3.6 BASE CASE MODEL EVALUATION PERFORMANCE EVALUATION

Model performance evaluation is a complex process, involving a variety of statistical and graphical methods. Thorough assessment of the model's behavior requires comparing model output to thousands of measurements of ozone, ozone precursors, and reaction products. Because the modeling episode was chosen to coincide with TexAQS 2000, the observational database is the most comprehensive ever used in this process. Aircraft and surface sites provided a variety of aerometric data rarely available. This data set thus provides an opportunity to study not only the model's ability to simulate atmospheric concentrations of ozone and its basic precursors, but also to evaluate its internal processes.

Discussion of much of the detailed performance evaluation conducted for the Phase 2 MCR modeling analysis are beyond the scope of this chapter. Here, only a brief summary of model performance is presented including commonly referenced statistical measures. Detailed information is provided in Appendix B, the "Phase 2 HGB Mid Course Review Base Case Model Performance Evaluation."

3.6.1 The Hybrid Base Case

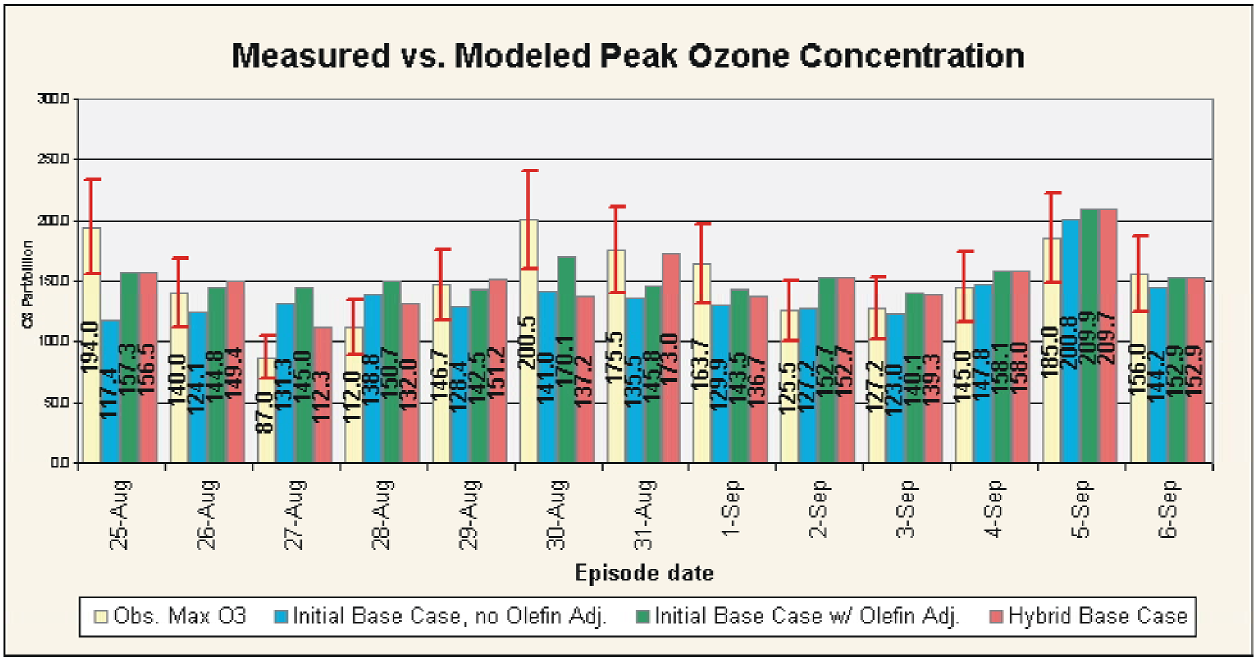
As described in Appendix B, the hybrid base case combines two meteorological characterizations; the original meteorology developed by ENVIRON/ATMET for the extended base case and a newly-developed characterization provided by Dr. Nielsen-Gammon of TAMU which incorporates observations from the GOES. The GOES-based meteorology is currently only available for the period of the original TexAQS2000 episode (August 22 - September 1, 2000), but provided better model performance during that period. The ENVIRON/ATMET characterization provides acceptable performance for the period September 2-6 (except for September 5), so is used for these days. However, the model under-predicted ozone during the first part of the extended episode (August 18-21), so those days were excluded pending development of improved meteorological data for those days. As a result, the hybrid base case begins on August 22 with the GOES-based meteorology and the ENVIRON/ATMET characterization is used from September 2 through 6.

Because the model consistently under-predicts peak ozone using the reported emissions inventory, the hybrid base case includes the terminal olefins-to-NO_x adjustment described in Section 3.5. Additional discussion of the terminal olefins-to-NO_x adjustment can be found in the December 2002 SIP Revision Technical Support Document.

3.6.2 Statistical Performance Evaluation

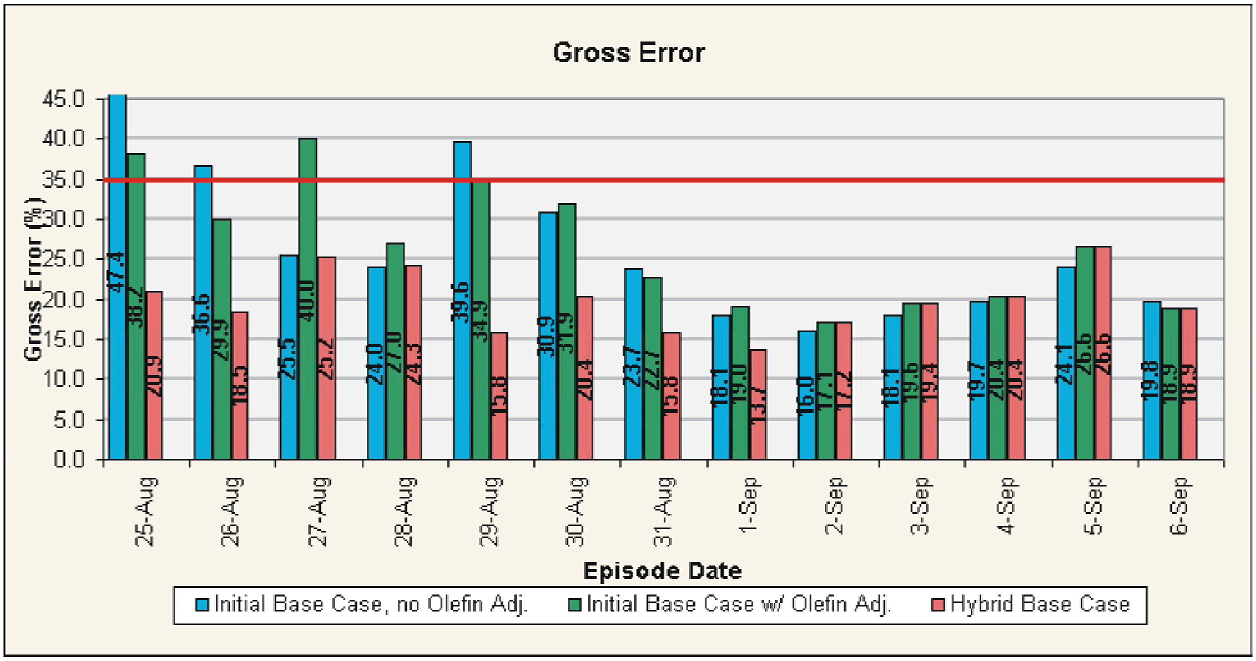
The most commonly used measures of model performance are the three statistics recommended in EPA's 1992 Modeling Guidance: Unpaired Peak Accuracy, Relative Bias, and Relative Gross Error. These statistics are calculated separately for each episode day. The first statistic, Unpaired Peak Accuracy, measures the model's ability to replicate the highest ozone observed on each day of the episode. A negative value of Unpaired Peak Accuracy is a clear indication that the model is not generating a high enough peak, but a positive value does not necessarily signal that the model is generating a peak that is too high, since the actual peak concentration generally does not occur at a monitoring location. Figure 3.6-1, *Measured and Modeled Peak Ozone for Three Base Case Configurations*, compares measured and modeled peak ozone concentrations for three configurations of the base case; the initial (ENVIRON/ATMET meteorology) base case, with no adjustment to industrial olefins, the initial base case with the terminal olefins-to-NO_x adjustment, and finally the hybrid base case, including the terminal olefins-to-NO_x adjustment. Although the Unpaired Peak Accuracy statistic itself is not shown, the red error bars represent the EPA's recommended ± 20 percent tolerance. Peaks that lie within the range shown meet this performance criterion.

Figure 3.6-1: Measured and Modeled Peak Ozone for Three Base Case Configurations



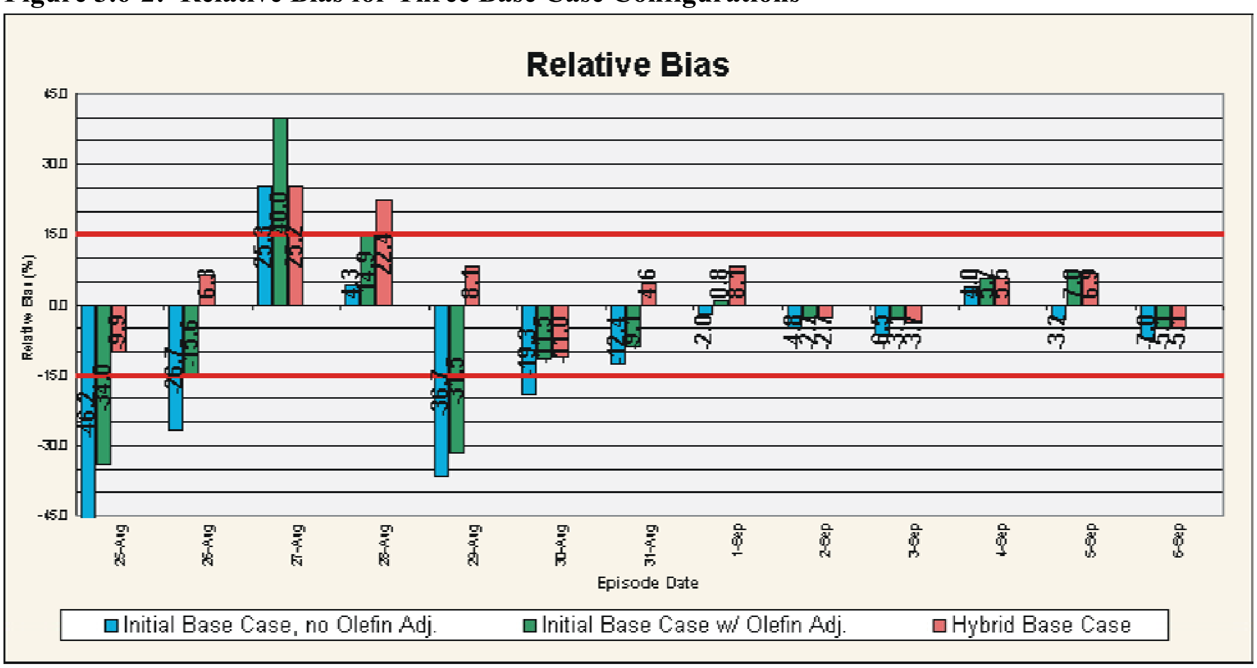
Relative Bias measures the model's ability to generate enough ozone across the monitoring network (positive bias = too much ozone, negative bias = too little). Figure 3.6-2, *Relative Bias for Three Base Case Configurations*, shows the Relative Bias for each episode day. The red lines represent EPA's recommended tolerance of ±15 percent. Relative Gross Error provides a measure of how closely the modeled ozone concentrations match the observations overall. This statistic is presented in Figure 3.6-3, *Gross Error for Three Base Case Configurations*. The red line represents EPA's recommended tolerance of 35 percent.

Figure 3.6-3: Gross Error for Three Base Case Configurations



Comparing the model performance statistics for the three base case configurations confirms that the hybrid base case delivers the best statistical performance. Extensive analysis of graphical performance measures and precursor analysis agrees that the hybrid base case is the most appropriate available modeling configuration for use in the current SIP revision.

Figure 3.6-2: Relative Bias for Three Base Case Configurations



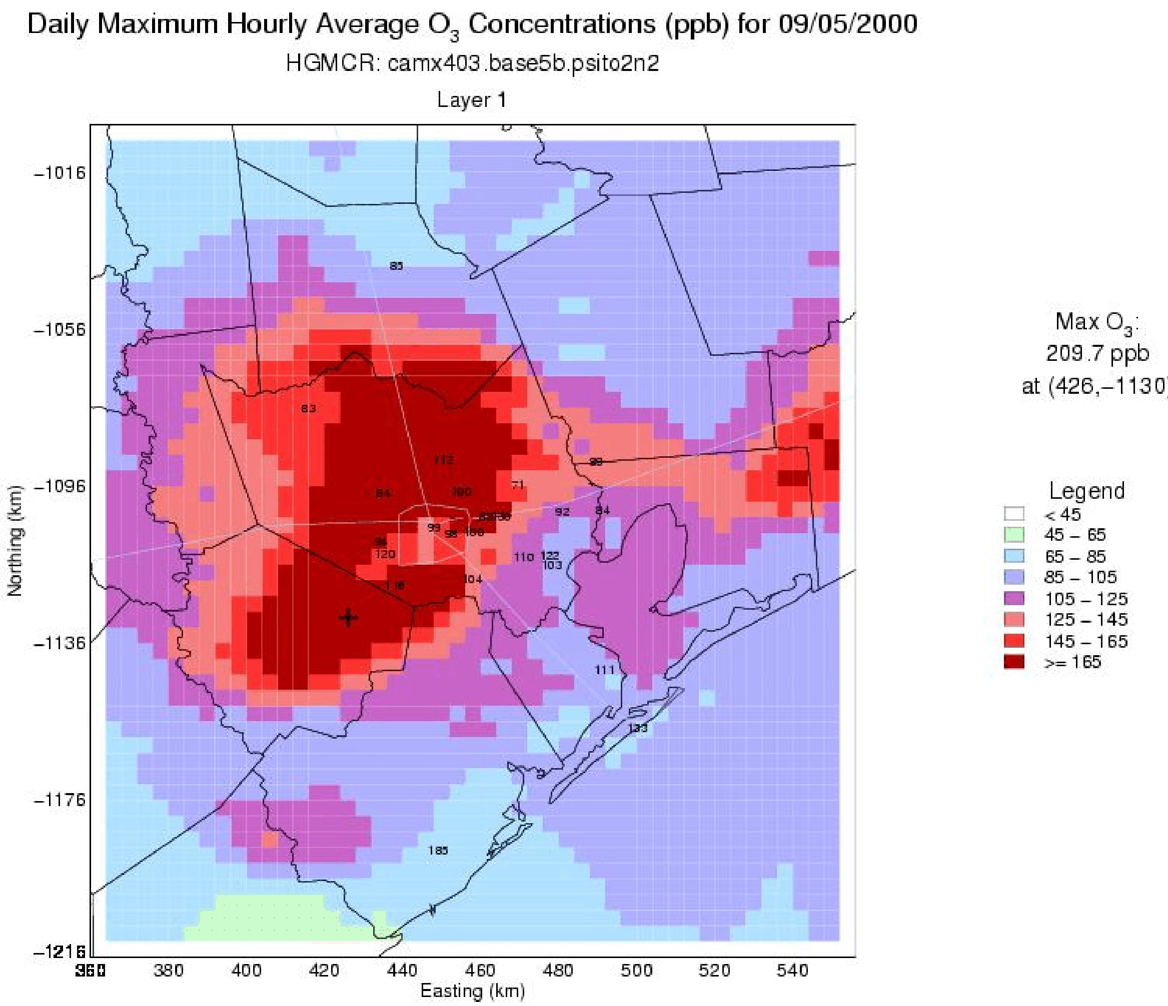
Based on the statistical performance evaluation, the model delivers acceptable performance on all days of the hybrid base case except August 27 and 28. Although the performance is otherwise acceptable on August 30, modeled peak ozone is too low. Evidence suggests that one or more emission events not accurately represented in the modeling inventory occurred on this day, which could explain the high peak ozone recorded. For this reason, August 30 is included in the modeling analysis, despite its inability to replicate the very high ozone value recorded on that day. The accompanying detailed model performance evaluation documentation (Appendix B) includes a discussion of a sensitivity analysis wherein hypothesized upsets were included on August 30.

3.6.3 Graphical Performance Evaluation

Statistical model performance evaluation is the first step in assessing the model's suitability for control strategy development. It is a useful screening step, but much more insight into how the model is working can be gleaned through a variety of graphical analysis techniques. Graphical assessments include time series, scatter plots, isopleth plots, and animations of ozone and precursors. The bulk of the discussion of graphical analysis is provided in Appendix B, but enough is included in this section to illustrate why September 5 was excluded from further consideration at this time, despite acceptable statistical performance.

Figure 3.6-4, *Daily Peak Ozone Concentrations on September 5*, is an isopleth plot showing daily maximum ozone concentrations simulated on September 5, along with monitored daily peak concentrations. The model simulated very high ozone concentrations across the urban areas, where only moderate concentrations were measured, yet simulated low ozone concentrations in the area where the observed maximum occurred.

Figure 3.6-4: Daily Peak Ozone Concentrations on September 5

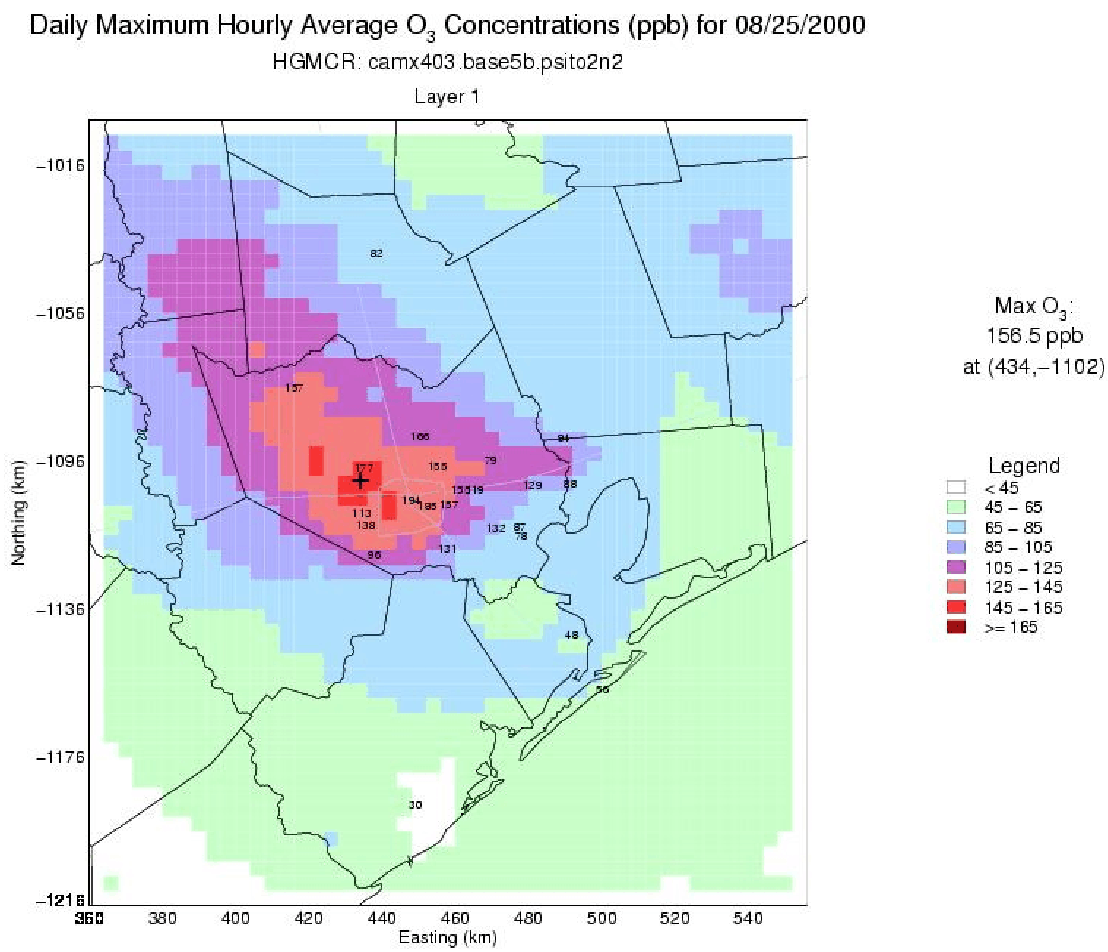


The poor model performance on September 5 is due to a convergence zone along the coast, causing air to be transported from the urban and industrial areas into southern Brazoria County. The stagnant conditions resulting from the convergence then allowed very high levels of ozone to build up. However, MM5 placed the convergence zone too far inland which resulted in the model predicting the peak ozone concentration in Fort Bend County. Similar analysis for the BPA region shows the model significantly over-predicts ozone concentrations in the area.

This analysis illustrates that performance statistics alone are not sufficient to assess model performance. The actual performance of September 5 is so poor that it has not been used for control strategy evaluation. Had the peak been merely displaced downwind of the source region, the day may have been useful for control strategy purposes. However, the peak is errantly located across the urban region and it would be impossible to accurately assess the response to emission reductions.

Figure 3.6-5, *Daily Peak Ozone Concentrations on August 25*, on the other hand, shows the modeled daily peak ozone concentrations on August 25, a day where the model replicated the observed ozone concentrations very well. The pattern seen arises from the ozone moving from its origin near the Ship Channel across town in a westerly, then northwesterly direction.

Figure 3.6-5: Daily Peak Ozone Concentrations on August 25

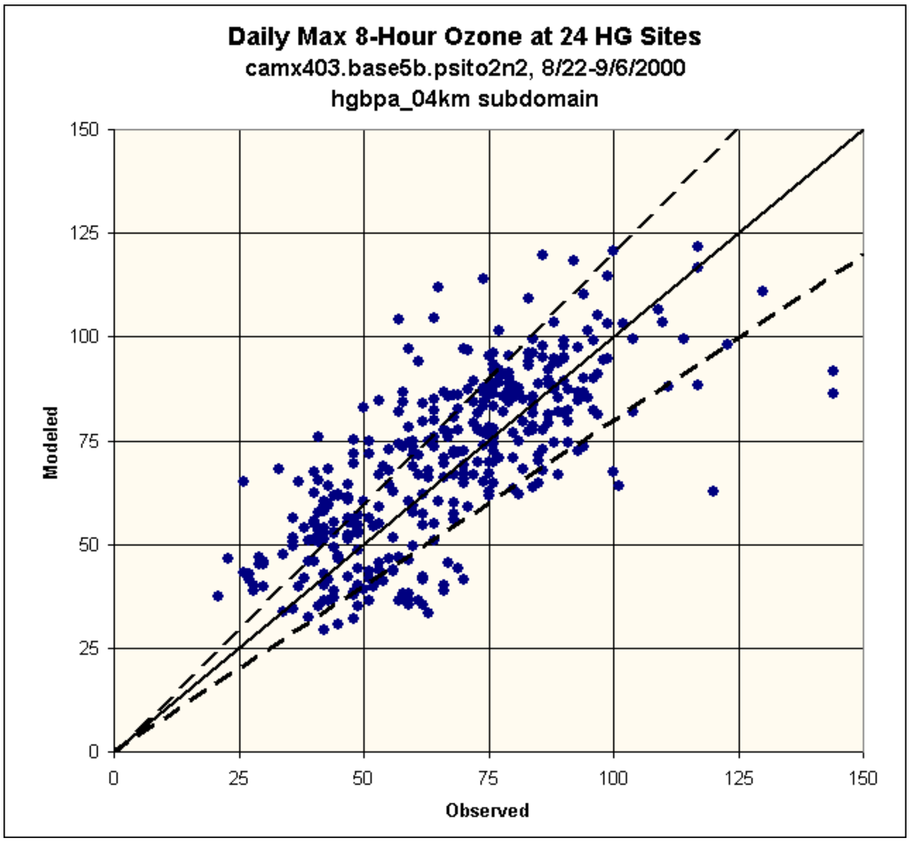


3.6.4 8-Hour Performance Metrics

A set of analyses targeting the evaluation of model performance in replicating 8-hour peak ozone concentrations, as described in EPA's 1999 Draft Guidance for 8-hour ozone attainment demonstrations, was conducted. Specifically, the relative bias was calculated for the entire set of daily 8-hour ozone peaks, as well as on a day-by-day and monitor-by-monitor basis. Scatter plots comparing the measured and modeled 8-hour peaks for the same sets of data were also created. One significant departure from the draft modeling guidance, however, is that instead of using the highest modeled peak value "near" each modeling site, the values at the actual monitoring locations were used. The reason for this deviation is that the majority of ozone monitors in the Houston area are located in a relatively small area encompassing the Houston Ship Channel. Within this area, local NO_x sources can greatly affect ozone concentrations, creating strong concentration gradients in the space of a few grid cells. Using the values at the monitoring locations thus provides a better assessment of how accurately the model actually replicates ozone concentrations than using the peak values "near" each monitoring location would.

Figure 3.6-6, *Modeled vs. Measured 8-Hour Peak Ozone Concentrations*, below shows the overall comparison of measured and modeled 8-hour ozone peaks. The solid line on the graph is the ideal; if the model perfectly replicated observed 8-hour peaks, all the data points would fall on this line. The dashed lines indicate a 20 percent departure from a perfect fit.

Figure 3.6-6: Modeled vs. Measured 8-Hour Peak Ozone Concentrations



While Figure 3.6-6 shows reasonable agreement between measurements and modeled concentrations, clearly there is considerable scatter in the data. Overall the model shows a positive bias of 3.23 ppb, which is less than 5 percent of the mean 8-hour ozone peaks across all monitors and days.

3.6.5 Summary Discussion of Model Performance

The hybrid base case provides generally acceptable model performance for August 25-September 6, with the exception of two days with low measured ozone concentrations, August 27 and 28. Despite acceptable statistical performance on September 5, graphical analysis shows that the model performs poorly on that day and it cannot be used reliably in control strategy evaluation. The 8-hour predicted ozone peaks show very little bias, but there is a fairly large amount of scatter in the data.

Continuing efforts to improve both the meteorological characterization and the modeling inventory are expected to provide improved model performance in future work.

3.7 FUTURE CASE MODELING FOR THE 1-HOUR OZONE STANDARD

To develop and evaluate control strategies designed to bring the HGB area into attainment of the 1-hour ozone standard by 2007, an emissions inventory must be developed for the attainment year. This inventory includes both projected growth and also planned controls. If the modeling does not show attainment, then additional controls are added to the model until attainment is demonstrated. The development of the future modeling inventories for point, area/nonroad, and onroad mobile sources is discussed in earlier sections of this chapter. Biogenic emissions were unchanged in the future inventory.

A "future base case" (i.e., an inventory that is grown to 2007 but not controlled) was not modeled. Instead the final control strategy modeled in the December 2002 SIP revision was used as a starting point. In the 2002 SIP revision, the commission adopted rules that replaced the previously-adopted 90 percent point source NO_x reduction with an 80 percent point source NO_x reduction coupled with reductions of four HRVOCs: ethylene, propylene, 1,3-butadiene, and butenes. Some minor modifications were made to this set of controls resulting in the future case called Control Strategy 3, or **CS-03**[1]**.** The specific modifications from the 2002 SIP revision are:

- Removed Chambers, Liberty, and Waller Counties from the vehicle I/M program
- Removed Commercial Lawn & Garden Restriction
- Removed Heavy-Duty Diesel Idling Restriction
- Removed 55 MPH speed limit (retains 65 MPH on appropriate roads)
- Included Portable Fuel Container Rule

A summary of VOC and NO_x emissions for CS-03 (as well as several other strategies) is presented below in Table 3.7-1, *Anthropogenic Emissions Summary for August 30, 8-County Area*.

3.7.1 Future Case Modeling Results

When Control Strategy 3 was modeled, it became apparent that some additional reductions to emissions would be necessary to develop an attainment demonstration. A series of runs, described below, were conducted to refine the final control strategy used in this attainment demonstration. The daily peak modeled ozone concentrations for each these model runs are shown in Table 3.7-2, *Modeled 2007 Peak Ozone Concentrations for Several Strategies.*

CS-05 - Control Strategy 5 was a sensitivity analysis in which point source VOC emissions in the 8-county area were reduced by 25 percent across-the-board. Results of this modeling run showed concentrations below or slightly above the 1-hour standard for all days except August 31 (see the Weight-of-Evidence section for a detailed discussion of August 31).

CS-06 - Control Strategy 6 replaced the across-the-board VOC reduction of CS-05 with a reduction to Harris County HRVOC emissions only. The level of HRVOC emission reductions was calculated to remove the same amount of reactivity from Harris County as was removed in

[1]Earlier control strategy runs CS-01 and CS-02 are not reported because they did not use the corrected meteorological characterization introduced with Base 5b.

the previous across-the-board VOC reduction sensitivity. Reactivity was calculated using Carter's Maximum Incremental Reactivity (MIR) scale (Carter, 2000).

In addition, 6 tpd of NO_x emissions were removed from the nonroad sources to be consistent with the goals set forth by the TERP. Because CS-06 provided modeled peak ozone concentrations that were only marginally over the 1-hour ozone standard and attainment can be demonstrated using weight-of-evidence as discussed in the next section, Control Strategy 6 was considered to be the "attainment" strategy. Table 3.7-2 shows peak modeled ozone concentrations for each episode day for CS-06 as well as the other strategies discussed here.

CS-07 - Control Strategy 7 is identical to CS-06 except that HRVOC reductions were taken in all 8 counties, for a total reduction of 33 tpd. This strategy offered only minimal benefits over CS-06.

CS-06a - Once the attainment strategy was defined, the maximum allowable credit for VMEP could be calculated and compared to the previously modeled VMEP reduction of 23 tpd of NO_x. The new maximum VMEP figure is considerably smaller than the previously assumed value and is now only 12.9 tpd.

However, at this time only 7 tpd of quantifiable VMEP reductions have been identified, 3.6 tpd from onroad sources and 3.4 tpd from nonroad sources. This tonnage is modeled in CS-06a as opposed to the maximum allowable credit for VMEP.

Table 3.7-1 summarizes the anthropogenic emissions modeled in the 8-county area for the Base Case and several reduction scenarios.

Table 3.7-1: Anthropogenic Emissions Summary for August 30, 8-County Area

Emission Category	NO_x (tpd)						VOC (tpd)					
	Base 5B	CS-03	CS-05	CS-06	CS-07	CS-06a	Base 5b	CS-03	CS-05	CS-06	CS-07	CS-06a
Onroad Mobile	342	169	169	169	169	175	151	90	90	90	90	89
Point Sources	492	174	174	174	174	174	384	275	206	255	244	255
Area/Nonroad	184	146	146	146	146	155	254	234	234	234	234	234
Total	1018	489	489	489	489	504	789	599	530	579	568	578

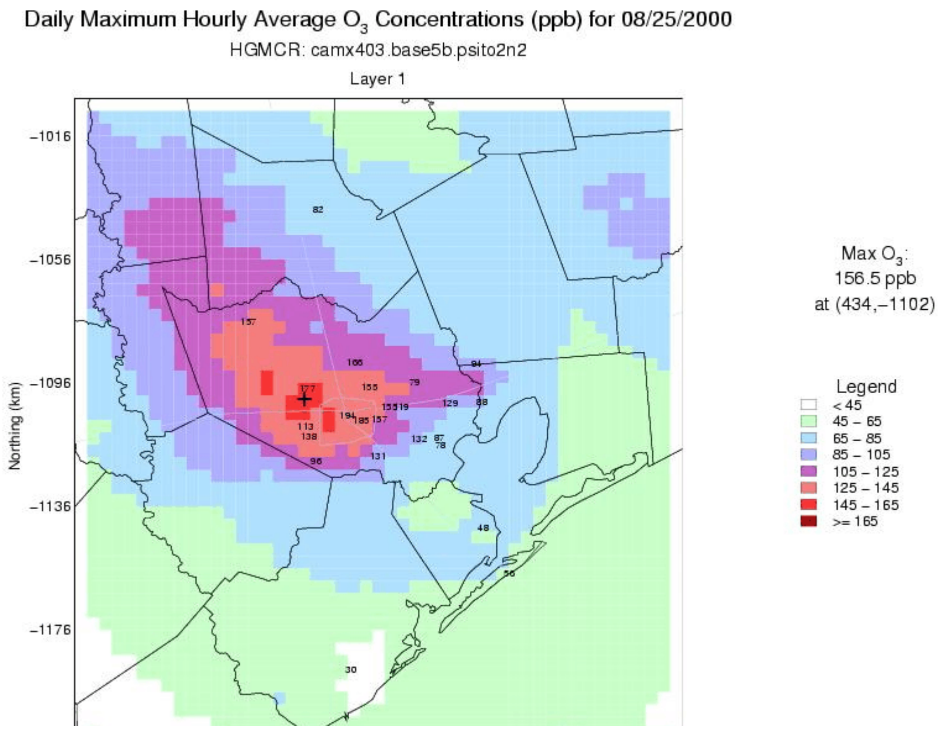
Table 3.7-2 summarizes peak 2007 ozone concentrations modeled on each day (excluding the two low-ozone days August 27 and 28, and September 5 which had unacceptable performance) for several control strategies. Base case peak ozone concentrations are shown for comparison.

Table 3.7-2: Modeled 2007 Peak Ozone Concentrations for Several Strategies

Episode Day	Modeled Peak Ozone Concentration (parts/billion)					
	Base 5b	CS-03	CS-05	CS-06	CS-07	CS-06a
August 25	156.5	134.1	123.9	123.2	123.1	123.4
August 26	149.4	116.9	112.6	113.6	113.5	114.7
August 29	151.2	120.2	115.7	114.4	114.3	114.9
August 30	137.2	128.8	123.9	123.7	123.7	123.4
August 31	173.0	152.1	149.0	148.2	148.1	148.6
September 1	136.7	123.7	120.6	119.9	119.9	120.3
September 2	152.7	133.0	129.6	129.7	129.7	130.2
September 3	139.3	119.1	116.5	115.5	115.5	116.0
September 4	158.0	127.9	126.6	126.5	126.5	127.3
September 6	152.9	129.6	129.0	126.6	126.6	127.6

Table 3.7-2 shows that CS-06a achieved large reductions in peak ozone from the base case on every day, although peak ozone concentrations on 4 days were above the 1-hour ozone NAAQS of 125 ppb. The results in Table 3.7-2 can be easily visualized by comparing peak ozone isopleth plots of the base and future control cases. Figures 3.7-2 through 3.7-6 show daily peak modeled ozone concentrations on all days with future predicted concentrations over 123 ppb. Note that the plots show the model file names, not the control strategy names - “fy07j.CS03_harCap” is the file name for CS06a. A complete set of base case peak ozone isopleths can be found in Appendix H, “Peak Ozone Isopleth Plots for August 25 through September 6.”

Figure 3.7-1: Peak Ozone Isopleths for Base & Future Control Cases for August 25



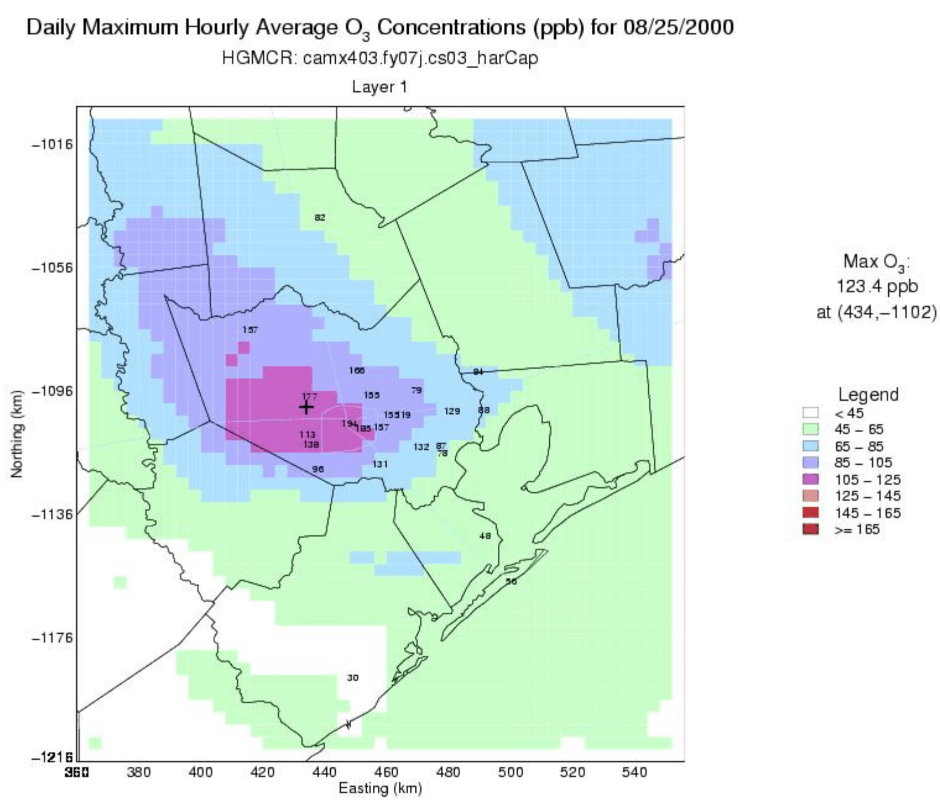
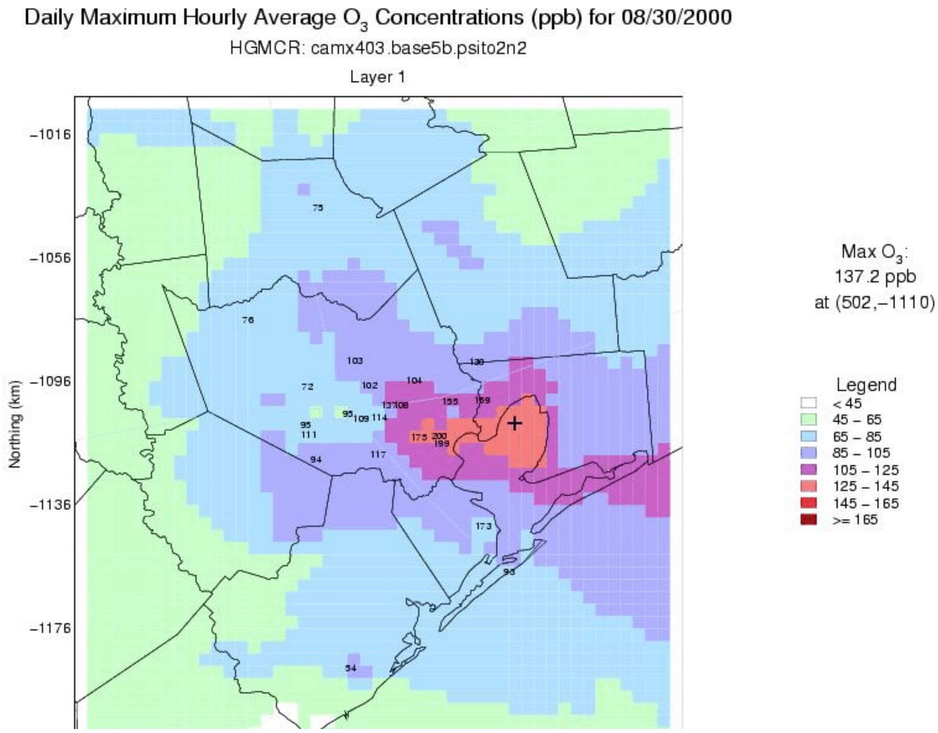


Figure 3.7-2: Peak Ozone Isopleths for Base & Future Control Case for August 30



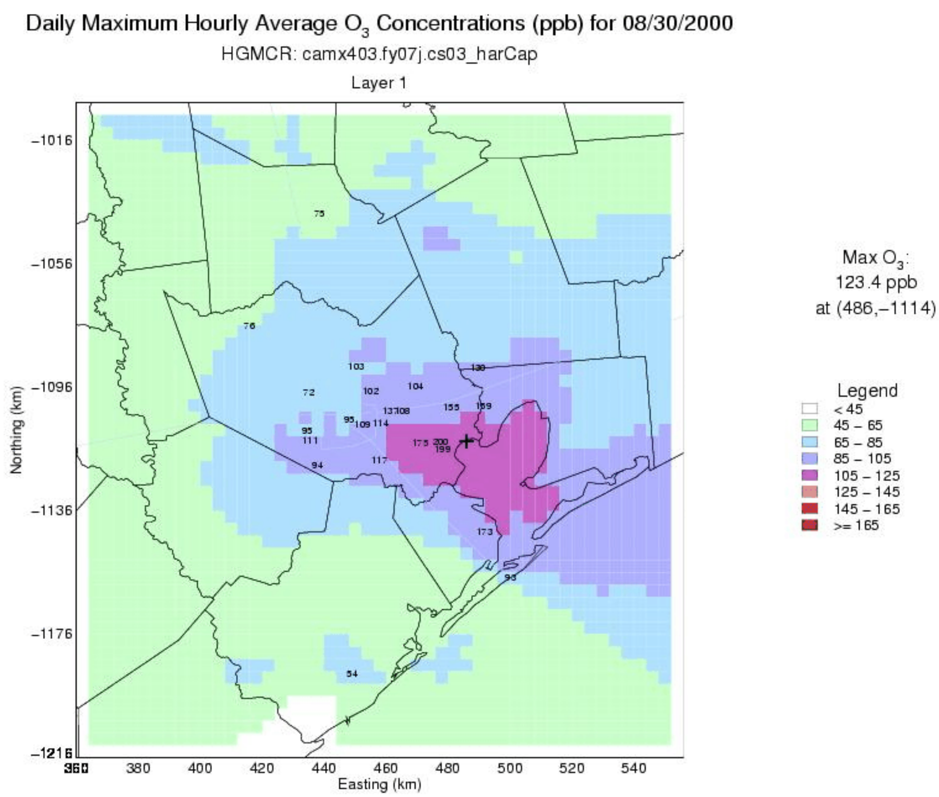
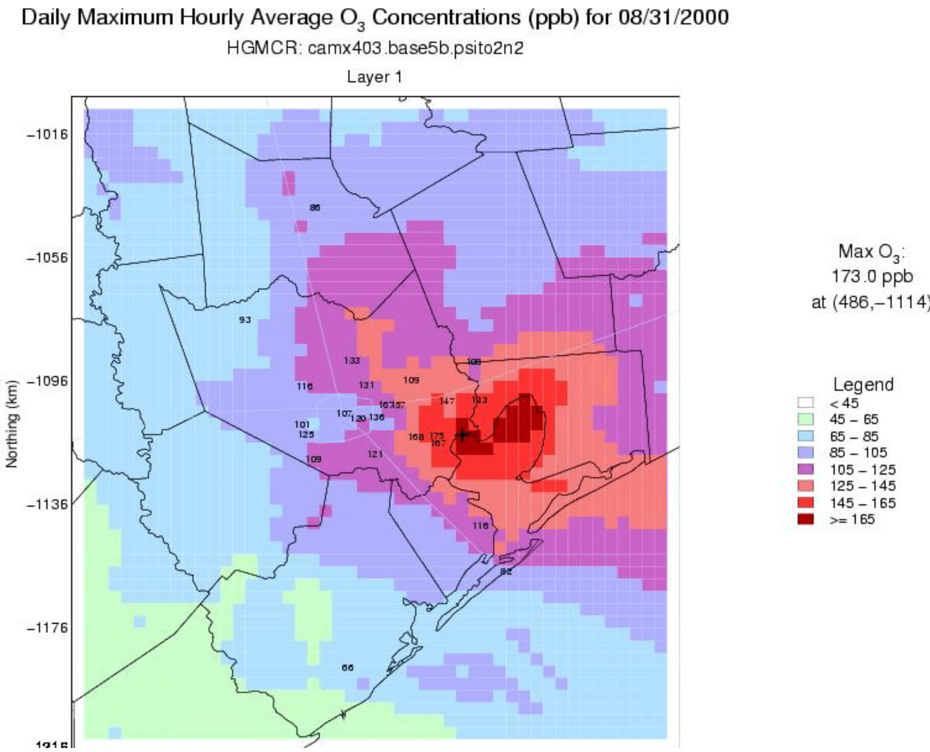


Figure 3.7-3: Peak Ozone Isopleths for Base & Future Control Cases for August 31



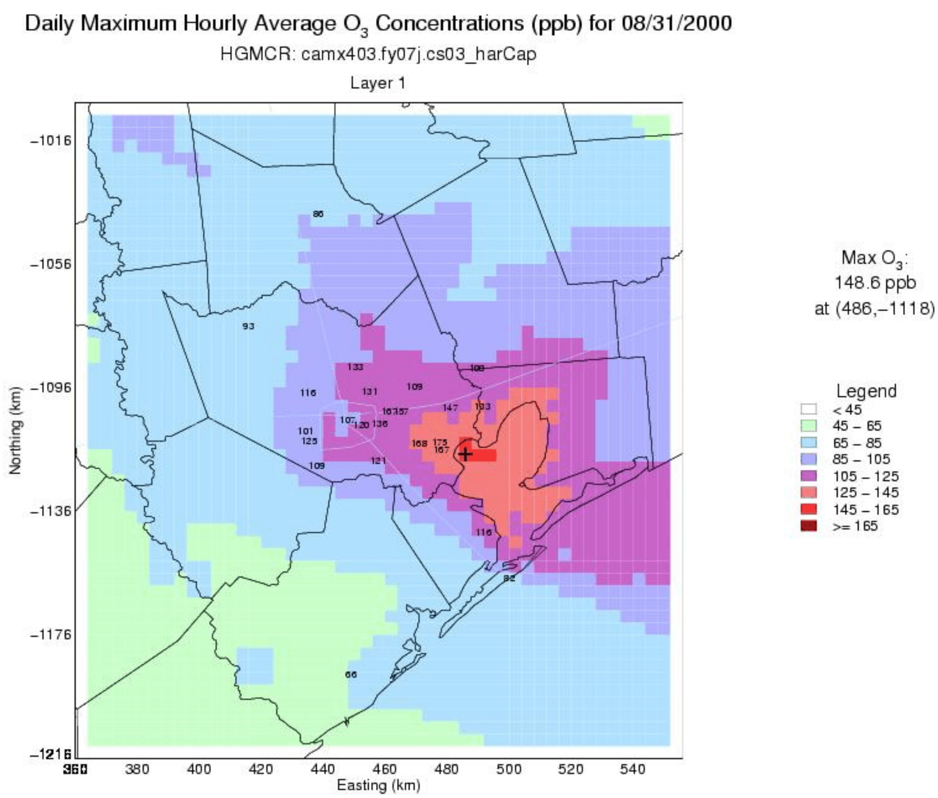
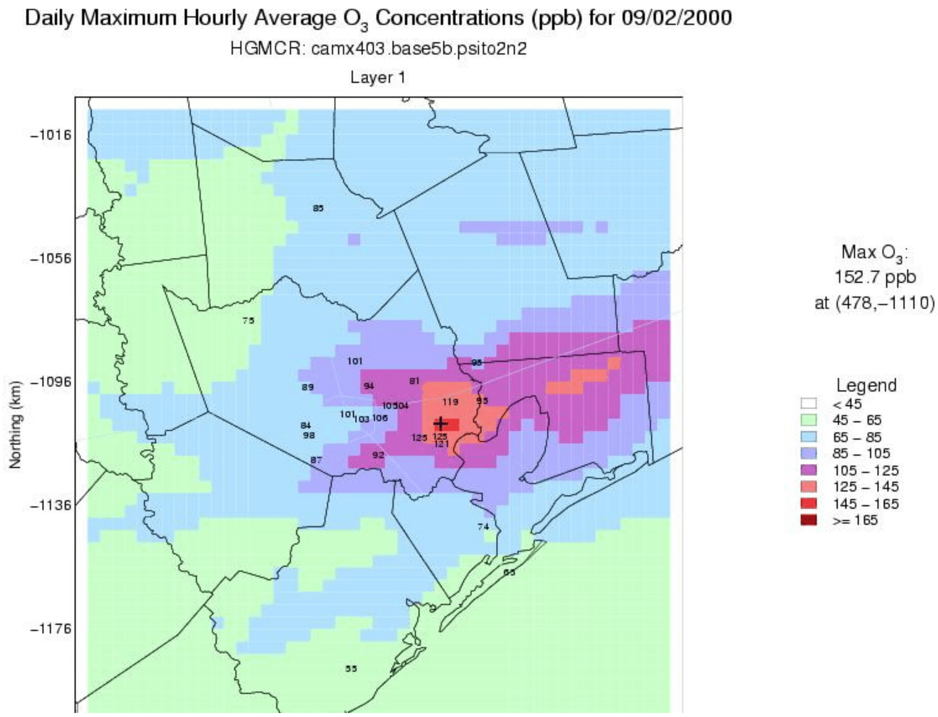


Figure 3.7-4: Peak Ozone Isopleths for Base & Future Control Cases for September 2



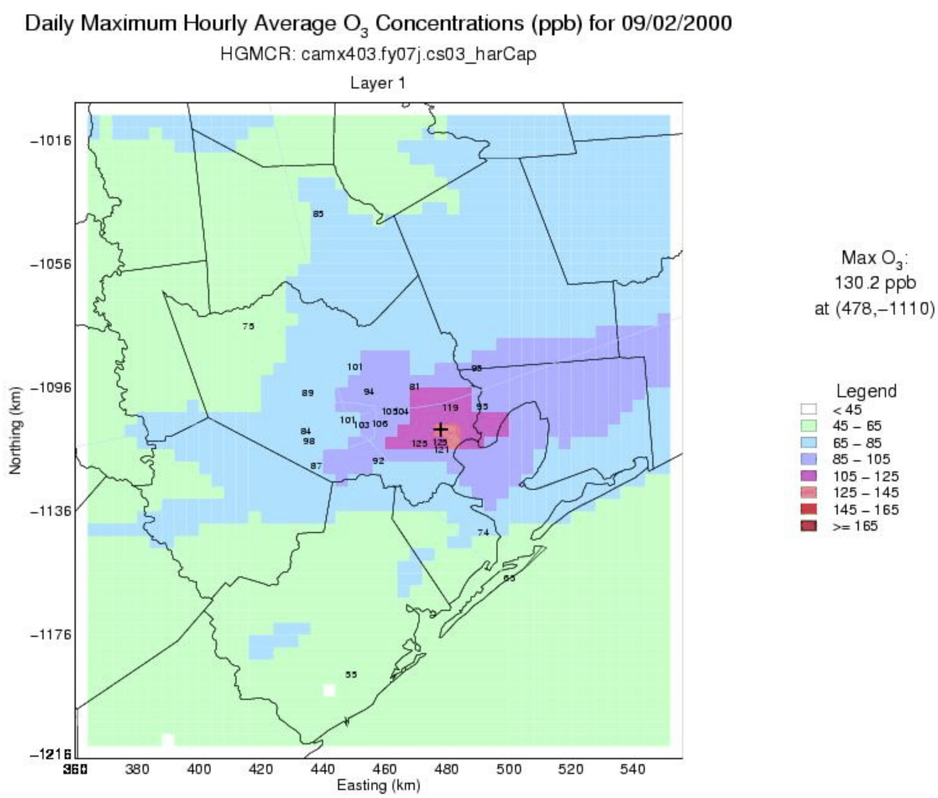
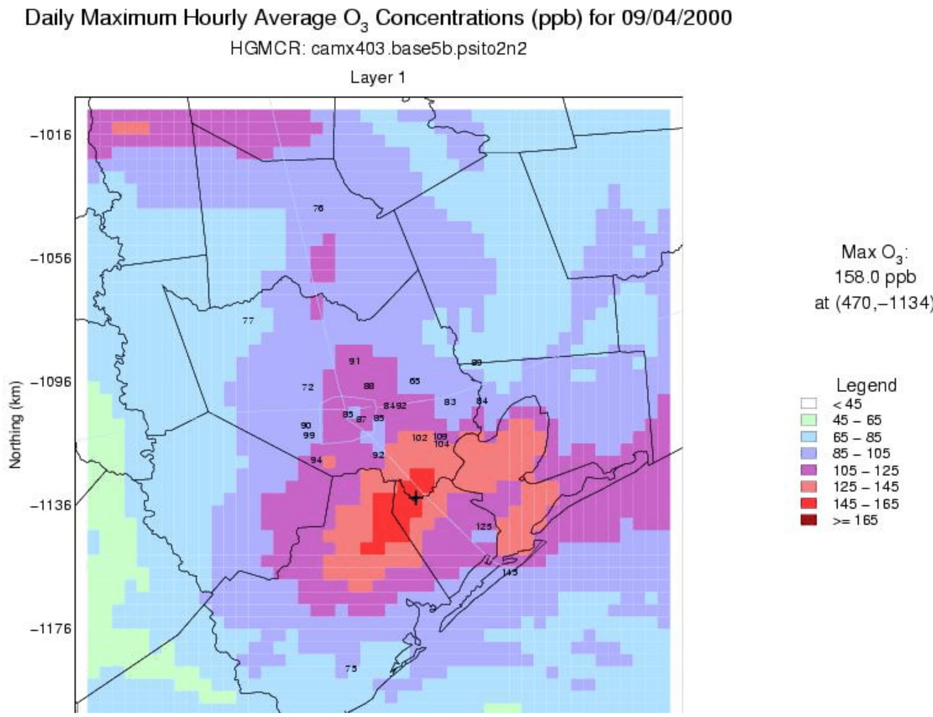


Figure 3.7-5: Peak Ozone Isopleths for Base & Future Control Cases for September 4



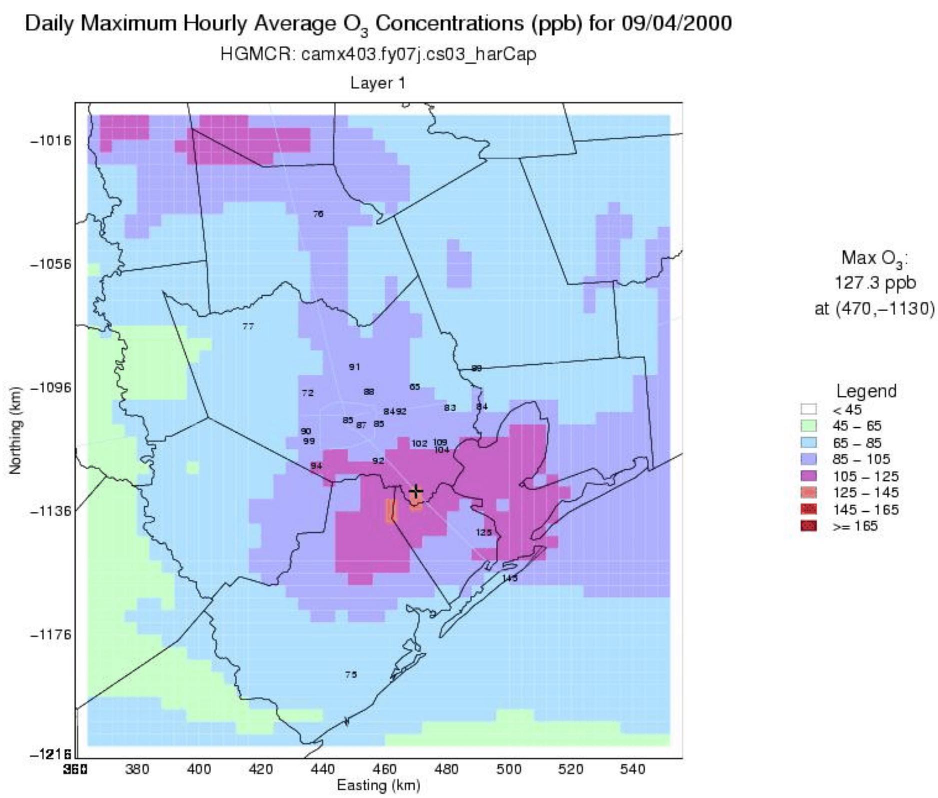
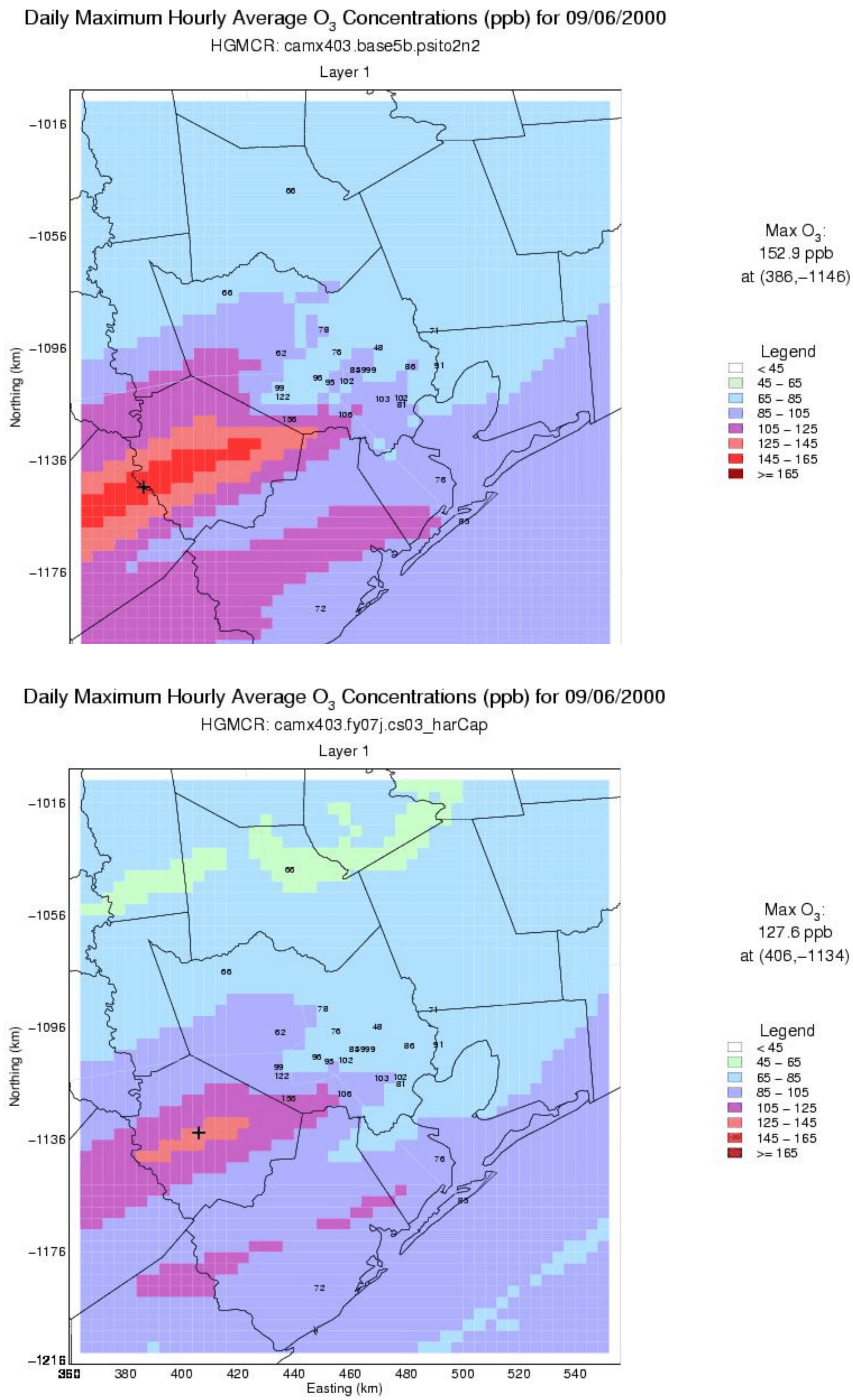


Figure 3.7-6: Peak Ozone Isopleths for Base & Future Control Cases for September 6



3.7.2 Modeling for Latest Control Strategy

Subsequent to the June 2004 proposal, additional future-case modeling was conducted. The controls proposed in June were essentially unchanged, but other aspects of the modeling were enhanced as a result both of comments received and additional development work by TCEQ. The latest control strategy, named CS-08, includes the following modifications:

- Control strategy CS-06a included emissions from planned new electricity generation units (EGUs) in the 8-county HGB area. Adding planned units is the preferred method of modeling future emissions from this sector, since it allows emissions to be simulated in the correct geographic location with realistic stack parameters and emission rates, as opposed to growing units in-place via the traditional EGAS-type growth factors (as is applied to NEGUs). In the HGB area, the MECT program includes all of the potential growth in the 8-county area, i.e., all new EGUs (or NEGUs) in the HGB area are already accounted for in MECT. Therefore, new EGUs previously added to the HGB area were removed from modeling files in CS-08. This change reduced 2007 predicted NO_x emissions in the 8-county area by 15.4 tpd and VOC by 9 tpd.

- In CS-06a, new EGUs in the attainment counties were modeled at their maximum permitted emission rates. Analysis of Electricity Reliability Council Of Texas (ERCOT) trend data showed that in recent years, and projection into the near future, the demand for electricity from ERCOT EGUs has been, and is expected to be, approximately 75 percent of the capacity of these EGUs. Given that new EGUs are permitting at their capacity, their actual emissions would follow the demand for electricity. In CS-08, these new EGUs were modeled at 75 percent of their projected capacity to more accurately reflect these demand (actual emissions) projections. This change did not affect emissions in the HGB nonattainment area (see previous bullet), but resulted in a NO_x reduction of 27.7 tpd across Texas.

- In CS-06a, the cement kilns in the Midlothian area (southwest of Dallas) were modeled at their permitted capacity. In CS-08, the existing kilns were modeled with growth estimates from the Texas Industrial Production Index (TIPI). One new kiln (since 2000) was added to the inventory at its permit allowable emission rate and incorporated permit conditions. These changes were made primarily for the benefit of modeling the DFW area and should have only minimal impacts on modeling for southeast Texas.

- Software upgrades to the EPS2x suite resulted in approximately 5 tpd more NO_x and 1 tpd less VOC emissions than the levels documented for CS-06a.

- New funding from TERP has been made available to the BPA area. This program is expected to reduce NO_x emissions in the BPA area by 3 tpd in 2007. In CS-08, 2.75 tpd of NO_x reductions were applied to nonroad emission sources and 0.25 tpd were applied to onroad emissions.

- Projected construction emissions from three new planned liquid natural gas (LNG) terminals were added to CS-08. No operating emissions were modeled since none of these facilities are expected to be in service in 2007. The units are:

  Golden Pass - Located south of Port Arthur, Texas, with projected 2007 construction emissions of 1.99 tpd of NO_x and 0.25 tpd of VOC

Freeport - Located near Freeport, Texas, with projected 2007 construction emissions of 0.52 tpd of NO_x and 0.05 tpd of VOC

Cheniere - Located in Cameron Parish, Louisiana, east of Golden Pass. At the time CS-08 was run, no projected construction emissions from this facility were readily available, so its projected 2007 construction emissions were modeled as the average of the two units above: 1.25 tpd of NO_x and 0.15 tpd of VOC.

Finally, additional analysis of the future model results showed that wildfires contributed to ozone formation in Southeast Texas during early September of 2000. Figure 3.7-7, *Wildfire activity in 2000 compared with four other recent years*, shows acres of land burned in Texas and Louisiana for a five-year period including 2000. It is clear from the figure that an extraordinary amount of fire activity, and thus, fire emissions, occurred during this period. Because the fire emissions constitute an exceptional event, CS-08 was run with no fires. Strategy CS-08 was also run with wildfire emissions, and these results are presented in the Weight-of-Evidence section of this chapter.

Figure 3.7-7: Wildfire activity in 2000 compared with four other recent years

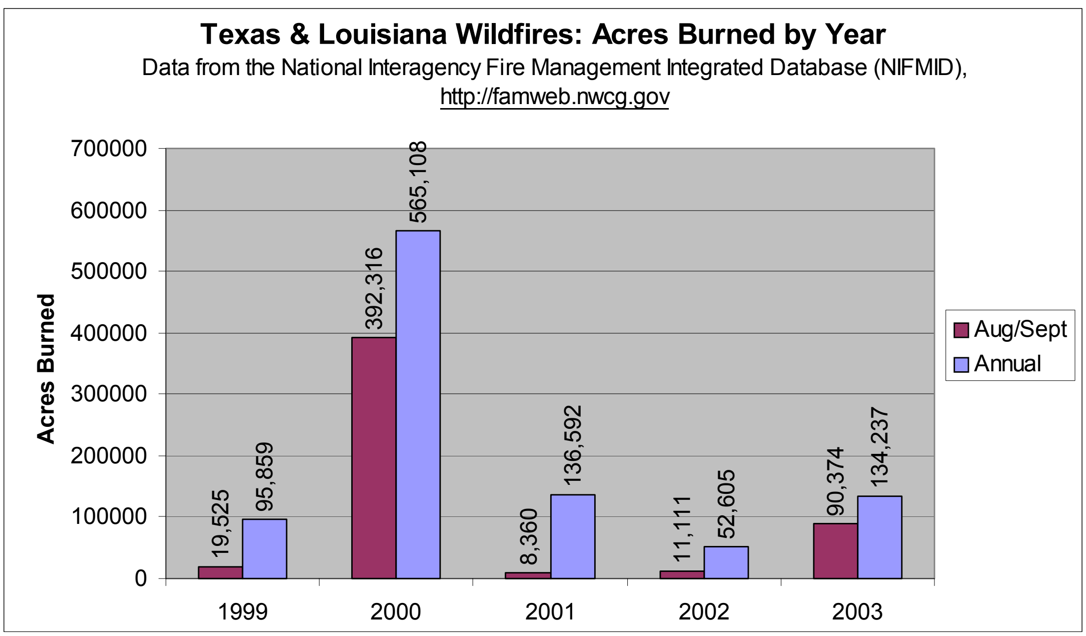


Table 3.7-3, *Anthropogenic Emissions Summary for August 30, 8-County Area*, compares emissions in the 8-county area between CS-08 and CS-06a. As is evident in the table, the only change in the nonattainment area was a small reduction in point source NO_x emissions.

Table 3.7-3: Anthropogenic Emissions Summary for August 30, 8-County Area

Emission Category	NO_x (tpd)			VOC (tpd)		
	Base 5B	CS-06a	CS-08	Base 5b	CS-06a	CS-08
Onroad Mobile	342	175	175	151	89	89
Point Sources	492	174	163	384	255	245
Area/Nonroad	184	155	156	254	234	234
Total	1018	504	494	789	578	568

Table 3.7-4, *Modeled 2007 Peak Ozone Concentrations CS-06a and CS-08*, summarizes peak 2007 ozone concentrations modeled on each day (excluding the two low-ozone days August 27 and 28, and September 5 which had unacceptable performance) for CS-06a and CS-08. Base case peak ozone concentrations are shown for comparison.

Table 3.7-4: Modeled 2007 Peak Ozone Concentrations CS-06a and CS-08

Episode Day	Modeled Peak Ozone Concentration (ppb)		
	Base 5b	CS-06a	CS-08
August 25	156.5	123.4	121.6
August 26	149.4	114.7	113.6
August 29	151.2	114.9	113.6
August 30	137.2	123.4	122.5
August 31	173.0	148.6	147.6
September 1	136.7	120.3	119.5
September 2	152.7	130.2	128.6
September 3	139.3	116.0	115.0
September 4	158.0	127.3	125.2
September 6	152.9	127.6	125.1

CS-08 produces lower 1-hour ozone peaks than did CS-06a. In fact, two days - September 4 and 6, are now less than half of one ppb over the standard, and September 2 is well below 130 ppb. These results, together with the expanded Weight-of-Evidence arguments presented later in this chapter, make a compelling argument for attainment of the 1- hour standard in 2007. August 31 remains over the standard, but this day is very usual, and for that reason should not be the driver for developing controls for the area.

Figures 3.7-8 through 3.7-14 show predicted daily peak ozone concentrations in the HGB area for August 25, 26, 30, and 31, and September 2, 4 and 6 for 2007 future case CS-08 (without wildfires). These may be compared with the peak ozone plots presented for the base case and for control strategy CS-06a presented earlier.

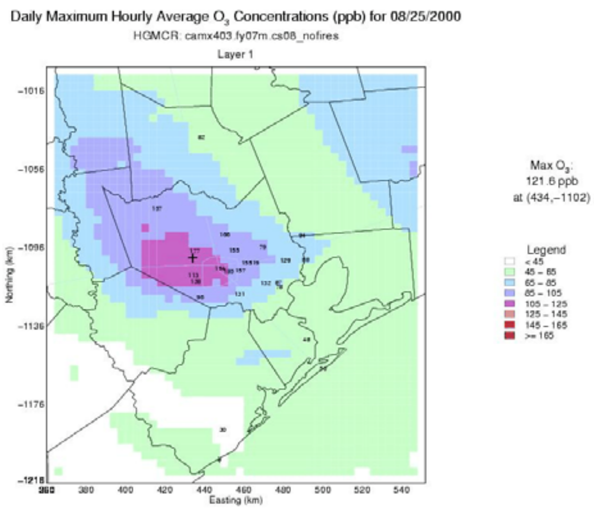


Figure 3.7-8: August 25, 2000

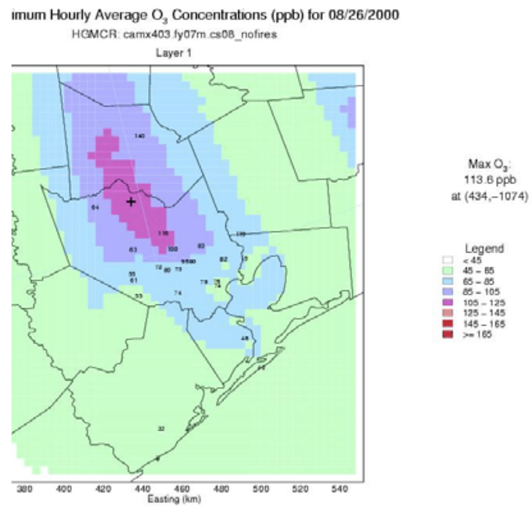


Figure 3.7-9: August 26, 2000

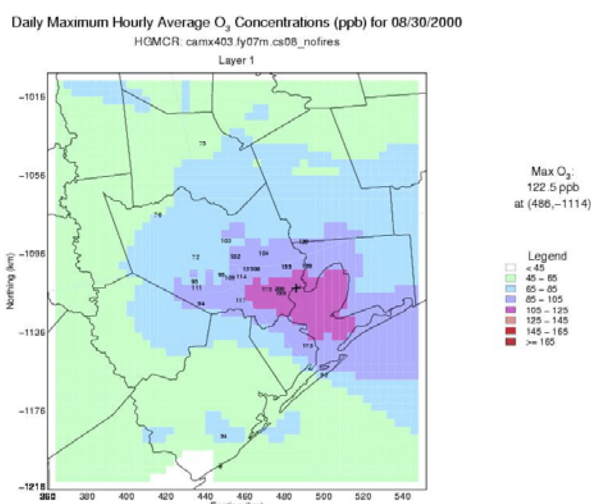


Figure 3.7-10: August 30, 2000

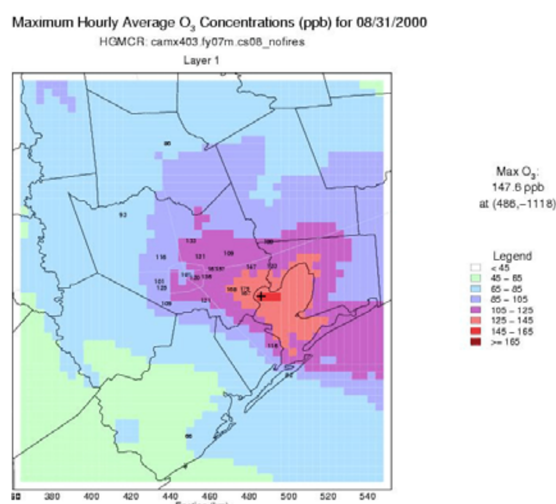


Figure 3.7-11: August 31, 2000

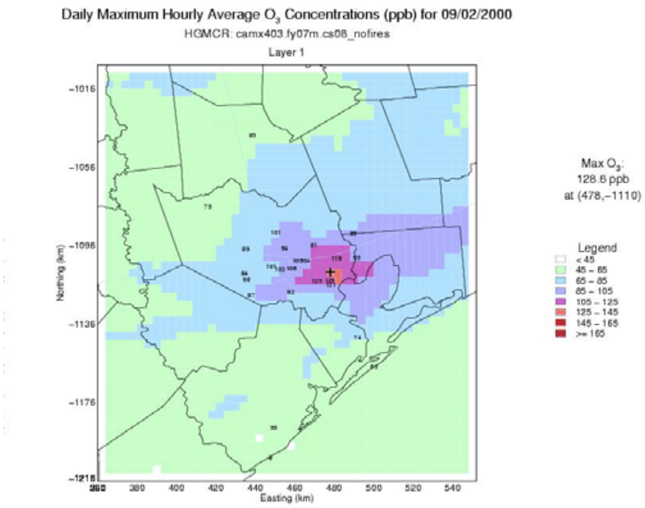


Figure 3.7-12: September 2, 2000

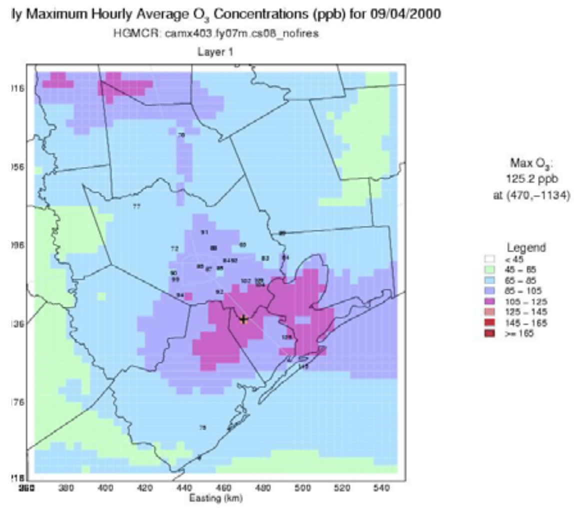


Figure 3.7-13: September 4, 2000

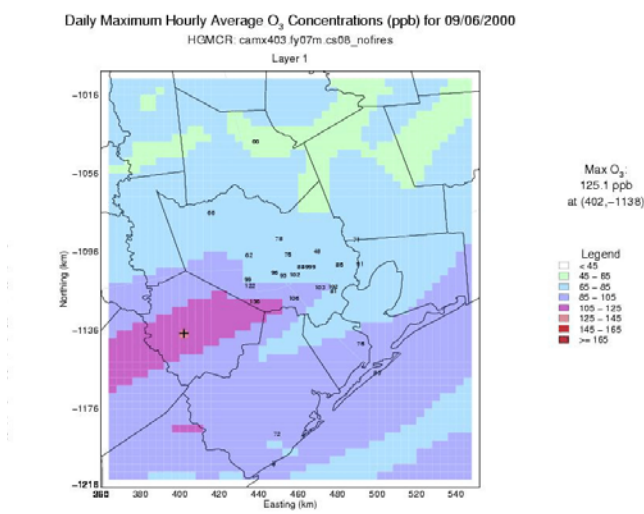


Figure 3.7-14: September 6, 2000