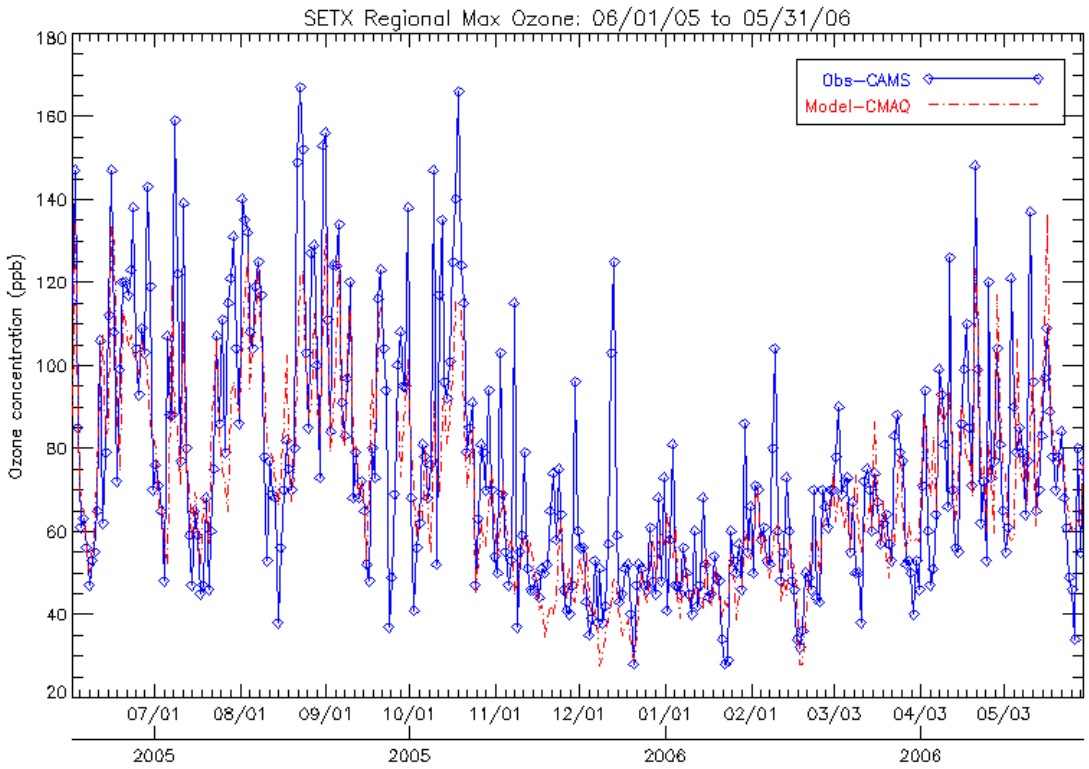

Figure 4.1 Yearly time series of regional max ozone

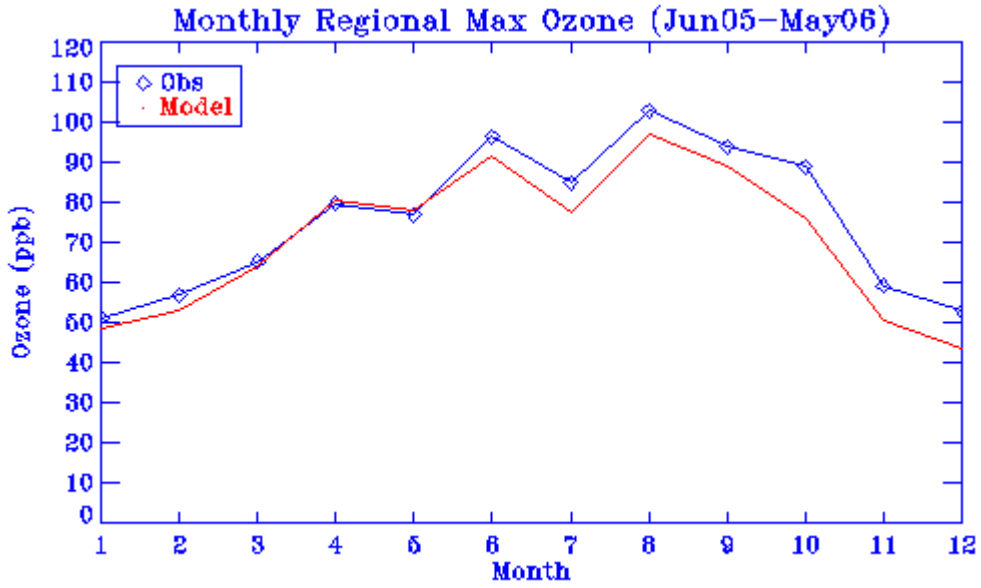

Figure 4.2 – Monthly mean regional max ozone

4.1.3 *Average day-time (8-19 CST) ozone*

Since ozone chemistry has a distinct diurnal cycle, it is useful to separate the day-time and night-time ozone and examine them individually. Average day-time ozone is calculated by averaging the 08-19 CST hourly ozone for each site. So there is one data point each day for each site.

Figure 4.3 shows the monthly mean day-time ozone.

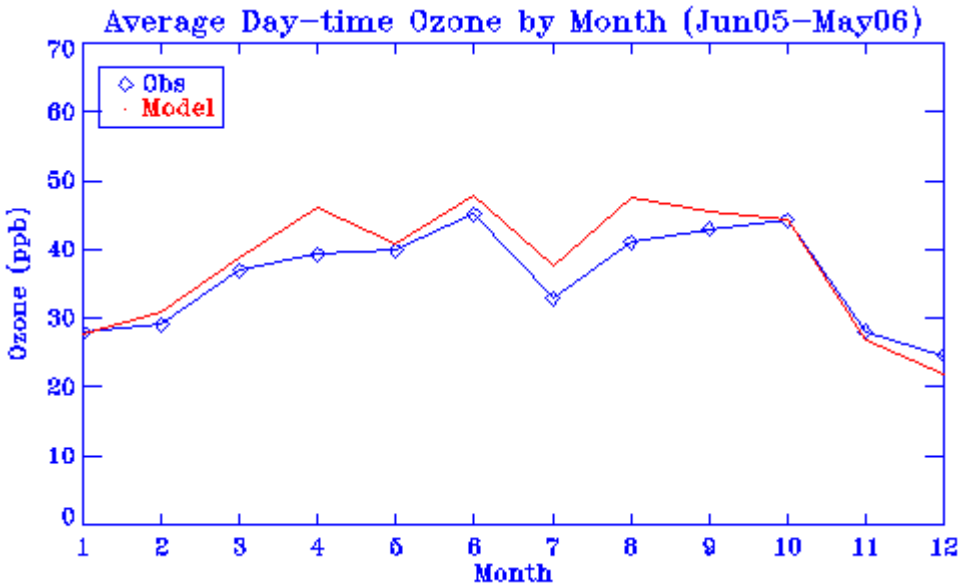

Figure 4.3 – Monthly mean day-time (08-19 CST) ozone

Figure 4.3 is important in showing that the model's day-time ozone simulation seems to be adequate, which in turn suggests that the emission level in the simulation is appropriate. Though there are small model overpredictions in summer months, the seasonal pattern and magnitude match nicely. The small summer overprediction is likely due to a higher model boundary condition. An internal IMAQS study suggests that the model's south boundary ozone may be around 10 ppb higher than the actual.

4.1.4 *Average night-time (0-7, 20-23 CST) ozone*

Figure 4.4 shows the monthly mean night-time ozone. Apparently, there is a significant overprediction for all months, with higher biases in summer months.

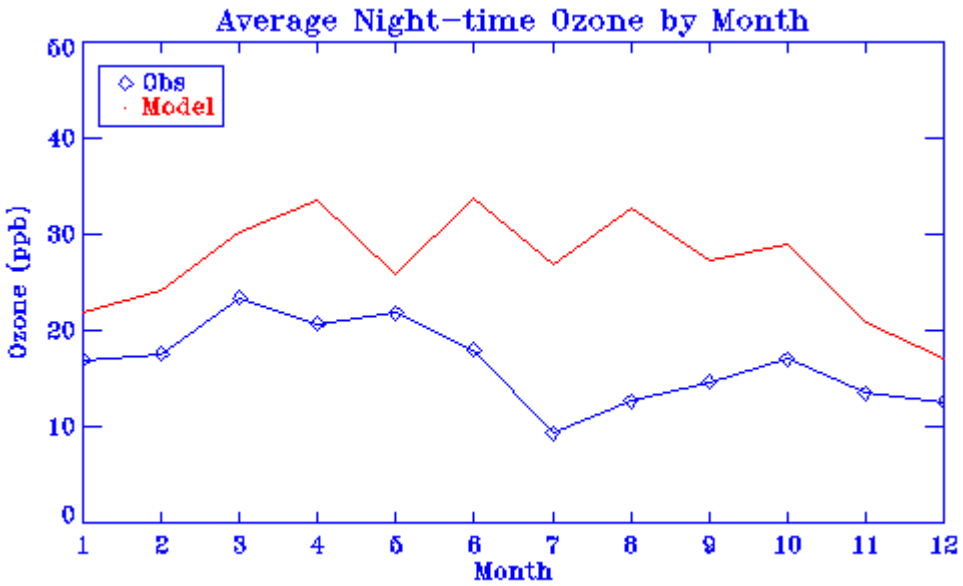

Figure 4.4 – Monthly mean night-time (0-7, 20-23 CST) ozone

There are several possible explanations for the large night-time biases, especially in summer. The night-time ozone chemistry is dominated by the NO titration which is responsible for the observed low night-time ozone. An overprediction usually suggests reduced titration rate. In summer time, the prevailing night-time wind is south wind (S, SSE, SE, EES) which brings in the air from the Gulf. The model predicted wind at land is stronger at night (Harvey, 2005) and it also has higher background ozone from the Gulf of Mexico. The stronger model night winds can have two mechanisms to raise the night-time ozone. First, stronger night-time winds bring in more air with higher background ozone, mixing with local air mass with low ozone. Secondly, the stronger wind can decrease the NO concentration, hence the titration rate. Other likely causes include the height mismatch of observed and modeled first layer ozone and the possible higher model PBL height after sunset.

4.2 *Diurnal Variations*

Figure 4.5 shows the seasonal and yearly diurnal profile of ozone. The seasonal diurnal variation is calculated by averaging all data points of the season at each hour. We present the seasonal rather than monthly variation primarily due to the small monthly variations within each season.

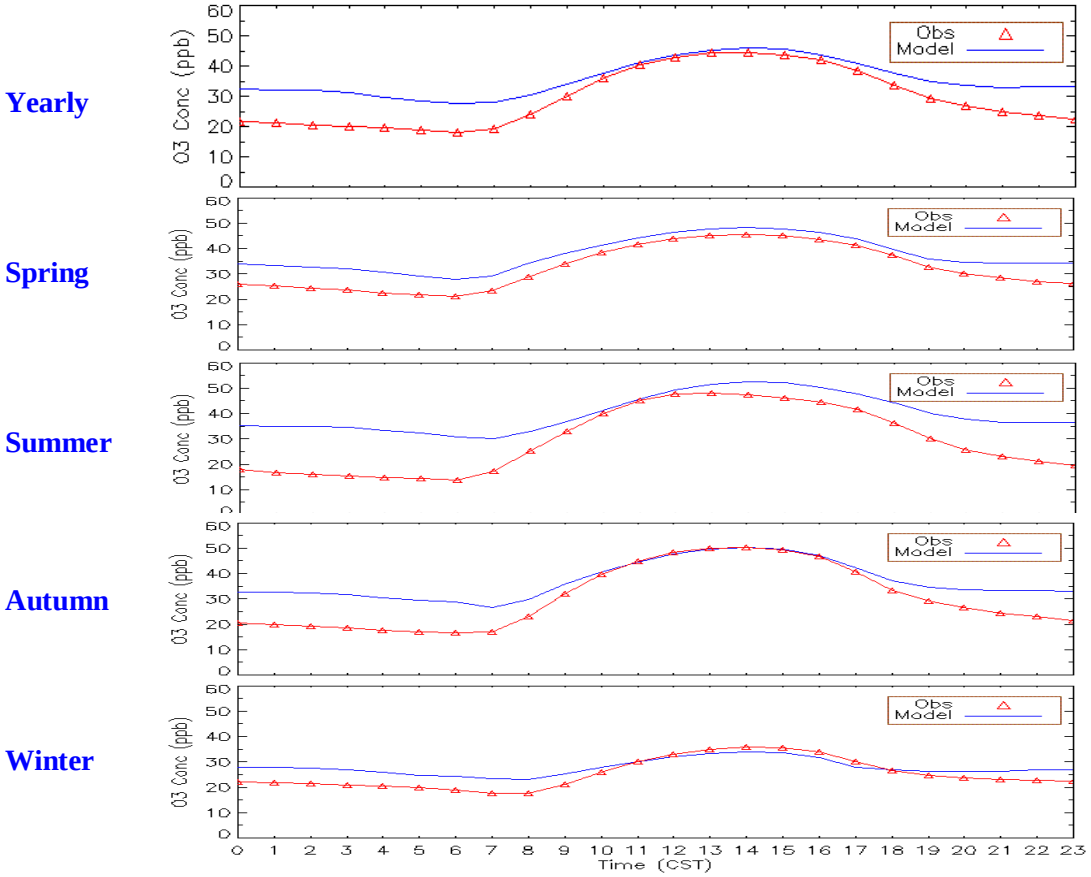

Figure 4.5 – Diurnal Profile (All site average)

The yearly and seasonal diurnal ozone profiles show the almost perfect agreement of model and observation on the peak of the curves, with the best match in Fall (SON). In summer, the observations peak slightly earlier with a slightly lower top. In spring, model slightly overpredicts all day long. At night time, as already mentioned above, the model has a positive bias which is the most pronounced in summer and the least in winter.

In summer, observed ozone rises more rapidly from 6 CST to 10 CST. Since morning wind is not strong and almost all the sites have low ozone at night, the rapid rise in observed

ozone is largely the result of local production, not transport. Plausible causes for the low model morning production are: the stronger model winds hinder the local precursor build-up; the model's mixing height rises faster than observed, leading to lower precursor concentration. More detailed analysis of vertical distribution of ozone and precursors is required to pinpoint the exact cause. Observed ozone also decreases faster than simulated starting from around 17 CST, and the trend continues till midnight. The model prediction only drops in the first few hours after peaking, and the drop gradually levels out around 20 CST. The slower drop of model ozone from 17 CST to 20 CST could be caused by the slower collapse of the simulated PBL. The leveling off of model ozone after 20 CST is likely due to the higher model night-time wind and higher background ozone from the Gulf. In real world, observed night-time wind is low, the transport is much less a factor than titration. Therefore observed ozone at night time can drop further than model ozone. At downtown Houston, ozone often drops to lower single digits, or even zero.

4.3 *Spatial Patterns*

One of the most important spatial patterns is the spatial distribution of monthly mean ozone and the model mean bias. The spatial color contour plots can reveal hidden spatial features that are otherwise difficult to find. To create spatial color contour plots, gridded data must be first generated. The model data are gridded by nature so no extra work is needed. The observed data, however, only have values at each CAMS site. To generate the spatially gridded data from a handful of discrete points, several techniques may be employed. The most common gridding method is Kriging since it is fast and accurate in most occasions. Kriging is a linear interpolator which provides the smoothest spatial transition from one observed point to another. The flat terrain and generally homogenous LULC in the region makes Kriging a good choice for gridding. Although LULC varies in the metro Houston and Beaumont areas, the dense CAMS network in the two areas makes it a minor issue.

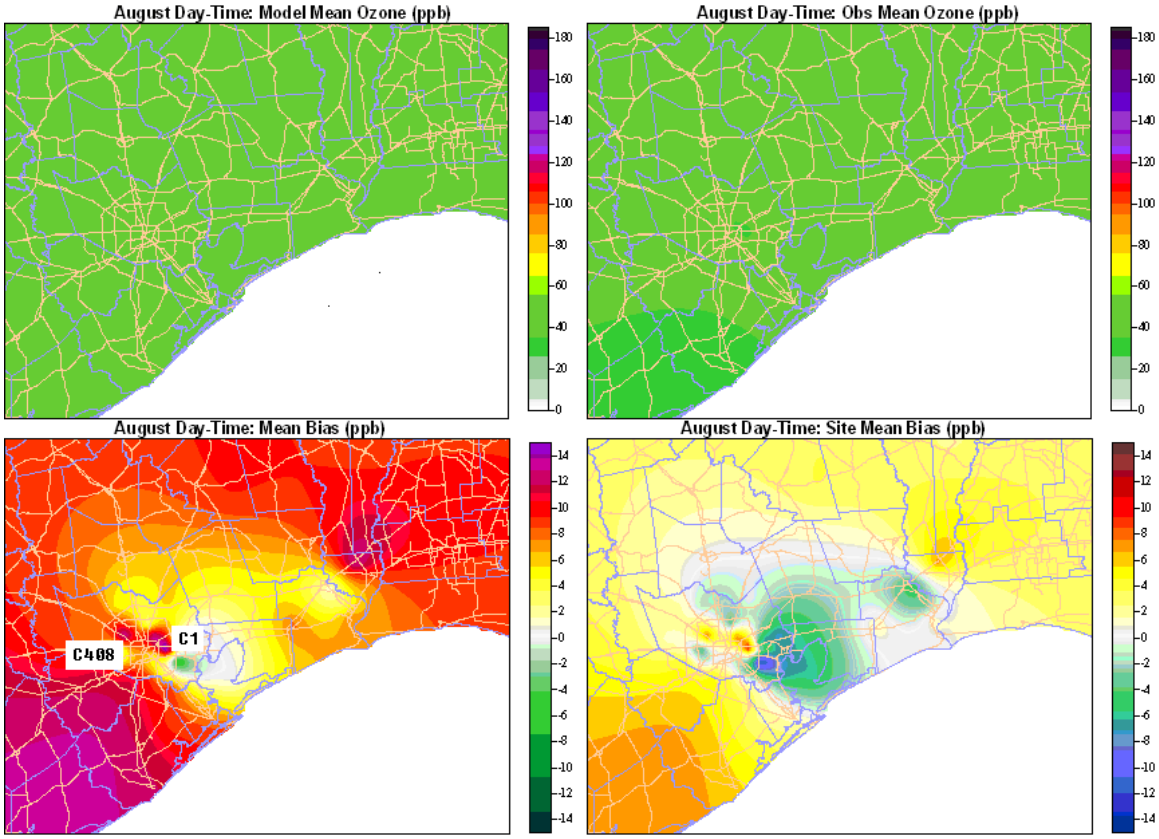

Figure 4.6 – August mean day-time ozone

We have created color contour plots for each month and the whole year, however, only August plots are presented here due to the page constraint. August is chosen because historically it is the month with the highest ozone exceedances.

Figure 4.6 presents the spatial plots of August mean day-time ozone. There are four plots in Fig. 4.6: A) Model mean ozone (Upper-Left) B) Observed mean ozone (Upper-Right) C) Mean bias (model mean - observed mean) (Lower-Left) D) Site-specific mean bias (mean bias – all site average mean bias).

Plot D shows the site-specific bias by subtracting the monthly regional bias from the actual site mean bias. While the monthly regional bias is more related to monthly meteorology, ozone boundary conditions and large scale transport, the site-specific bias can give clues of possible emission inventory problems.

The spatial plots in Fig. 4.6 revealed a number of interest features. In plot C (LL plot), there is a model overprediction spot (shown as purple) at site C408 (Houston, Lang) in NW Houston. A close check shows that the overprediction was caused by an observed ozone drop at the spot. The low observed ozone is probably caused by the NO titration. The site is just off Highway 290 where traffic is very heavy and NO level is high. However, the model did not have the low ozone spot.

Another high positive bias around site C1 has not been carefully studied yet. It may be related to a slight difference in simulated wind direction which veered more westerly in the area.

Also in plot C, there is a large red and purple area at lower left which represents coastal Brazoria and Galveston. The positive bias is likely the result of the model's high background ozone and the August prevailing south and southeasterly winds. As for the positive bias at rural sites C311, C1027 (plot C, purple spot at upper right), currently there is no concrete explanation yet.

In plot D, there is a blue area which represents model underprediction just below the C1 orange area. The underprediction at the area (site C35) may suggest possible VOC emission inventory problem.

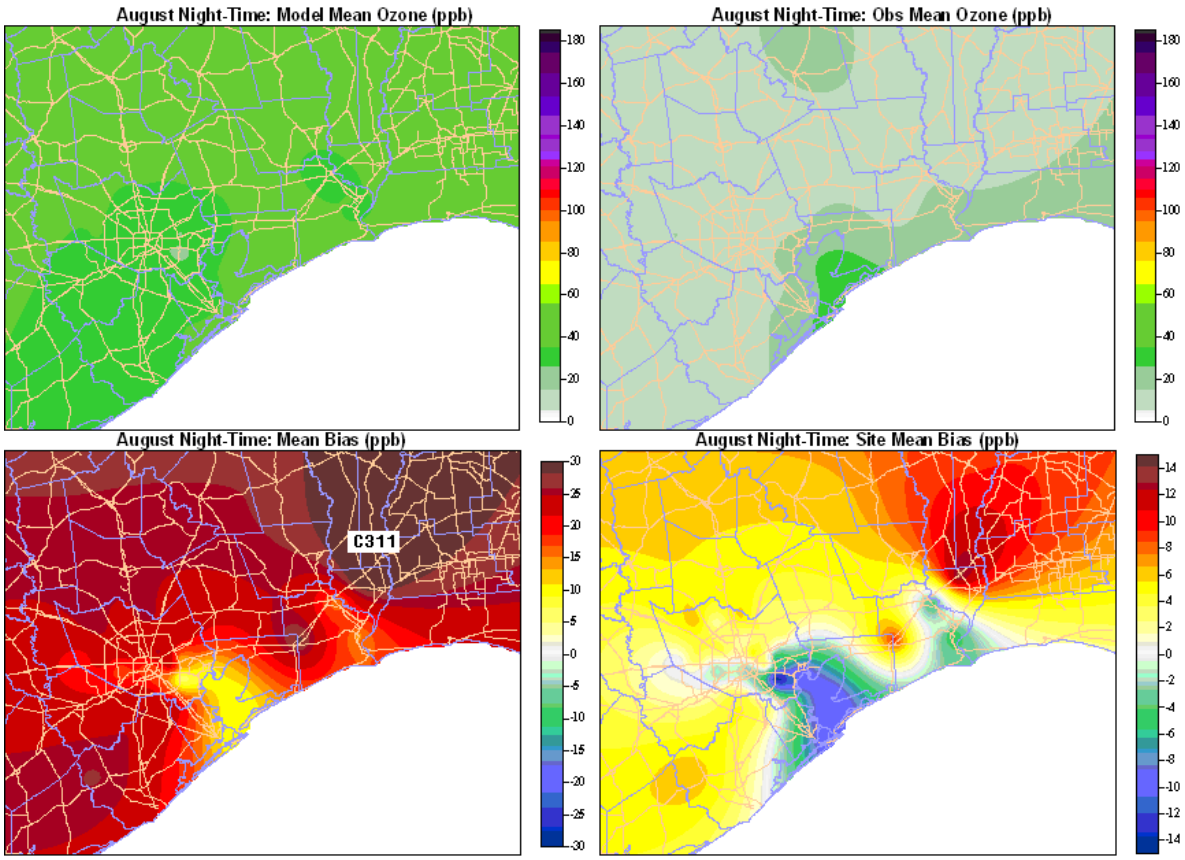

Figure 4.7 – August mean night-time ozone

Figure 4.7 presents the spatial plots of August mean night-time ozone. The 4 plots in Fig. 4.7 are similar to those in Fig 4.6 except that they are for night-time.

The night-time ozone biases are large in most areas, reaching 30 ppb at certain places, such as site C311. The high bias at C311 is characterized by an observed low ozone area. The model prediction at C311 is not much different from surroundings – suggesting the bias is more related to the observed low instead of the predicted high. Again, further study is needed to pinpoint the cause for the observed low ozone.

Another interesting feature is the lower bias at coastal sites – this may be explained by the fact that the stronger model winds can penetrate deeper inland, making the ozone more uniform at inland sites. Also the higher winds likely decrease the titration rate and keep the ozone from falling too low. In reality, however, the winds are weaker inland and titration is a much more important process than transport. Only a few sites that are very close to the coastline are subject to the transport. The night-time transport brings in background ozone from the Gulf and raises the ozone level at these sites. As a result, the biases at coastal sites are lower than inland sites.

5. DISCUSSION AND CONCLUSION

Our evaluation of the MM5-SMOKE-CMAQ modeling system shows that the system is capable to forecast the day-to-day ozone variation in the 4-km SETX domain. The overall forecasting performance is only slightly behind our analysis-mode performance. Overall the model has a positive bias in simulating the regional ozone for most of the year. The positive bias is due to the night-time and early morning overprediction.

For regional daily max ozone, however, model shows slight underprediction which is probably due to the wilder swings in observed max ozone. It is hard to improve the model performance in simulating the max ozone because the small-scale meteorology and other events which contribute to the wilder swing are not reflected in the model.

The model shows better performance in predicting temporally averaged ozone such as monthly average than individual daily ozone as the average smoothes out the instant fluctuations.

The seasonal cycle is well captured by the model. The observed double peak in late spring and fall, as well as the dip in July are present in the model. The forecast monthly mean max ozone as well as day-time average ozone matches the observation quite well, indicating the emission inventory is adequate.

The forecasted diurnal cycle generally agrees with observation, especially during the daytime. However, the night-time ozone is overpredicted in the model. The large night-time positive bias in summer may be caused by a combination of factors, such as the high night-time winds and high background ozone in the model.

While not elaborated in the evaluation section, our analyses show that the ozone forecasting performance is closely associated with the performance of forecasted meteorology. Most of the model errors can be traced back to the imperfect meteorology from MM5.

The evaluation also reveals a number of interesting spatial features. The August mean bias plot showed a day-time overprediction at C408 which is a heavy traffic site, and at C1 which is near ship channel. The August night-time mean bias plot shows that the model tends to have smaller biases at coastal sites.

6. REFERENCES

Byun, D.W., and J. K. S. Ching, 1999: Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *EPA/600/R-99/030*, U.S. EPA

Byun, D.W., S.T. Kim and S.B. Kim, 2004: Evaluation of CMAQ Results for the Simulation of a High Ozone Episode in the Houston-Galveston Metropolitan Airshed. *2004 Models-3 Conference*, October 18-20, Chapel Hill, NC

Cowling, E.B., C. Furiness, B. Dimitriades, D. Parrish et al, 2006: Preliminary Findings from the Second Texas Air Quality Study (TexAQS II), A Report to the Texas Commission on Environmental Quality, Section C1, *http://www.tceq.state.tx.us/assets/public/implementation/air/am/workshop/20061012-13/RSST_Preliminary_Findings_Report_20061031.pdf*.

Jeffries, H. E., 2005: Analysis of HGMCR 2000 SIP episode modeling with CAMx and CMAQ. Presented in IMAQS weekly seminar.

Mellqvist, J., J. Samuelsson, et al. 2007: Measurements of industrial emissions of VOCs, NH3, NO2 and SO2 in Texas using the Solar Occultation Flux method and mobile DOAS. *2007 AGU Fall Meeting*, December 10–14, San Francisco, CA, USA

Smyth, S. C., W. Jiang, D. Yin, H. Roth, and E. Giroux, 2006: Evaluation of CMAQ O3 and PM2.5 performance using Pacific 2001 measurement data. *Atmospheric Environment*, 40**,** 2735-2749.

Willmott, C. J., 1981: On the validation of models, *Phys. Geog.*, 2, 184–194.

Yu, S., B. Eder, R. Dennie, S.-H. Chu, and S. Schwartz, 2003: New unbiased symmetric metrics for evaluation for the air quality model, *the 2nd Annual CMAS Models-3 User's Conference*, Research Triangle Park, NC, October 27-29

Zhang, Y., P. Liu, A. Queen, C. Misenis, B. Pun, C. Seigneur, and S.-Y. Wu, 2006: A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode-Part II: Gas and aerosol predictions. *Atmospheric Environment*, 40, 4839-4855.