

WOE Chapter 2

Comparison of CMAQ-calculated ozone to surface and aloft measurements

March 9, 2007

Charles Piety, Jennifer Hains, Dale Allen, Jeffrey Stehr

Department of Atmospheric and Oceanic Sciences
University of Maryland
College Park, MD

Abstract

In an effort to assess the ability of the Community Multi-scale Air Quality model (CMAQ) to replicate ozone patterns, particularly high ozone events over the Ozone Transport Region (OTR), comparisons are performed between surface and aircraft ozone measurements and CMAQ ozone simulations using the 2002 base case B1 emissions inventory. Overall, CMAQ does an excellent job of capturing the mean distribution of surface layer ozone during the ozone season. However, the success is somewhat misleading. EPA performance criteria may appear to be independent or offer different information, but in reality, nearly all criteria are strongly geared toward average performance at the surface. This focus steers model developers toward making models that hit the mean very well. However, CMAQ performance aloft which is vital to capturing transport, CMAQ capture of extreme values which is vital to calculating design values for an exceedance-based standard, and model sensitivity, are all neglected. In this analysis, UMD explores several other means of evaluating the CMAQ model by examining its performance only on high ozone days, by separating performance at rural, suburban, and urban sites, and by comparing CMAQ to aloft ozone data from aircraft campaigns. The mixed results of these comparisons show that CMAQ has critical shortcomings (e.g., transport appears to be underrepresented) that appear to be magnified during periods when high ground level ozone concentrations are a concern.

Comparison with aircraft profiles from 136 Regional Atmospheric Measurement Modeling and Prediction Program (RAMMPP) flights reveals that CMAQ has an overall high bias of ~15% from the surface to ~500 meters above sea level (ASL) and a low bias aloft (600-2600 meters ASL) of ~10%. Agreement between CMAQ-calculated and aircraft-measured ozone varies substantially from flight to flight.

Inspection of the surface maps for the OTR reveals that CMAQ, in general, replicates the spatial pattern of high ozone events but often does not capture the full spatial extent or magnitude of the high ozone patterns. Mean CMAQ-calculated and measured 8-hour ozone values from 66 surface ozone monitors in the Baltimore, Washington DC, and Philadelphia nonattainment areas are highly correlated (correlation coefficient, R , of 0.92) over the ozone season (May 15 – September 15) and well correlated ($R=0.81$) when a subset of 38 high ozone days (i.e. days when the peak daily 8-hour average ozone in Maryland exceeded 85 ppbv) are compared. Biases between CMAQ-calculated and observed 8-hour ozone mixing ratios are minimal (-1.6 ppbv)

when averaged over the entire ozone season. However, larger negative biases are seen during high ozone days (-2.2 ppbv at urban sites and -7.7 ppbv at rural/suburban sites).

The high bias near the surface and low bias aloft is indicative of an underestimation of transport by CMAQ. Aloft is where most transport occurs; ground-level air does not move as readily. On the highest ozone days, CMAQ's performance is not as good as on lower ozone days. This is a statistical reflection of CMAQ's inability to capture large-scale deviations from average or median conditions. These deviations occur on days with poor air quality. CMAQ performs better at urban sites than at suburban and rural areas. This bias provides more evidence that CMAQ is missing incoming ozone, possibly transport. Often these rural/suburban areas are dominated by power-plant emissions more than they are dominated by motor vehicle emissions. The bias may also indicate that CMAQ's relatively coarse vertical resolution is unable to resolve the transport of point source (i.e. power plant) emissions. In particular, performance at upwind sites with fewer nearby sources is poorer on the whole than it is at other sites (see Chapter 9).

None of these shortcomings are reflected in EPA's traditional ozone model performance measures. However, these shortcomings make it necessary to consider CMAQ output and other evidence when evaluating the probability of success of State Implementation Plans (SIPs). This Chapter goes with Chapter 9 on model performance, uncertainty, and responsiveness.

Introduction

The following analysis details comparisons between ozone observations and CMAQ ozone simulations using the 2002 "base case B1" emissions inventory. This analysis is important in understanding the limitations of CMAQ and its strengths and weaknesses in simulating air quality over the Mid-Atlantic in particular and the Ozone Transport Region in general. Understanding basic CMAQ model performance, compared to surface and aloft observations, will provide a foundation for additional commentary on CMAQ simulations related to future emissions strategies. Some key questions that will be addressed by this analysis include:

- *How well does CMAQ replicate the mean transport of ozone and its precursors?*
- *How well does CMAQ capture temporal fluctuations in ozone concentrations both at the surface and aloft?*
- *What do identified CMAQ limitations imply for modeled attainment demonstrations using future emissions strategies?*

Many of the images of surface comparisons presented in this study employ a map which covers a large portion of the Ozone Transport Region (OTR) (see Figure 1). The domain in Figure 1 was selected in an effort to present as much of the OTR modeling domain as possible while keeping the focus on the Baltimore, Washington, D.C. and Philadelphia nonattainment areas (NAA). Additional comparative images were derived from CMAQ ozone vertical profiles and aircraft ozone vertical profiles collected by the University of Maryland research aircraft. Other ozone images include comparisons for ozone monitors within the Baltimore, Washington D.C., and Philadelphia NAAs. Time

series plots were derived from CMAQ simulations of the entire summer of 2002. Because there is keen interest in CMAQ model performance when high ozone was observed over Maryland, as a subset of the model evaluation, the 38 days when the observed peak daily 8-hour average ozone exceeded 85 ppbv (i.e. the NAAQS for 8-hour average ozone) somewhere in Maryland were analyzed separately.

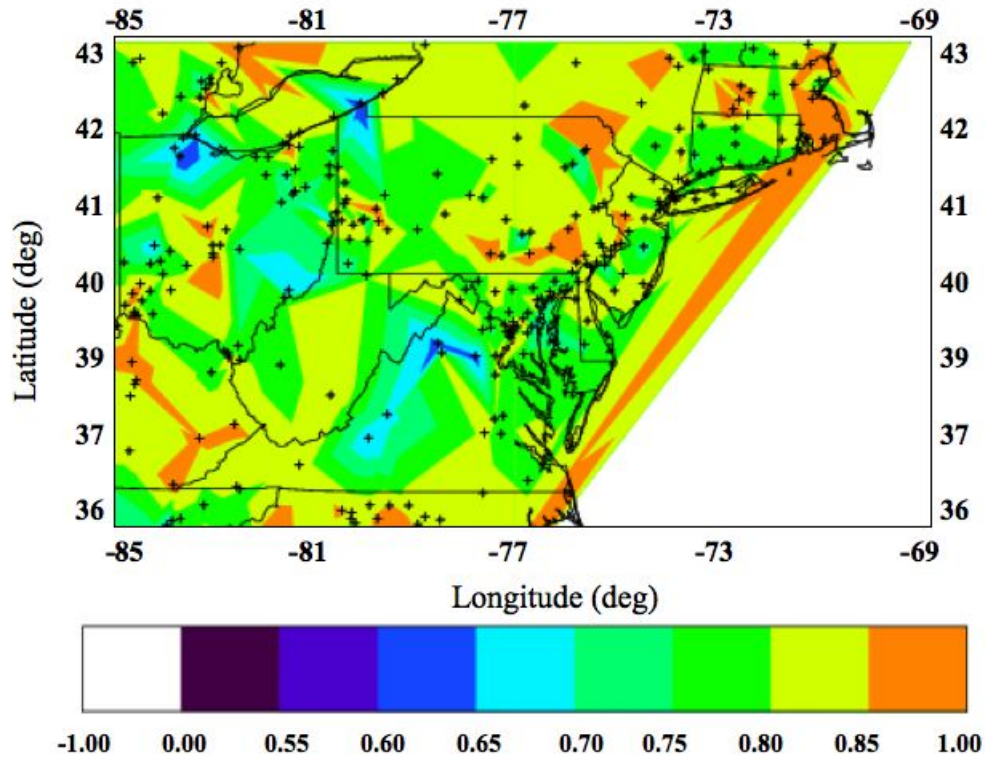


Figure 1. A plot of the correlation coefficients (R) for daily maximum 8-hour ozone over the OTC modeling domain (base case B1 vs observations). Time period for the plot is the summer of 2002 (May-September).

Methods

As mentioned previously, the evaluation of the CMAQ base case B1 simulation was performed using surface and aloft observations. Comparisons with surface data were conducted using surface ozone maps covering much of the OTR modeling domain, while time series plots were produced for ozone monitors within the Baltimore, Washington D.C., and Philadelphia NAAs. For surface maps (i.e. 2002 ozone observations, base case B1 simulations and difference plots), Delaunay Triangulation was used to map irregularly spaced observations and model output onto the desired grid. The surface plots present the peak daily 8-hour average ozone and do not represent one particular time of day. The timing of peak 8-hour ozone concentrations varies with location. For comparisons of time series of surface ozone, CMAQ-simulated surface ozone was sampled at locations corresponding to the latitude and longitude of surface ozone monitors.

For CMAQ comparisons to aloft data, the lowest 16 layers of the CMAQ model were assumed to be 10, 24, 68, 116, 185, 282, 398, 544, 727, 949, 1212, 1523, 1886, 2312, 2820, and 3393 meters above ground level (typical altitudes for summertime conditions in the northeastern U.S.). These elevations were then converted to meters above sea level by adding the surface elevation of each grid point. CMAQ modeled ozone was linearly interpolated in altitude and time to best match the location and time of aircraft measurements.

The differences between aircraft and model profiles were calculated, accounting for differences in shape (the locations of the minima and maxima in the profiles) and magnitude (absolute differences in mixing ratio). All aircraft spiral measurements and corresponding CMAQ-calculated ozone fields were initially averaged into 100 m altitude bins. This allowed for consistent comparisons between pairs of modeled and measured ozone. To calculate the difference between modeled and measured ozone, UMD accounted for the shape and morphology of the profiles by looking at the absolute difference in the mixing ratio as well as the slope and correlation coefficient between modeled and measured ozone in four altitude bins (250-650 m, 651-1150 m, 1151-1650 m, and 1651-2150 m). The altitude bins were selected in an effort to obtain a statistically significant number of data points in each bin while also allowing for some partitioning between the planetary boundary layer and the lower free troposphere. Equation 1 (Taubman et al., 2006, Hains et al., 2007) was used to calculate the difference (D_{ij}) between each pair of modeled (i) and measured (j) ozone profiles:

$$D_{ij} = \sum_{k=1}^{k=4} \left\{ \underbrace{\left(\sum_{a=n}^{a=m} \text{abs}(c_{ia} - c_{ja}) \right)^2}_{\text{Part}_1} * \underbrace{\left(1 + [1 - R] + [1 - \exp(-(s - 1)^2)] \right)}_{\text{Part}_2} \right\} \text{ Equation 1}$$

Here k represents the four different altitude bins, a is an index that can take on values between 1 and the total number of data points between 250 and 2150 meters, n (m) is an index that point to the first (last) data point in each bin. The number of data points within each altitude bin ($m-n+1$) varies with altitude bin (k). The ozone mixing ratio is represented by c for the i^{th} (modeled) and j^{th} (measured) profiles. A regression between CMAQ calculations and observations is calculated, giving the slope, s , and correlation

coefficient, R , for each of the four layers. The first part of Equation 1 calculates the square of the sum of the differences between values (from observations and CMAQ) at each altitude level, k . The second part of the equation multiplies the difference by one plus differences associated with the correlation and slope. When the correlation is small or negative, the profiles are very different and the $1 - R$ portion increases, which increases the total difference D_{ij} . The exponent of the slope portion is used to account for the slope of the best-fit line between CMAQ and observations in each of the profile altitude bins. A slope near one suggests that the profiles are similar and should therefore add little to the total difference. The exponent of the slope was used to guarantee that slopes much different from one would make the exponential term small, thereby increasing the $(1 - \text{exponent})$ term and increasing the total difference. Taken as a whole, Equation 1 becomes large when absolute differences between predicted and observed concentration are large,

Results

Figure 1 is a plot of R , the correlation coefficient, between peak 8-hour ozone surface observations and peak 8-hour ozone values from CMAQ. The correlation coefficient indicates the strength and direction of a linear relationship between two variables. These correlations were calculated, site-by-site over the ozone season, May 15- September 15, 2002. Figure 1 shows values of R ranging from 0.55 – 0.90, with the poorest correlation occurring over northwestern OH, northern VA and northwestern PA. Correlations are better over much of the Interstate-95 corridor from Washington, DC to Boston, MA with values ranging from 0.75 - 0.90. Over the Mid Atlantic region, R ranges from 0.60 – 0.90, with the poorest correlations occurring over central-western VA. Model performance is poorer over the Ohio River Valley ($R=0.65-0.80$). Figure 1 demonstrates that on average, CMAQ simulates peak 8-hour ozone reasonably well. The relatively low correlations over some portions of the Ohio River Valley and central VA are interesting. These regions are often upwind of the Mid-Atlantic region during major pollution events. The poorer performance in these regions could indicate that CMAQ is not accurately representing ozone pollutant transport into the Mid Atlantic region from these two regions. More conclusive statements must await comparisons with aircraft measurements taken downwind of the Mid-Atlantic region.

Table 1 lists the surface monitors used in the surface comparisons presented in Figures 2 and 3. Information on how CMAQ performs at individual monitors is also shown in Figure 1. For example, CMAQ-calculated and measured 8-hour ozone at Edgewood is well correlated ($R=0.84$). The mean bias (-2 ppbv) is small but larger on high ozone days (-9 ppbv). The best fit line to the relationship between CMAQ-calculated (x) and measured (y) 8-hour ozone at Edgewood is given by $y = 10.0 + 0.88x$. The null hypothesis that CMAQ-calculated ozone equals measured 8-hour ozone ($y = 0.0 + 1.0x$) at Edgewood is shown to be untrue at the 95% (2 sigma) confidence level. Figure 2 is a time series of the average, peak 8-hour ozone values, from surface comparisons of 66 ozone monitors located within Virginia; Maryland; Washington, D.C.; and the Philadelphia non-attainment area. Figure 3 shows surface ozone monitor data categorized by EPA's classification of monitor location (urban, suburban and rural). Table 2 presents summary average statistics associated with Figure 2 and Figure 3.

Considering 66 monitors, CMAQ displays a 1.6 ppbv low bias and a 7.0 ppbv centered root mean square error (RMS_c) for the entire ozone season and a low bias of 7.0 ppbv and an RMS_c of 5.8 ppbv when only the 38 high ozone days are included. The correlation coefficients, R , are 0.92 and 0.81 for the entire 2002 ozone season and the high ozone days respectively. Figure 3 shows that biases are small when averaged over the ozone season ranging from a 2.4 ppbv low bias at suburban sites to a 1.5 ppbv low bias at urban sites. Biases are larger on high ozone days ranging from a 2.2 ppbv low bias at urban sites to a 7.7 ppbv low bias at suburban and rural locations. The larger biases at suburban/rural locations exist because CMAQ tends to underestimate the spatial extent of high ozone events. i.e., it tends to underestimate the regional character of ozone episodes. Both deficiencies in the chemical algorithm (the CB4 mechanism was originally designed for urban locations) and transport (it is often underestimated) are responsible for these differences in biases. Dale/Jeff. Correlation coefficients are insensitive to monitor location. Overall, urban performance is better than suburban, which is in turn better than rural performance, showcasing the tendency for CMAQ to miss incoming and upwind ozone.

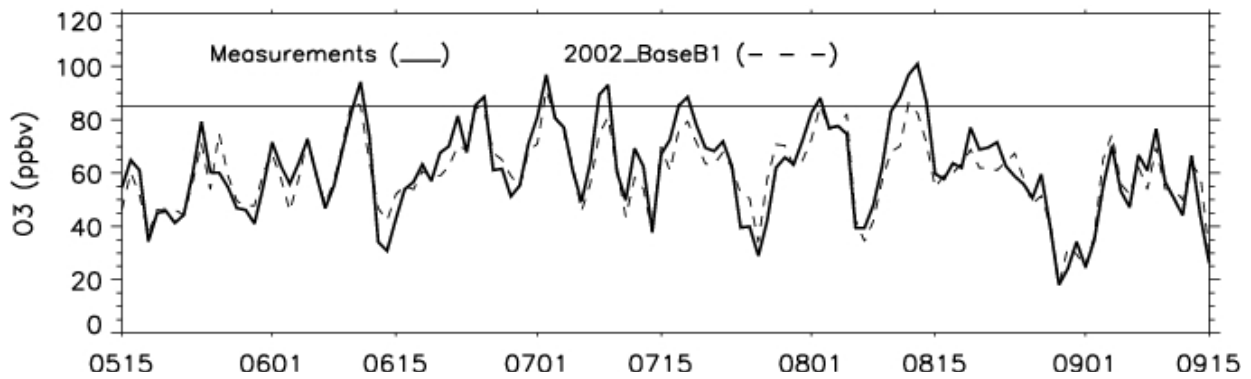


Figure 2. Comparison of the average peak 8-hour ozone from 66 surface ozone monitors in Virginia, Maryland, Washington D.C., and the Philadelphia NAA. Out of the 123 days, 38 days are considered “high ozone days” when 8-hour peak ozone is greater than or equal to 85 ppbv at one or more monitors within the Baltimore NAA. CMAQ is generally lower than the peak observations and higher than the nighttime minima.

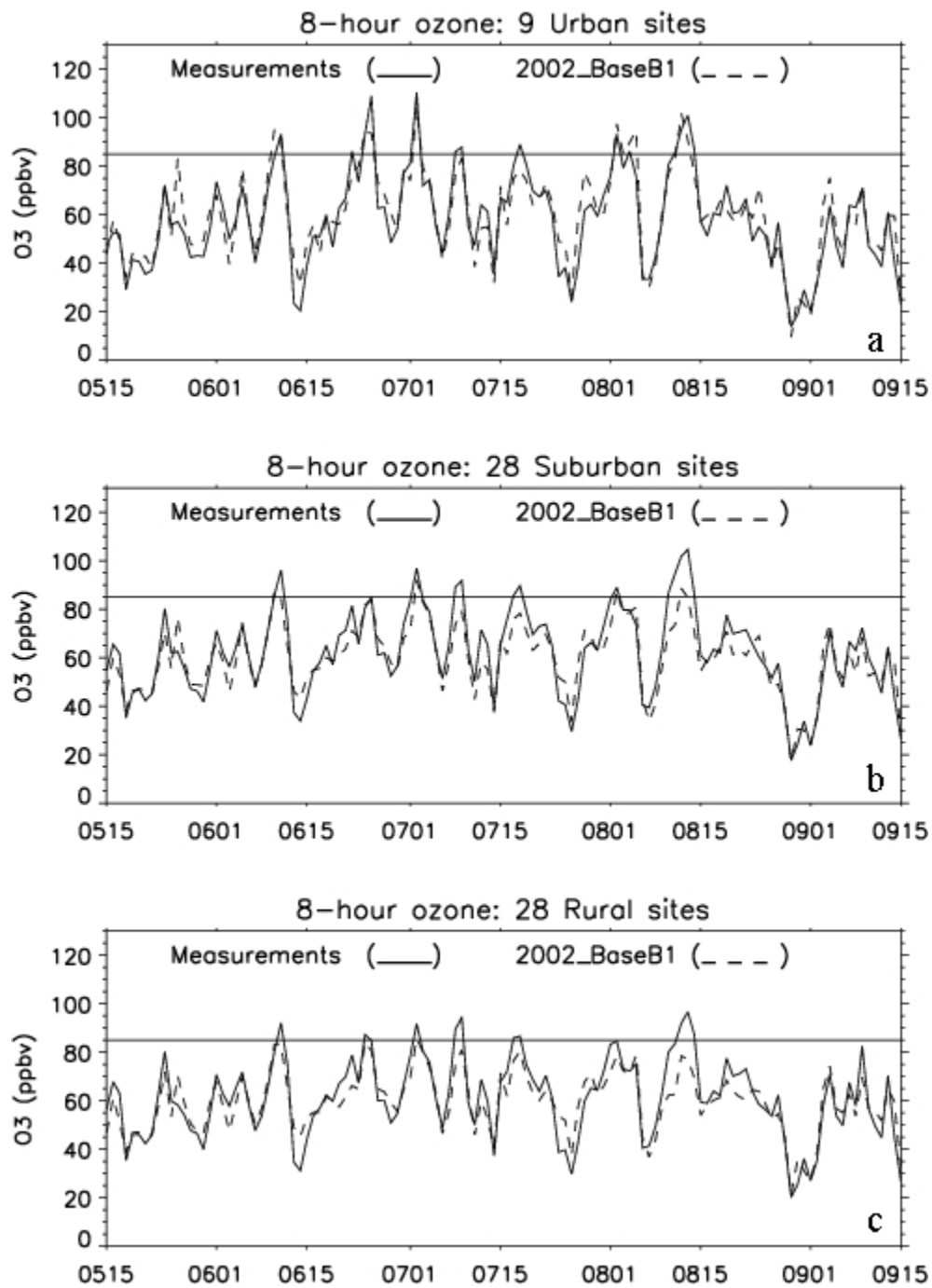


Figure 3. Surface ozone comparisons for (a) urban, (b) suburban, and (c) rural surface sites in Virginia, Maryland, Washington D.C., and the Philadelphia non-attainment area (n = the number of monitors). Redo panel letters (a, b, c) with reasonable font size and font type.

Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
100010002	Killens Pond	R	-75.56	38.98	0.82	10.2	-2.1	-8.7	-1.5 + 1.06x	Y
100031007	Lums Pond	R	-75.73	39.55	0.84	10.7	5.3	-0.9	-5.7 + 1.01x	Y
100031010	Brandywine	R	-75.56	39.82	0.85	11.7	-2.3	-8.7	-4.3 + 1.11x	Y
100031013	Bellefonte	S	-75.5	39.77	0.85	10.6	3	-0.2	0.3 + 0.94x	Y
100051002	Seaford	S	-75.61	38.64	0.82	10.5	-3.4	-10.1	-4.0 + 1.13x	Y
100051003	Lewes	R	-75.16	38.78	0.85	9.6	-0.3	-6	-6.3 + 1.11x	Y
110010025	Takoma Park	U	-77.02	38.98	0.89	9.2	4	1.5	3.1 + 0.89x	N
110010041	River Terrace	U	-76.95	38.9	0.89	9.4	-2.2	-6.9	5.4 + 0.94x	N
110010043	McMillan Reser	U	-77.01	38.92	0.9	9.2	-8.1	-13.8	8.5 + 0.99x	N
240030014	Davidsonville	R	-76.65	38.9	0.88	10.3	-4.5	-10.4	0.7 + 1.06x	Y
240030019	Fort Meade	S	-76.73	39.1	0.88	9.8	-2.4	-5.8	7.2 + 0.93x	N

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Table 1 (continued) CMAQ performance at individual monitors										
Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
240051007	Padonia	S	-76.63	39.46	0.83	11.5	-1.2	-6.7	5.4 + 0.93x	Y
240053001	Essex	S	-76.47	39.31	0.82	13.6	2.4	1.9	15.7 + 0.72x	N
240130001	South Carroll	R	-77.04	39.44	0.82	9.9	0.3	-4.6	2.4 + 0.96x	Y
240150003	Fair Hill	R	-75.86	39.7	0.85	12	-1.6	-10.5	-8.3 + 1.16x	N
240170010	S. Md (Hughesville)	R	-76.81	38.5	0.84	10.3	0.1	-6.6	-4.7 + 1.07x	Y
240210037	Frederick Apt	S	-77.38	39.41	0.85	9.6	-0.4	-5	3.9 + 0.94x	Y
240251001	Edgewood	R	-76.3	39.41	0.84	12.4	-2	-9	10.0 + 0.88x	N
240259001	Aldino	S	-76.2	39.56	0.87	11.1	-3.2	-11	3.1 + 1.00x	Y
240290002	Millington	R	-75.8	39.31	0.83	11.3	-2.8	-11.5	-2.4 + 1.09x	Y
240313001	Rockville	R	-77.11	39.11	0.86	9.9	5.3	2.5	1.5 + 0.89x	N

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Table 1 (continued) CMAQ performance at individual monitors										
Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
240330002	Greenbelt	R	-76.83	39.02	0.87	10.3	1.2	-2.4	3.4 + 0.93x	Y
240338003	PG Equest Ctr	R	-76.74	38.81	0.83	12.1	2.7	-4.7	-7.5 + 1.07x	Y
240430009	Hagerstown	R	-77.72	39.57	0.87	9	-2.7	-6.7	1.1 + 1.03x	Y
245100053	Baltimore County	U	-76.55	39.29	0.78	14.7	5.9	5.7	6.8 + 0.79x	N
340010005	Nacote Creek	R	-74.46	39.53	0.85	9.6	3	-1.5	-4.7 + 1.03x	Y
340070003	Camden Lab	S	-75.1	39.92	0.84	12.2	-3.3	-7.5	6.8 + 0.94x	Y
340071001	Ancora Hospital	R	-74.86	39.67	0.88	10.7	-3.5	-8.7	-4.2 + 1.12x	N
340110007	Millville	R	-75.03	39.42	0.87	10.3	-1.6	-8.4	-8.1 + 1.16x	N
340210005	Rider U	S	-74.75	40.28	0.86	11.6	-2.6	-8.5	1.5 + 1.02x	Y
340290006	Colliers Mills	R	-74.45	40.07	0.85	13.2	-4.6	-12.7	-7.8 + 1.20x	N

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Table 1 (continued) CMAQ performance at individual monitors										
Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
420170012	Bristol	S	-74.88	40.11	0.87	11.6	-0.5	-4.8	1.3 + 0.99x	Y
420290050	West Chester	S	-75.6	39.94	0.87	10.1	-5.7	-11.7	0.7 + 1.09x	Y
420290100	New Garden (Ai)	R	-75.77	39.83	0.85	11.8	-7.7	-17.3	-3.5 + 1.19x	N
420450002	Chester	U	-75.37	39.84	0.86	11.4	-4.1	-10.1	2.5 + 1.03x	Y
420910013	Norristown	S	-75.31	40.11	0.87	10.4	0.1	-3.7	-0.2 + 1.00x	Y
421010004	Frankford (Lab)	U	-75.1	40.01	0.84	11.2	13.6	13.3	0.8 + 0.76x	N
421010014	Northwest (Rox)	S	-75.24	40.05	0.84	11	-1.9	-4.3	8.8 + 0.88x	N
421010024	Northeast (Air)	S	-75.01	40.08	0.88	11.2	-6.6	-12.9	3.5 + 1.05x	Y
421010136	Southwest (Elm)	U	-75.22	39.93	0.86	11	5.3	1.9	-0.4 + 0.92x	Y
510130020	Arlington Co.	U	-77.06	38.86	0.9	9.7	-4.5	-11.3	2.7 + 1.03x	Y

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Table 1 (continued) CMAQ performance at individual monitors										
Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
510330001	Caroline Co.	R	-77.38	38.2	0.76	10.1	2.1	1.3	5.3 + 0.88x	Y
510360002	Charles City	S	-77.26	37.34	0.84	12.1	-4.6	-11.9	-9.4 + 1.23x	N
510410004	Chesterfield	R	-77.59	37.36	0.76	11.2	-4.6	-12.7	3.9 + 1.01x	Y
510590005	Chantilly	R	-77.47	38.89	0.86	9.3	-5.6	-10.1	7.4 + 0.97x	N
510590018	Mt. Vernon	S	-77.08	38.74	0.85	11.7	2.1	-4.4	-1.6 + 0.99x	Y
510590030	Lee Park	S	-77.11	38.77	0.88	10.3	-3.6	-10.3	2.9 + 1.01x	Y
510591005	Annandale	S	-77.16	38.84	0.83	12.1	-1.8	-7.1	4.9 + 0.95x	Y
510595001	McLean	S	-77.2	38.93	0.79	12.3	6.4	3.7	9.0 + 0.76x	N
510610002	Fauquier Co.	R	-77.77	38.47	0.77	9.7	-0.2	-7	-6.8 + 1.12x	Y
510690010	Frederick Co.	R	-78.08	39.28	0.8	10.6	-0.9	-7	-5.0 + 1.10x	Y
510850003	Hanover Co.	S	-77.22	37.61	0.8	11.6	-5.3	-14.3	-2.8 + 1.13x	Y
510870014	Henrico Co.	S	-77.4	37.56	0.85	9.9	-5.1	-9.1	5.8 + 0.99x	Y

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Site ID	Monitor Name	Classification ^a	Latitude	Longitude	R ^b	RMS ^c	Bias (all days)	Bias (high O ₃ days)	Intercept + (Slope)x	y = x ? ^d
511071005	Loudoun Co.	S	-77.49	39.02	0.85	10.3	-6.6	-13.7	0.2 + 1.11x	Y
511130003	Madison Co. -	R	-78.44	38.52	0.75	9.3	-9.1	-15.1	4.0 + 1.09x	Y
511390004	Page Co.	R	-78.5	38.66	0.66	10.4	-2	-7.4	4.9 + 0.95x	Y
511530009	Prince William	S	-77.64	38.86	0.81	10.3	-3.4	-12.1	-8.3 + 1.20x	N
511611004	Roanoke Co.	S	-79.88	37.29	0.8	9.2	-4.9	-9.8	0.9 + 1.07x	Y
511630003	Rockbridge Co.	R	-79.51	37.63	0.61	11.2	-2.2	-7.8	1.2 + 1.02x	Y
511790001	Stafford Co.	S	-77.37	38.48	0.87	9.2	-0.3	-4.6	3.1 + 0.95x	Y
511970002	Wythe Co.	R	-81.25	36.89	0.72	9.8	-6.6	-12.4	-4.1 + 1.20x	Y
515100009	Alexandria	U	-77.04	38.81	0.88	9.7	4.9	-0.3	-2.7 + 0.96x	Y
516500004	Hampton	S	-76.4	37	0.72	14.8	-9.9	-11.7	24.4 + 0.71x	N
518000004	Suffolk - TCC	S	-76.44	36.9	0.86	9.4	0	-5.1	-3.1 + 1.05x	Y
518000005	Suffolk - Holl	S	-76.73	36.67	0.79	11.9	-3.4	-12.6	-11.0 + 1.27x	N

^a Monitor classification (Urban, Suburban, or Rural) from <http://www.epa.gov/air/data/index.html>

^b Correlation coefficient (R)

^c Centered root mean square error after removing bias (a positive bias indicates that CMAQ-calculated 8-hour ozone exceeds measured ozone)

^d At each site, a regression line was fit to the relationship between 8-hour measured (x) and CMAQ calculated (y) ozone. The resulting slope and intercept were tested against the null hypothesis that the actual slope was one and actual intercept zero. Sites where neither hypothesis can be rejected are identified with a Y (good fit). Sites where either the slope or intercept are outside two sigma confidence limits are identified with N (poor fit).

Table 2. Comparison of 2002 Base B1 Simulation and 2002 Observations			
	Bias (ppbv)	RMS error [*]	R ^{**}
<i>All 66 Sites</i>			
Entire ozone Season	-1.6	7.0	0.92
High ozone Days	-7.0	5.8	0.81
<i>Urban Sites</i>			
Entire ozone Season	-1.5	7.7	0.92
High ozone Days	-2.2	7.1	0.86
<i>Suburban Sites</i>			
Entire ozone Season	-2.4	7.0	0.92
High ozone Days	-7.7	6.3	0.78
<i>Rural Sites</i>			
Entire ozone Season	-1.7	7.5	0.90
High ozone Days	-7.7	6.1	0.76
* Centered root mean square error in ppbv			
** R is the correlation coefficient			

Figure 4 shows the median, 25th and 75th percentiles for all aircraft-measured ozone (136 profiles) and matching CMAQ ozone predictions for 2002. While differences between the model-calculated and observed profiles are substantial, the model-calculated profile always remains between the 25th and 75th percentile of the observed profile. The large width of the observed profile (roughly 25 ppbv between the 25th and 75th percentile) shows that lower tropospheric ozone amounts vary substantially even on days when observed ozone amounts were expected to be large (most flights were made on days when high ozone concentrations were forecast). The large variability in observed ozone concentrations is also a function of the large spatial and temporal range of the data. The aircraft measurements were taken at locations extending from North Carolina to Maine, and therefore sometimes sampled very different chemical regimes. Another factor that can create variability is the occasional natural variability in ozone with altitude that produces scattered data at the high time resolution of the ozone instrument onboard the aircraft. Longer averaging periods would produce smoother data sets, but at the expense of vertical resolution in the boundary layer (the aircraft is ascending or descending when collecting data so that a longer sampling period implies more vertical distance will be covered per sampling period). The aircraft ozone instrument collects a running 1 minute average ozone value, which is obtained from 10-second sampling intervals. The data in Figure 4 suggests CMAQ has a high bias of ~15 % from the near surface to ~500 m above ground, and the aircraft profiles have on average 10% more ozone than the CMAQ profiles aloft, from 600 – 2600 m.

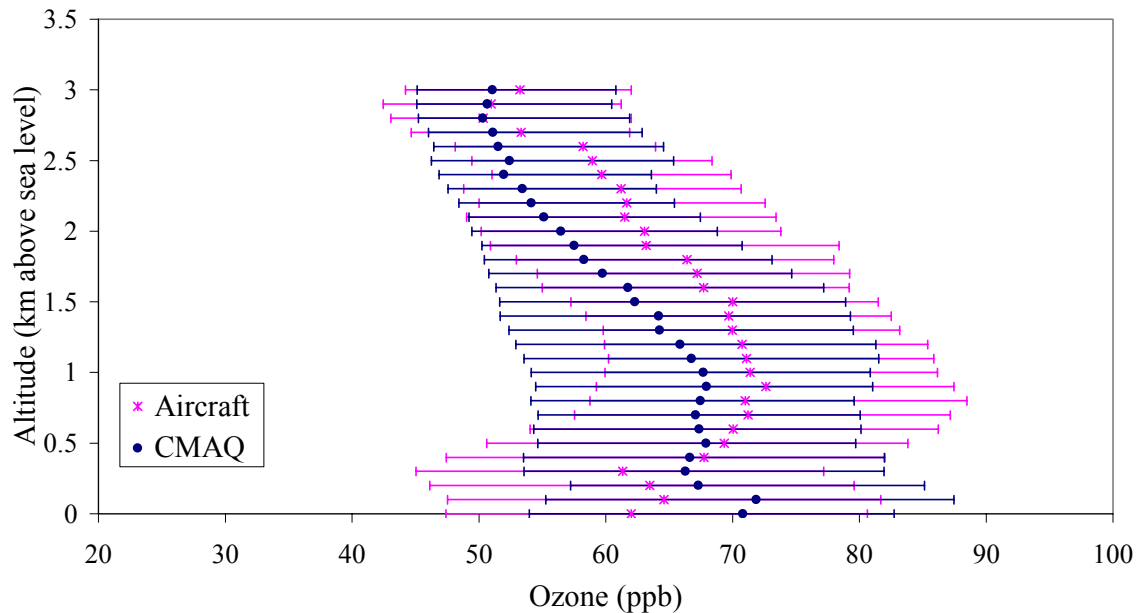


Figure 4. Median CMAQ and aircraft O₃ profiles from 2002 (June–August, 136 profiles). The ends of the horizontal bars represent the 25th and 75th percentiles.

Grand statistical averaging (i.e. over an entire year or ozone season) is helpful in identifying major problems associated with performance, (e.g. a large error in an emissions inventory or improper land surface use) but less helpful from a forensic perspective when subtle errors or compensating errors are present that may change from ozone episode to ozone episode. Types of error that may evade inspection by grand statistical averaging, include biases in the timing and spatial coverage of convection and/or biases in the timing and extent of planetary boundary layer ventilation. To further elucidate CMAQ performance during specific ozone episodes, several case studies were performed. For organizational purposes, selected case studies include periods when CMAQ performance was characterized, based on aircraft data, as *below-median* (the 95th percentile differences as calculated by equation 1), *average* (the 50th percentile differences) and *above-median* (the 5th percentile differences). For each case, surface layer ozone distributions and vertical profiles are compared to measurements. In addition, 24-hour back trajectories were calculated from 40 km North American Weather Model (NAM) data using modeled wind velocities from HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory, [Draxler, 1991]). To complement this analysis, general weather conditions will be discussed to elaborate on the history of air parcels in the context of air quality. Through presentation of CMAQ performance during individual ozone episodes (occurring over the Baltimore-Washington, D.C. non-attainment areas), the ability of CMAQ to replicate conditions during “base case” emissions conditions is evaluated.

Stoeckenius and Kemball-Cook (2005) developed a classification scheme and applied it to all OTR ozone episodes during the summer of 2002. They used a clustering

technique to classify ozone episodes based on spatial ozone patterns and weather conditions (wind direction, temperature, etc.). Table 3 below is a guide to episode types and composite ozone patterns determined in their work. To better identify the episodes discussed, the episode type as classified by Stoeckenius and Kemball-Cook are reported in this study.

Table 4 is a summary table of the ten flights selected for closer inspection in the following section. The ten flights contain three instances each of occasions when the CMAQ performance (based on the definition of Hains et al., 2006) is classified as *above median* and *median*, and four flights for which ozone performance was *below median*. Twice, multiple flights were performed on the same day (June 25 and August 2). CMAQ performance during each flight is evaluated separately. Table 5 and Figure 5 provide summary data regarding aircraft spiral and airport locations.

Table 3. Composite Pattern and Episode Type determined by Stoeckenius and Kemball-Cook (2005).

Composite Pattern	Episode Type
3	Type A: High ozone throughout the OTR
2	Type B: High ozone confined to extreme southeastern OTR
5	Type C: High ozone along I-95 corridor and northern New England
4	Type D: High ozone in the western OTR
1	Type E: Generally low ozone throughout the OTR

Table 4. Summary of CMAQ Comparisons of ozone vertical profiles to aircraft ozone vertical profiles

CMAQ Performance	Flight #	Date/Time	Spiral Location	Composite ozone Pattern
Above Median	RF09	06/11/02 13:00 UTC	Louisa, VA	5
Above Median	RF15	06/25/02 19:00 UTC	Churchville, MD	4
Above Median	RF41	08/02/02 15:00 UTC	Cumberland, MD	4
Median	RF29	07/16/02 20:00 UTC	Tappahannock, VA	2
Median	RF30	07/17/02 14:00 UTC	Crewe, VA	3
Median	RF10	06/11/02 20:00 UTC	Ashland, VA	5
Below Median	RF42	08/02/02 19:00 UTC	Fort Meade, MD	5
Below Median	RF14	06/25/02 14:00 UTC	Winchester, VA	4
Below Median	RF17	08/12/02 18:00 UTC	Bennington, VT	4
Below Median	RF49	08/13/02 14:00 UTC	Morrisville, VT	3

Table 5. Reference Guide for Aircraft Spiral Locations

ID #	Spiral	Airport	Town	Latitude	Longitude	Elevation*
1	LKU	Louisa Co. Freeman	Louisa, VA	38.01°N	77.97°W	150
2	0W3	Harford Co.	Churchville, MD	39.56°N	76.20°W	121
3	CBE	Cumberland Regl.	Cumberland, MD	39.62°N	78.76°W	237
4	W79	Tappahannock Mun.	Tappahannock, VA	37.92°N	76.87°W	10
5	W81	Crewe Mun.	Crewe, VA	37.18°N	78.10°W	131
6	0FP	Hanover Co. Mun.	Ashland, VA	37.71°N	77.44°W	62
7	FME	Tipton AFB	Fort Meade, MD	39.09°N	76.76°W	42
8	OKV	Winchester Regl	Winchester, VA	39.14°N	78.14°W	222
9	DDH	William Morse St.	Bennington, VT	42.89°N	73.25°W	252
10	MVL	Morrisville-Stowe	Morrisville, VT	44.54°N	72.61°W	223

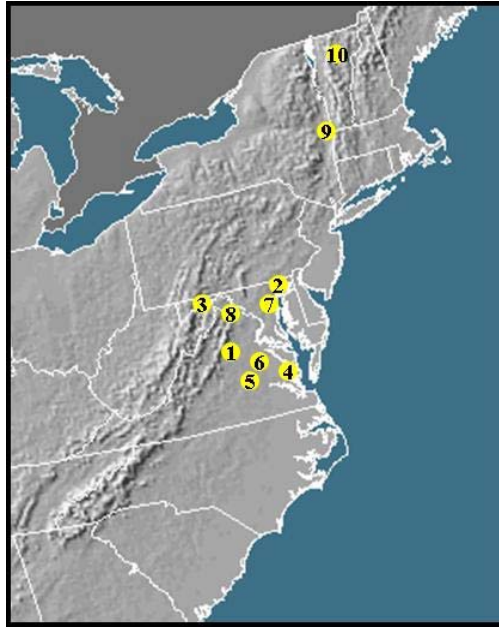


Figure 5. Airport spiral locations of ozone profiles collected by the UMD aircraft presented in this analysis. * Elevation in meters above Mean Sea Level.

Figure 6 is a legend for HYSPLIT trajectories and aircraft/CMAQ ozone profiles presented in the subsequent pages. For each case study, a back trajectory is calculated using the HYSPLIT program to present a history of the air entering the area where an aircraft spiral was performed. Back trajectories for all cases extend 24-hours, and are performed at 500, 1000 and 1500 meters. These levels were selected to obtain information about the boundary layer (500, 1000 meters) and lower free troposphere (1500 meters). The 24-hour duration was selected to minimize model (HYSPLIT) uncertainties that can result from random errors contained in the numerical weather model used to determine the back trajectories, which may propagate when longer back-trajectories are employed. Figure 6 also provides a legend called “Ozone Data” for panels showing a comparison of aircraft profiles (pink stars) against CMAQ profiles (blue diamonds). Occasionally, CMAQ profiles will show discontinuities of 5-10ppbv that repeat back and forth throughout the profile. These discontinuities occur at sampling locations near the edges of CMAQ grid boxes. At these locations, slight changes in horizontal location result in sampling from a different CMAQ grid box.

Legend for HYSPLIT Back Trajectories and Aircraft/CMAQ Ozone Profiles.

The guide below will help orient the reader to a series of figures associated with aloft comparisons between ozone observations and CMAQ simulations.

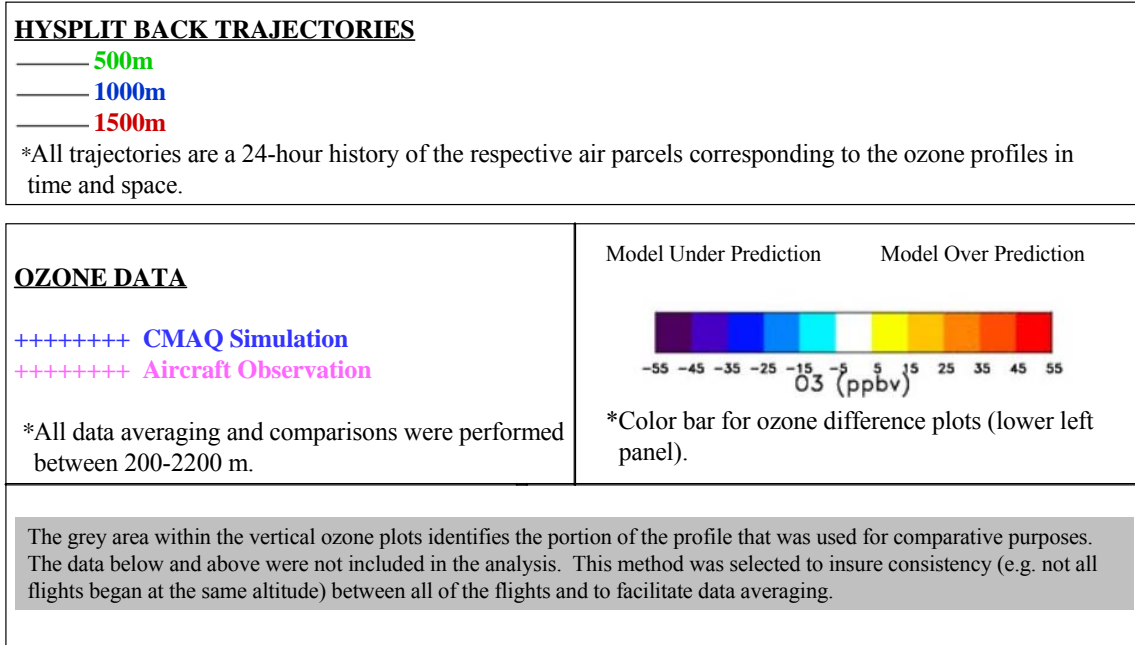


Figure 6. Guide to viewing HYSPLIT back trajectories, aircraft profiles and surface ozone difference plots.

Figure 7 is a sample 4-panel surface plot to familiarize the reader with subsequent figures that will be presented. The upper left panel is a map of observed 8-hour maximum ozone over the OTR, the upper right panel shows CMAQ-calculated 8-hour maximum ozone interpolated to ozone measurement sites within the OTR, and the lower left panel is a difference plot created by subtracting the CMAQ-calculated ozone (upper right) from measured ozone (upper left). The lower right panel shows CMAQ-calculated 8-hour ozone on the original model grid (meaning that the ozone values have not been interpolated to the measurement sites). Comparison of the upper-right and lower-right hand plots shows that interpolation of model output onto the OTR domain causes minimal loss of information. For simplicity, subsequent figures will exclude the plot showing model output on the original grid.

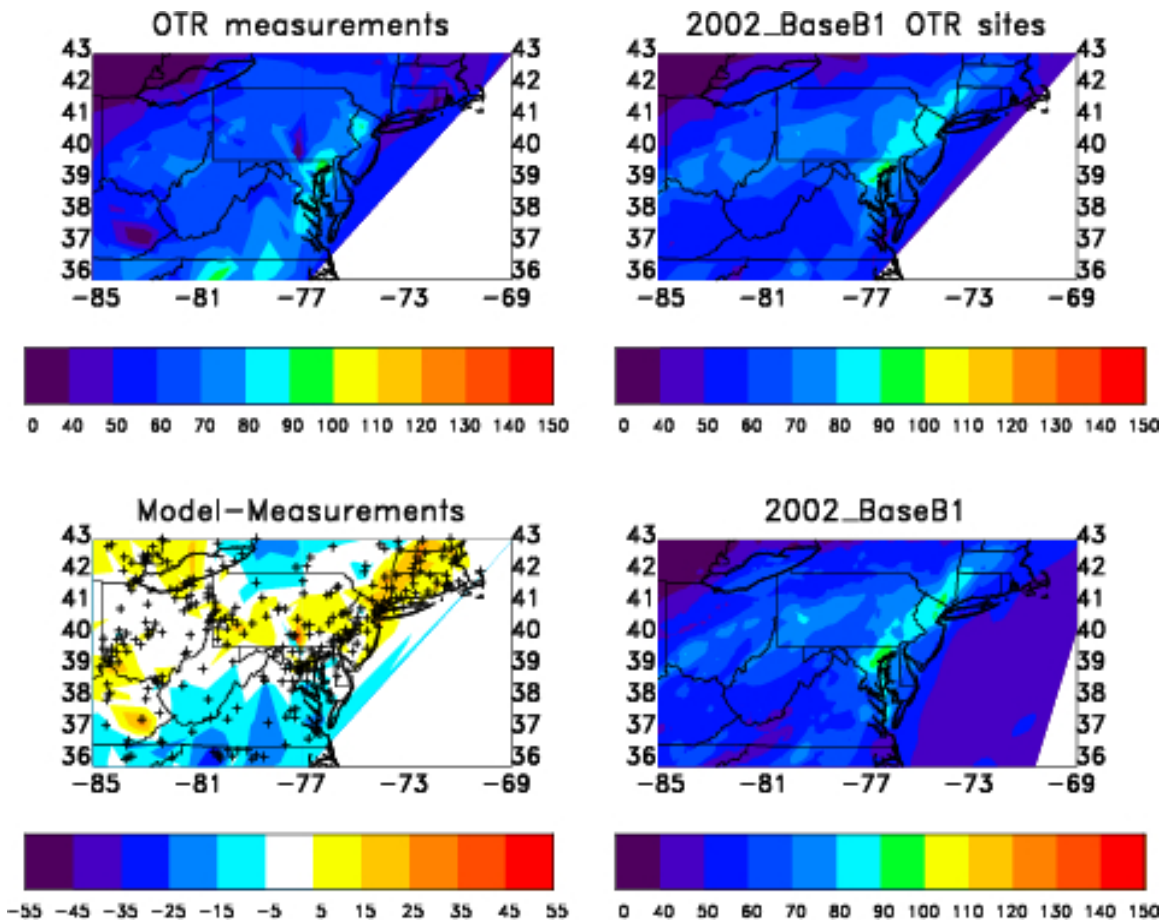


Figure 7. Sample plots for June 5, 2002. Latitude and longitude coordinates are provided on the x axis and y axis of each plot. The upper left plot shows the observed daily 8-hour ozone maxima (ppbv). Ozone (ppbv) from the 2002 base B1 simulation interpolated to the locations of ozone monitors is shown in the upper right panel. The difference between measurements and model outputs is in the lower left panel in ppbv, and the 2002 base B1 simulation on the original model grid is shown in the lower right panel, also in ppbv. Since the upper right plot is almost entirely the same as the lower right plot, the upper right plot is not be included in subsequent surface ozone comparison plots.

Above Median CMAQ Performance Cases in 2002

Tuesday, 06/11/2002, 13:00 UTC Case Study

June 11 was classified as a type 5 episode (high ozone along the I-95 corridor and northern New England). Figures 8a-b show back trajectories ending on June 11 over northern VA (Louisa County) and a corresponding aircraft profile collected on the morning (09:00 EST) of the 11th, respectively. The back trajectory showed that winds were light and that the air mass was over southwestern VA and eastern TN 24-hours earlier. CMAQ performance is almost uncanny on this occasion, with CMAQ accurately replicating the boundary layer ozone morphology as near-surface minima of ~ 55 ppbv increase to near 85 ppbv by 300 m. It appears in this case, based on the vertical ozone profiles and back trajectories, that the boundary layer has yet to vent. Thus, the origin of the observed pollution aloft was likely from transport rather than from local emissions. The back trajectories show a shift from a westerly flow to a flow along the I-95 corridor. Figure 9 compares surface ozone data on the 11th to a CMAQ simulation. CMAQ accurately represents ozone over much of PA and MD but generally underpredicts ozone concentrations elsewhere with peak underpredictions (~ 30 ppbv) occurring over central OH and the southern extreme of the modeling domain.

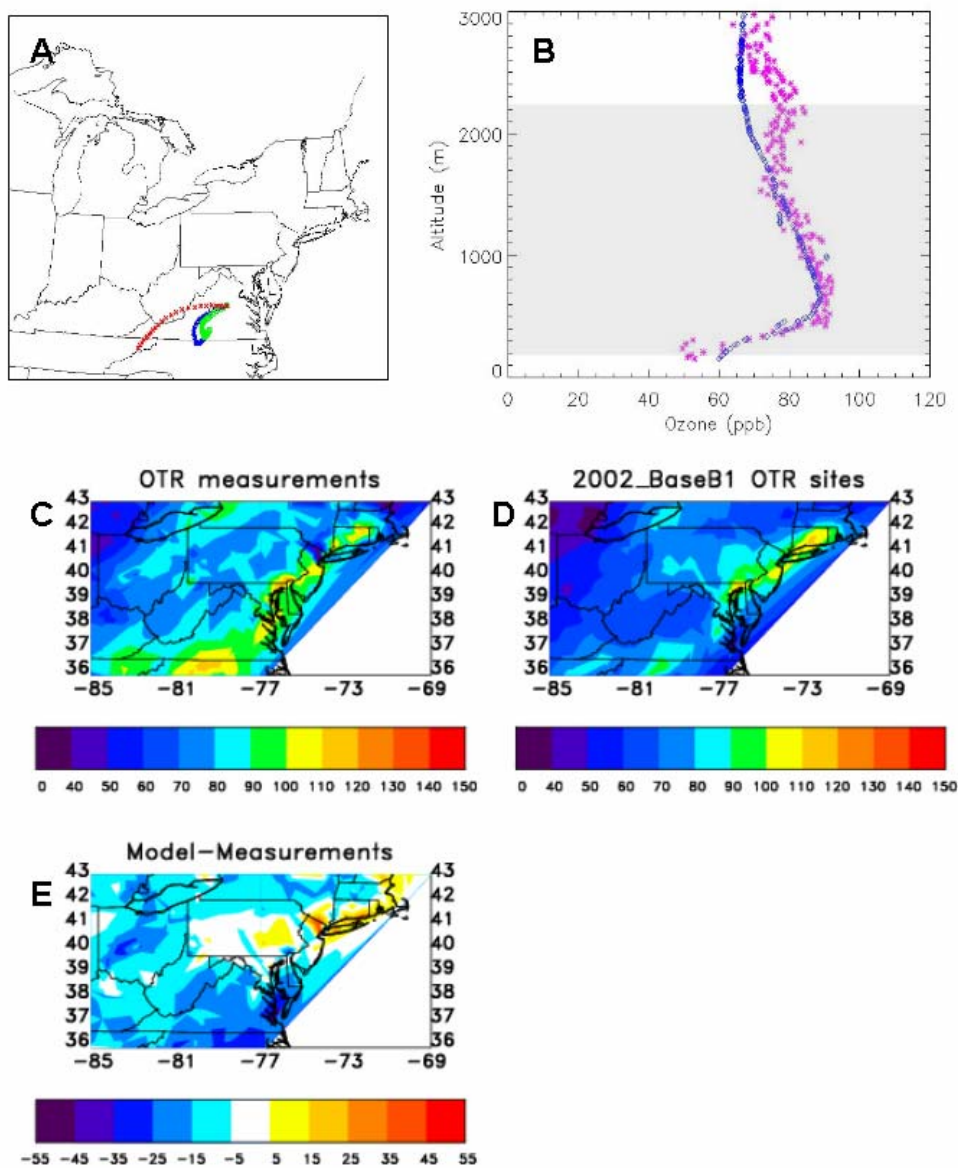


Figure 8. Tuesday, June 11, 2002, 13:00 EST, Louisa, VA Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Louisa, VA.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Tuesday, 06/25/2002, 19:00 UTC Case Study

June 25 was classified as a type 4 episode (high ozone in the western OTR). Like June 24, the highest ozone concentrations on June 25th were confined to a box south and west of New York City and north of VA, with widespread exceedances in the Ohio River Valley and scattered Code Red concentrations along the Washington-Philadelphia Corridor. Figure 10a shows back-trajectories terminating over northeastern MD in the early afternoon (15:00 EST) of June 25 and Figure 10b shows a corresponding aircraft profile. The back trajectory originates from eastern OH at the lowest levels and extends to near Detroit, MI farther aloft. Like the vertical profile from June 11th (Figure 8b) the boundary layer appears to be stratified, with minimum ozone values at the surface of ~70 ppbv increasing to a well-mixed layer of ~95 ppbv from 500 m to 2500 m. This is an interesting comparison because this is a case where CMAQ performance was defined as good, but inspection of the lowest portion of the aircraft and CMAQ profile suggests otherwise. CMAQ is overpredicting ozone from near the surface up to ~300 m, and only then begins to simulate ozone accurately throughout the rest of the profile. As noted previously, for consistency (and to accommodate for changes in elevations of ozone monitors), CMAQ and aircraft data were compared from 250 m to 2150 m, and therefore, the lowest portion of the boundary layer was not included in this comparison. To further complicate this analysis, it appears, based on surface ozone maps, that ozone was spatially highly variable in the area surrounding the aircraft profile (Churchville, MD). Thus CMAQ performance at the lowest levels, which initially did not seem favorable, is actually (on a slightly larger scale) reflective of the variable spatial ozone pattern in the area. This notion is supported when comparing modeled and observed surface ozone over the entire OTR (Figure 11) and underscores some of the challenges encountered when making aircraft comparisons. Clearly, subjective analysis is needed in complex cases such as this one.

CMAQ continued to underpredict peak values, but model performance improved compared to earlier days in this episode when underpredictions were severe. Differences between the model and observations are less than 15 ppbv over much of PA, OH and NJ, with differences in VA of ± 5 ppbv. There are multiple areas of model overprediction (~15-25 ppbv) found in western MD-eastern WV-northern VA, southwestern and central OH and northern NJ-Long Island-Boston, MA.

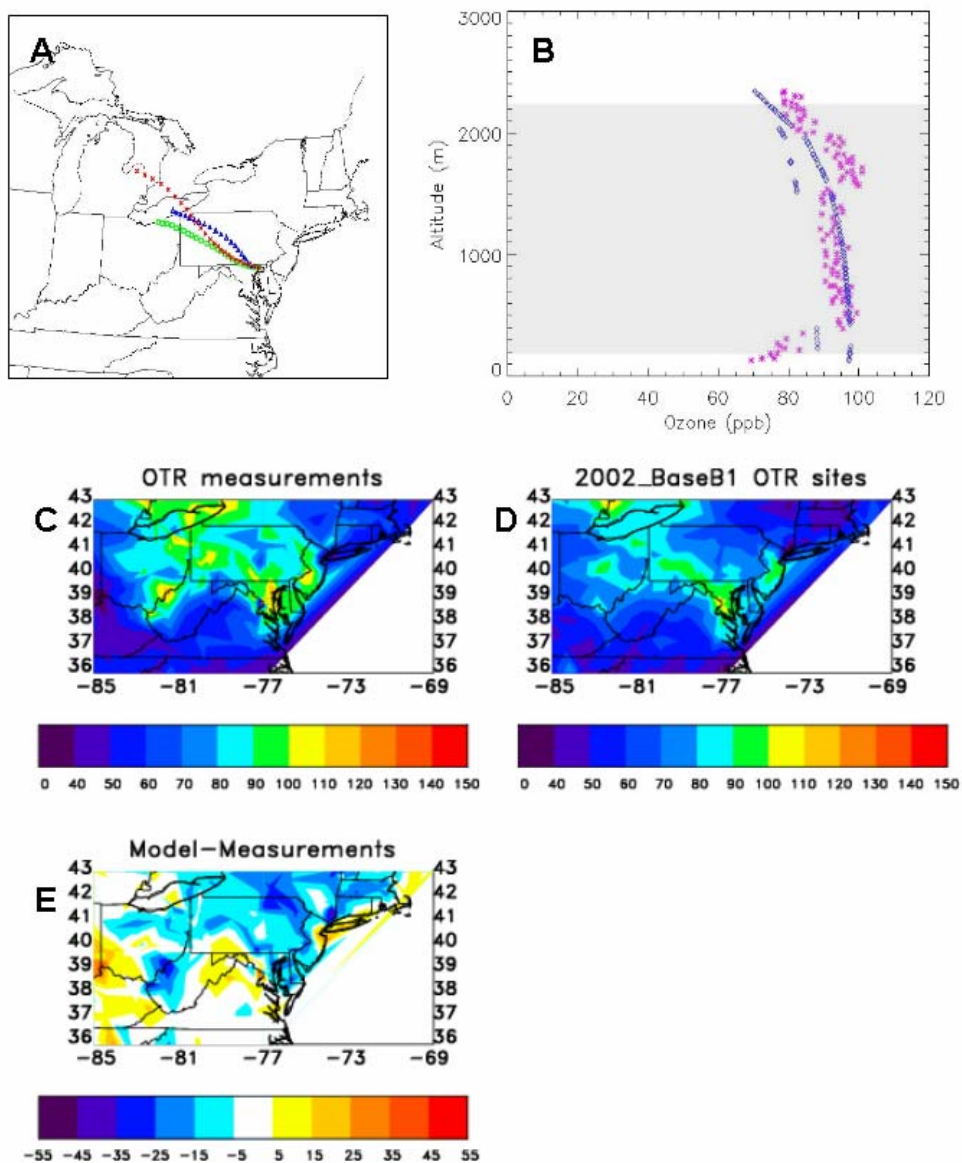


Figure 9. Tuesday, June 25, 2002, 14:00 EST, Churchville, MD. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over , Churchville, MD.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Friday, 08/02/2002, 19:00 UTC Case Study

August 2 was classified as a type 4 episode (high ozone in the western OTR). Figure 12a indicates a 24-hour back trajectory terminating over western MD that originated over western NY and southern Ontario. Figure 12b shows excellent agreement between CMAQ and aircraft observations except near 500 m where the aircraft encountered a thin layer of relatively clean air that CMAQ did not resolve. Vertical profiles show relatively constant, low ozone with values ranging from ~40 to 60 ppbv. Peak ozone values are located over southern OH, central PA and along the I-95 Corridor. CMAQ does not capture the very low ozone values over eastern KY, registering a high bias of 40 ppbv. To the north, CMAQ overpredicts (by 15-25 ppbv) over much of the southern New England Region. In general, CMAQ replicates the pattern, but underpredicts the magnitude of the area of high ozone over eastern OH, western PA, and along the I-95 corridor (Figure 13). However, at individual sites within a region, CMAQ sometimes performs differently. For example, on August 2, CMAQ overpredicts ozone amounts near Edgewood, MD (a site within the I-95 corridor). These small-scale differences in model performance are not surprising given small-scale variations in terrain and circulation (e.g., Edgewood is located near the Chesapeake Bay and is often subject to land-sea breezes). However, these differences do indicate that no single site should be given extra focus when evaluating model performance. For most applications, this is not a problem, but it is a concern for SIP modeling, since model performance at the worst air quality sites within a nonattainment area is relevant when determining if emission reductions are sufficient to bring a non-attainment area into compliance with the 8-hour ozone standard.

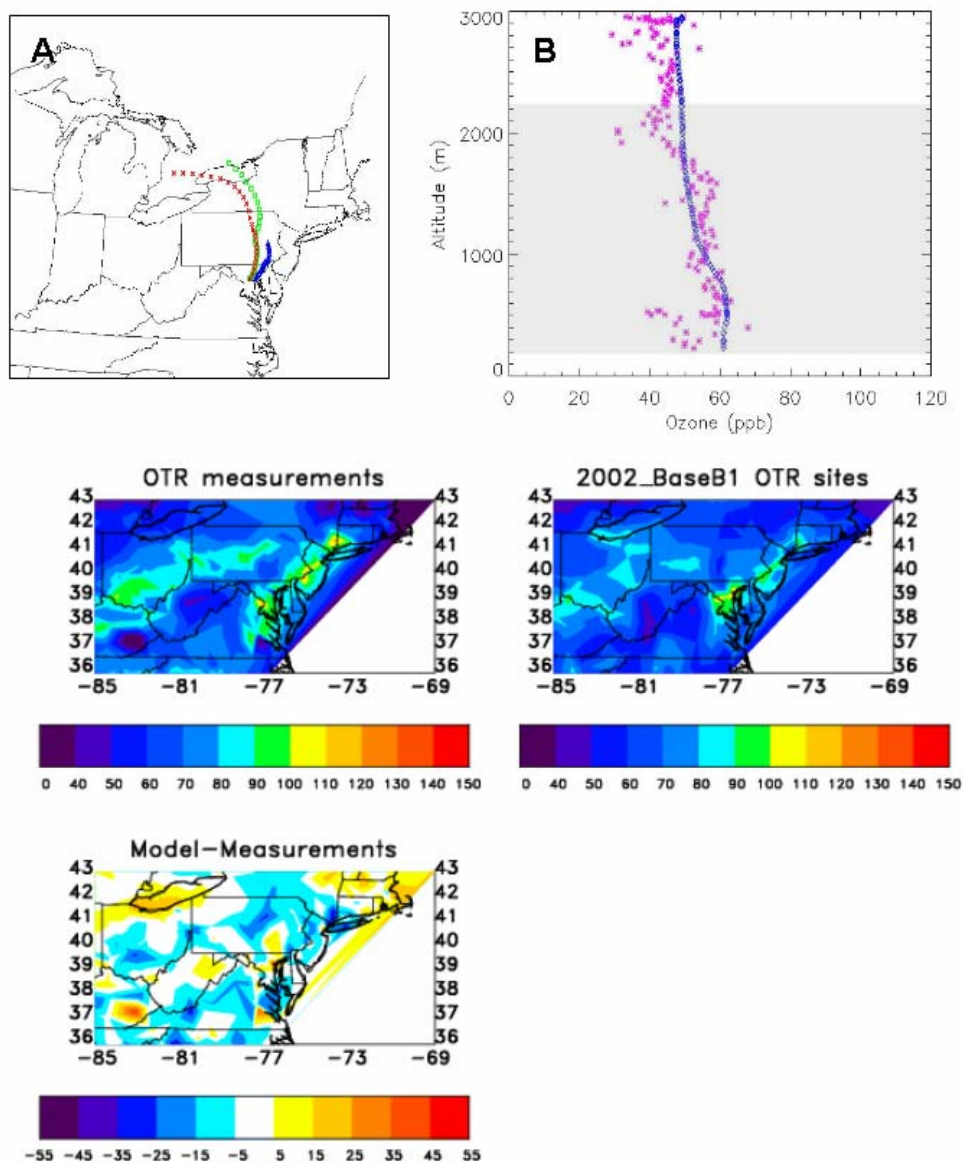


Figure 10. Friday, August 02, 2002, 10:00 EST, Cumberland, MD. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Cumberland, MD.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Median CMAQ Performance Cases in 2002

Tuesday, 07/16/2002, 20:00 UTC Case Study

July 16 was classified as a type 2 episode (high ozone confined to the extreme southeastern OTR). Figure 14a shows relatively fast 24-hour back trajectories (ending over Tappahannock, VA at 15:00 EST) extending northwest into northern MI and central Ontario. Where ozone levels were quite low (~40 ppbv), Figure 14b reveals CMAQ values twice as high as surface observations. As noted in the June 25 analysis (see Figure 6b), statistical comparisons begin at 250 m, and therefore, this “poor” comparison is not included. Starting at 500 m and looking upwards throughout the rest of the profile, CMAQ agrees with observations within 10 ppbv, with CMAQ showing consistently lower concentrations. From Figure 13 it can be seen that peak ozone levels on the 16th occurred along a line from the OH-KY border east-southeast across WV, VA and NC. Most of those values were in excess of 85 ppbv, with levels reaching 105 ppbv near Cincinnati, OH. On average the model performance appears to be good with differences over the domain mostly below 15 ppbv. The largest discrepancy is associated with the peak value near Cincinnati, OH; the local maximum that CMAQ places in that region is too low and displaced south of the actual peak. The result is an area of underprediction over southwestern OH and a small area of overprediction in northern KY. As has been noted, in other comparisons, CMAQ tends to overpredict (15-25 ppbv) surface layer ozone along the Ohio River and over portions of WV.

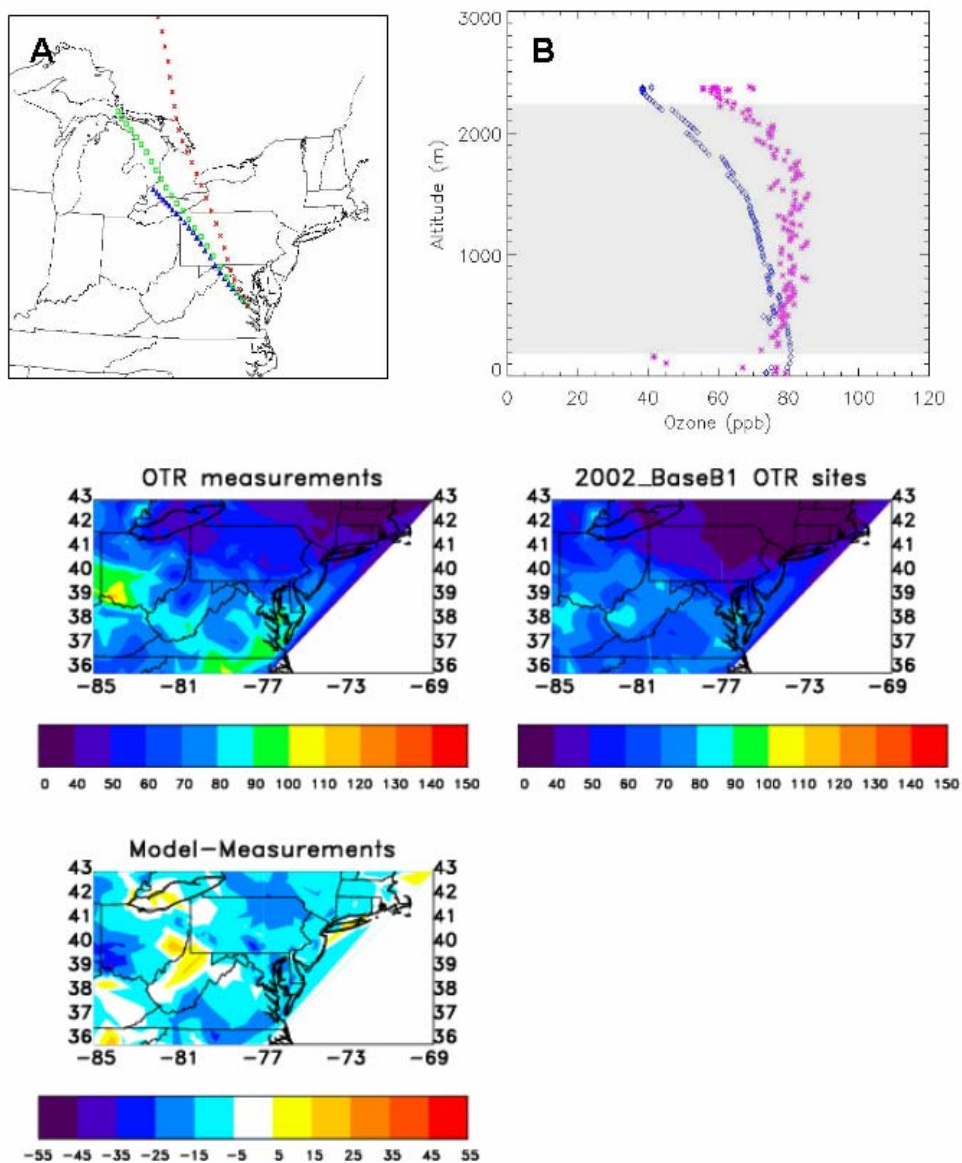


Figure 11. Tuesday, July 16, 2002, 15:00 EST, Tappahannock, VA. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Tappahannock, VA.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Wednesday, 07/17/2002, 14:00 UTC Case Study

July 17 was classified as a type 3 episode (high ozone throughout the OTR). Like the previous day, 24-hour back trajectories from central VA ending at 10:00 EST were fast (Figure 16a), but more northerly in nature, passing over central PA, western NY, and Lake Ontario. Figure 16b shows that ozone values begin fairly low (~40 ppbv) near the surface and increase to just above 50 ppbv by 600 m. Observations increase to ~80 ppbv by 1100 m and then tail off to a constant value of ~60 ppbv by 1500 m. On this morning CMAQ does an excellent job of replicating relatively low ozone levels near the surface, but it struggles to simulate the structure of ozone between 900m and 1200m. Instead, the ozone remains constant above 900 m. This is a case when it appears that CMAQ is not capturing a regional transport signal. Figure 17 shows that 8-hour ozone mixing ratios exceeding 85 ppbv were reported from SC to as far north as CT and westward into OH. Scattered peak 8-hour ozone concentrations exceeding 105 ppbv occurred along and east of the I-95 Corridor as well as near Cleveland and Buffalo. Surface model performance appears to be slightly less accurate than on the prior case study of July 16th, with the majority of differences in the range of 15 to 25 ppbv. Of note is the large area of underprediction covering much of the northern two thirds of OH and portions of southern NY, especially near the Canadian border.

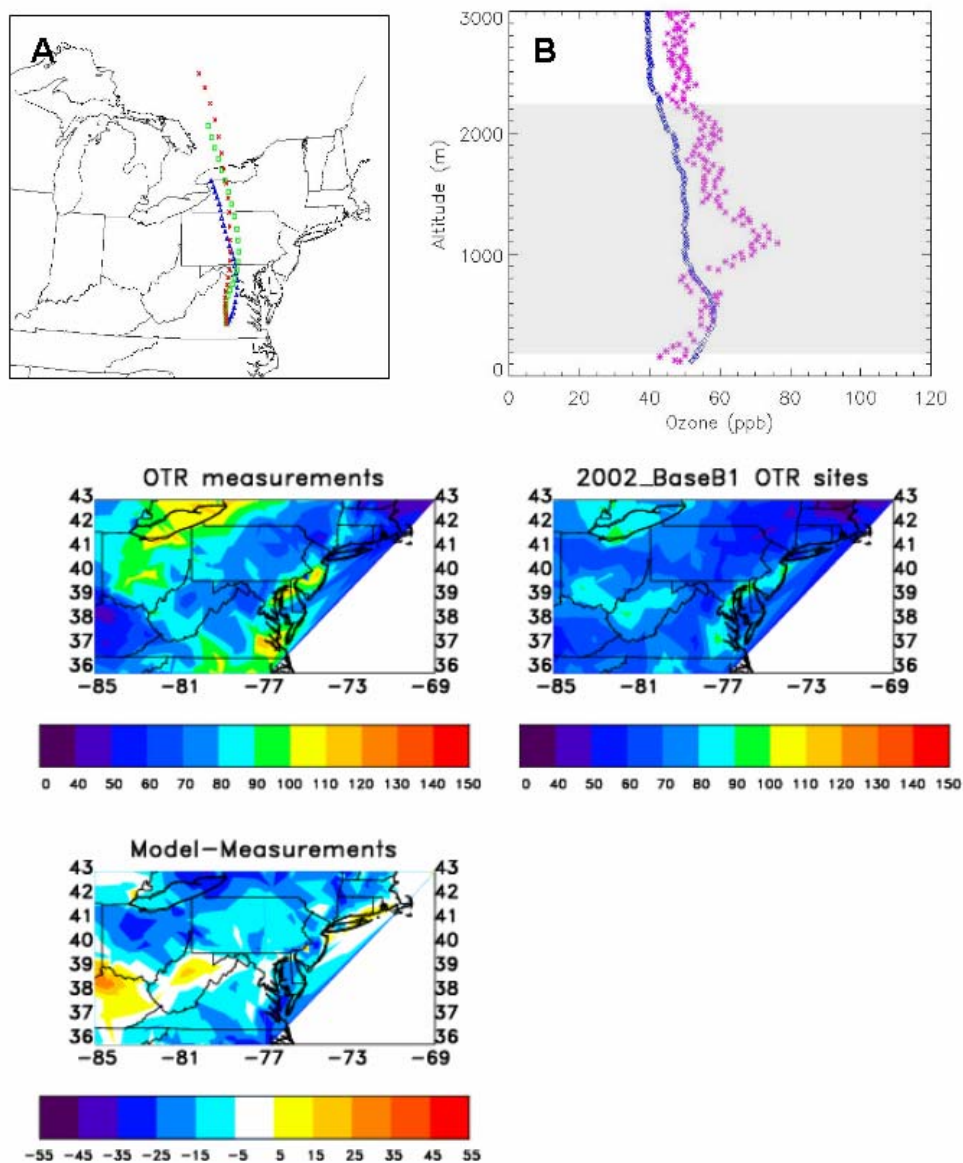


Figure 12. Wednesday, July 17, 2002, 09:00 EST, Crewe, VA. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Crewe, VA.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Tuesday, 06/11/2002, 20:00 UTC Case Study

June 11 was classified as a type 5 episode (high ozone along the I-95 corridor and in northern New England). Recall that CMAQ-calculated and observed ozone agreed well during a morning flight on June 11th. The flight addressed here is an afternoon flight at a different location (see Table 4). Figure 18a shows weak westerly winds that advected air parcels originating over southwestern VA and the central VA/NC border. Figure 18b indicates that CMAQ captures the general shape of the ozone profile nicely (i.e. fairly flat throughout the boundary layer) but underpredicts ozone throughout the profile by 10-20 ppbv. Given that this is an afternoon (16:00 EST) profile it is possible that CMAQ has vented the boundary layer too early or too vigorously which could account for the differences between CMAQ and observations. The synoptic situation included a near-stationary frontal boundary just northeast of Baltimore, MD at 23:00 EST on the June 10. By 8:00 EST the front had moved rapidly north to a location just south of New York City. The boundary, now in the nature of a retreating warm front, moved well into New England at 14:00 EST with an Appalachian lee-side trough analyzed from Baltimore, MD to northern NC. The preceding 3 sentence's explanation of the synoptic situation are not tied to the observations or CMAQ – please tie in / make relevant, or remove. Figure 19 shows high ozone concentrations reported along the I-95 Corridor, with scattered locations showing values exceeding 105 ppbv from central NC to northern CT. The most concentrated area of high ozone is in a band from just northeast of Baltimore into southeastern PA between New Castle, DE and Lancaster, PA. The back trajectories show a shift from a west-northwest flow to one along the I-95 corridor.

CMAQ underpredicts ozone concentrations on June 11 over much of the model domain, with the greatest underpredictions (~ 30 ppbv) occurring over portions of western PA and southwestern VA. Along the I-95 corridor, just north of Baltimore, MD extending to near Boston, MA CMAQ overpredicts ozone levels by 5-15 ppbv.

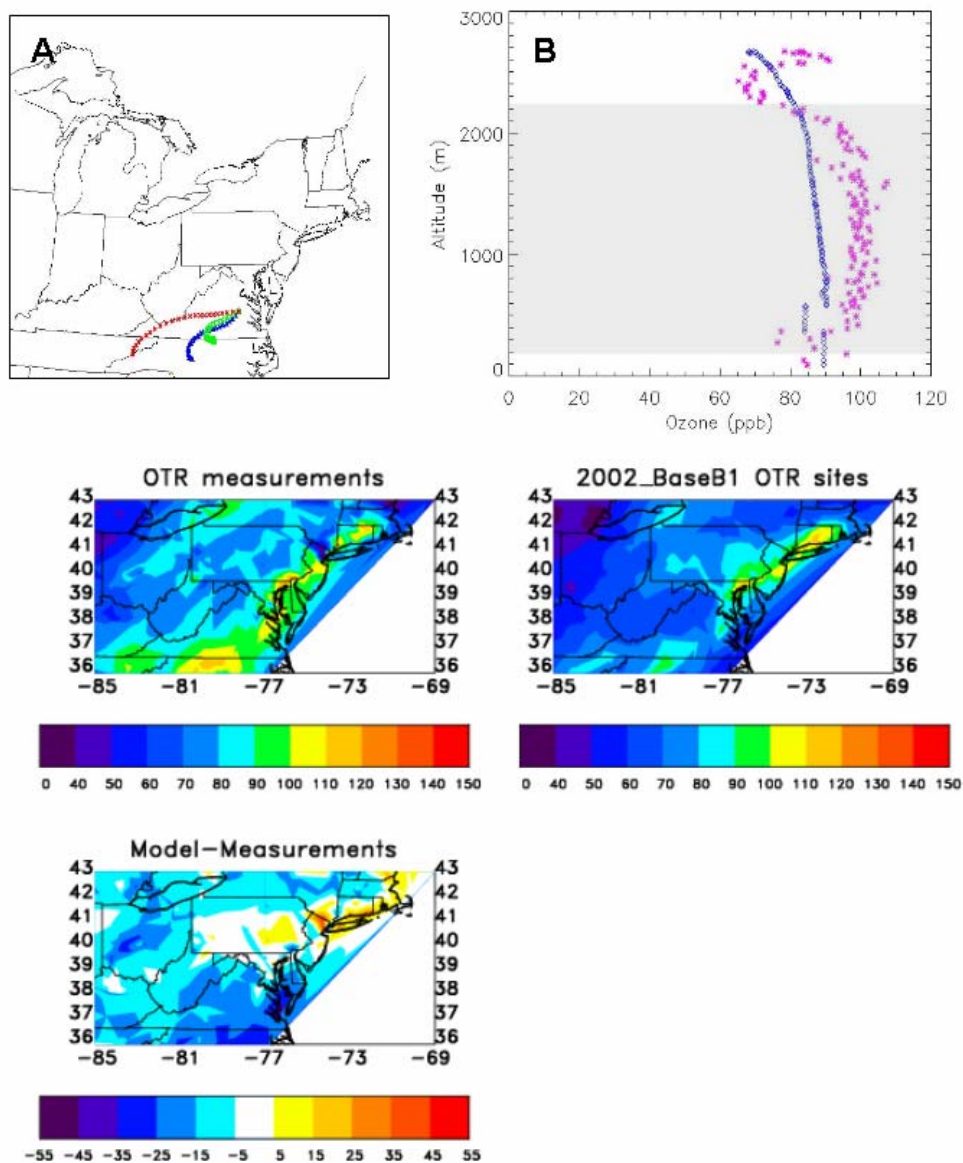


Figure 13. Tuesday, June 11, 2002, 09:00 EST, Ashland, VA Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Ashland, VA.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Below Median CMAQ Performance Cases in 2002

Friday, 08/02/2002, 19:00 UTC Case Study

August 2 was classified as a type 4 episode (high ozone in the western OTR). Figure 20a shows that 24-hour back trajectories ending over Fort Meade, Maryland in the afternoon (15:00 EST) show a high degree of variability. Recall that CMAQ-calculated and observed ozone agreed well during a morning flight on August 2. The flight addressed here is an afternoon flight at a different location (see Table 4). At the upper levels, air parcels originated over southern Ontario, Canada but back trajectories only extend to southeastern PA at the lowest levels. Figure 20b shows CMAQ greatly overpredicted ozone, with simulated values (~115 ppbv) well mixed throughout the boundary layer. Aircraft observations show slightly more variability, but only range from ~80 ppbv at the lowest levels to ~60 ppbv by 1500 m, before falling off to less than 40 ppbv. Figure 21c & d show that CMAQ is actually doing a reasonable job replicating the spatial pattern of ozone, especially along the I-95 corridor. The high variability in the observed surface maxima creates pockets of relatively large disagreement over small areas. The aircraft spiral was performed over one of those pockets. It is possible, given the proximity of the spiral to the Chesapeake Bay, that land-sea interactions may have affected the simulation locally. Peak ozone values appear over southern OH, central PA, and throughout the Interstate-95 Corridor. Isolated areas of CMAQ overprediction occur over eastern KY (~40-50 ppbv), the Philadelphia-Baltimore-Washington region, and much of southern New England (≤ 25 ppbv).

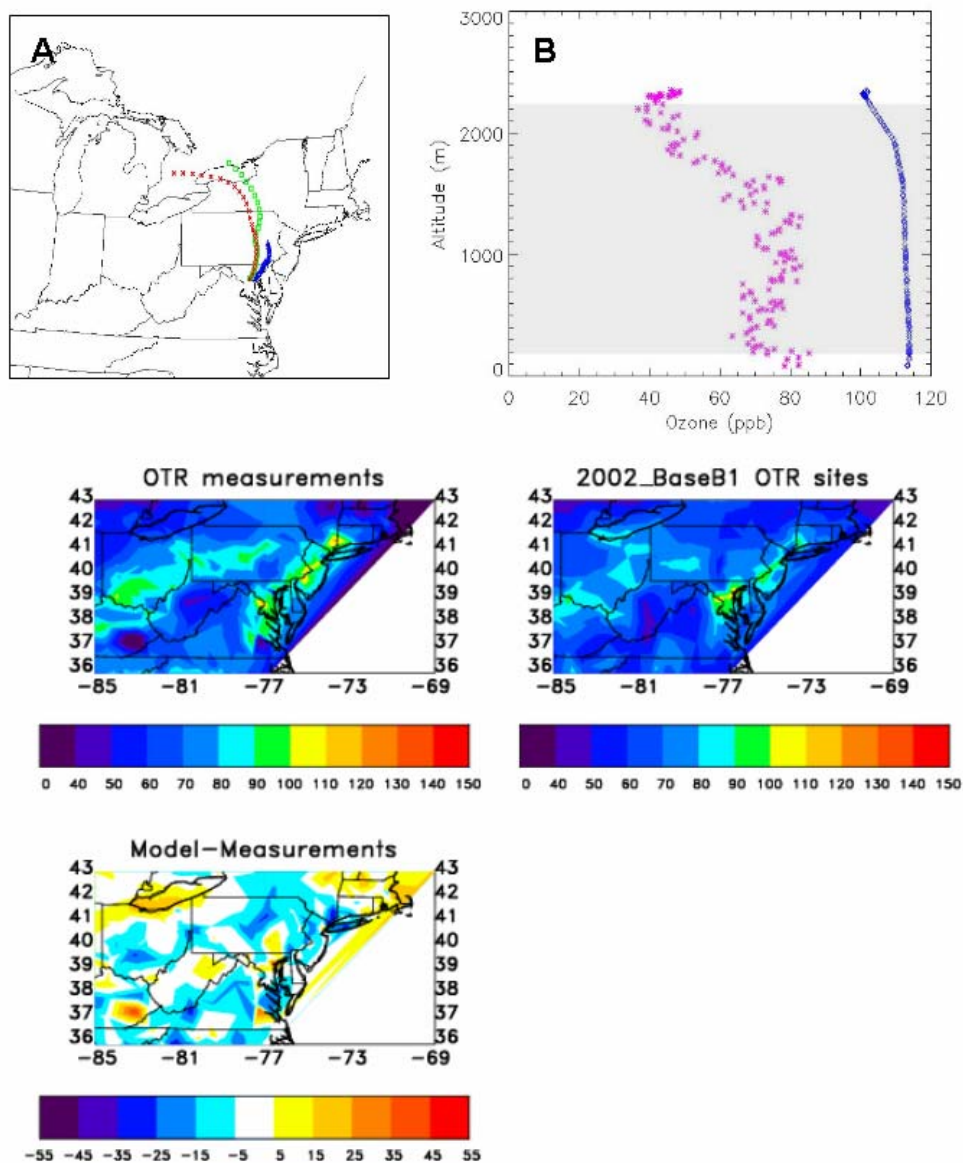


Figure 14. Tuesday, August 2, 2002, 14:00 EST, Fort Meade, MD. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Fort Meade, MD.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Tuesday, 06/25/2002, 14:00 UTC Case Study

June 25 was classified as a type 4 episode (high ozone in the western OTR). Figure 22a is a back trajectory for the morning of June 25 (10:00 EST) over Winchester VA. The trajectory is weak from the northwest reaching only southwestern PA and eastern OH. Recall that an afternoon (15:00 EST) flight on the 25th (see Figures 12a and 12b) was classified as a good CMAQ performance day. Analysis of the morning flight presented in this case study reveals that CMAQ overpredicted ozone aloft (~200 m through 700 m) by ~20 ppbv (Figure 22b). At 700 m, aircraft observations sharply increase to match CMAQ simulations of ~90 ppbv; farther aloft, the two measurements diverge again, with CMAQ still higher than observations (by as much as 50 ppbv). Based on surface comparisons from Figure 23 (also Figure 7) over that region, one conclusion is that CMAQ is just missing the location of a local ozone plume (which the aircraft interacts with briefly at ~700 m). Thus, spatially, CMAQ appears to be representing existing conditions with reasonable accuracy. However, because of the sharp ozone gradient over the measurement location and resolution limitations of the model, CMAQ does not compare favorably with aircraft observations.

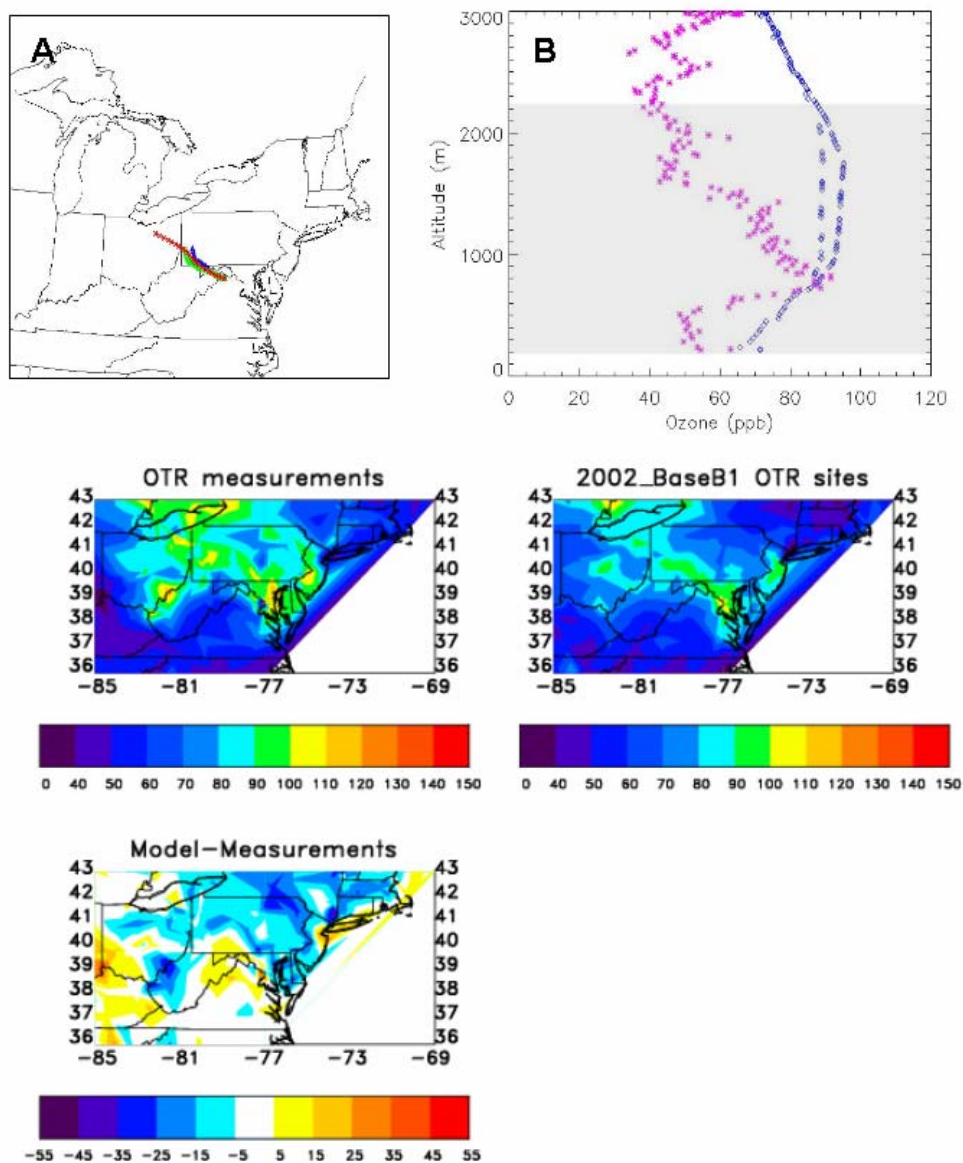


Figure 15. Tuesday, June 25, 2002, 14:00 EST, Winchester, VA. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Winchester, VA.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Monday, 08/12/2002, 18:00 UTC Case Study

August 12 was classified as a type 3 episode (high ozone throughout the OTR). A 24-hour back trajectory shows fast westerly flow ending over Bennington, VT on the 12th of August, originating from over MI and southern Ontario (Figure 24a). A vertical comparison from the early afternoon reveals CMAQ is underpredicting ozone compared to aircraft observations by as much as 40 ppbv in the layer between 200 m and 1000 m (Figure 24b). CMAQ reports a near-constant value of ~60 ppbv from the surface to ~2000 m, while aircraft observations show values closer to 100 ppbv from the ground to ~1000 m. This appears to be a case where CMAQ, rather than nearly missing an ozone plume, is not fully describing the spatial extent of the ozone episode. One possible reason why CMAQ underpredicted ozone is a failure to accurately simulate boundary layer mixing. Inspection of the August 12 synoptic conditions reveals strong, widespread convection developing late in the day. The underprediction of ozone and the apparent well-mixed boundary layer from CMAQ, suggest that convection was initiated earlier in CMAQ (in the parent MM5 simulation) than actually occurred. Another cause of CMAQ underprediction could be a failure to capture westerly transport of ozone and ozone precursors that originated over southern Ontario. Surface comparisons in Figure 25 are consistent with the aloft observations in that CMAQ underpredicts ozone over much of the eastern half of PA, NJ and portions of the southern New England area. (Note: Because of the relatively high latitude of the back trajectory terminations, the horizontal domains of Figures xx and yy have been expanded from 85°-69°W and 36°-43°N to 88°-68°W and 36°-46°N.) In contrast, CMAQ overpredicts ozone over much of Long Island, Coastal MA and southern CT. The locations of the overpredictions (along the East Coast) suggest that CMAQ fails to develop the boundary layer fully over water, which can lead to relatively high ozone values just offshore. This is a potential problem over the Chesapeake Bay. CMAQ is known to have problems with coastal locations, since CMAQ apportions emissions according to multiple land uses within a grid cell, while MM5 assumes that the dominant land use in the grid cell dictates boundary layer development. In a coastal grid cell, MM5 often produces a very tight boundary layer, while CMAQ often puts a lot of emissions (say from a small coastal city) into the very tight boundary layer that MM5 produces. Under these artificial conditions, ozone concentrations can soar unrealistically in the model. It is possible that biases at coastal locations are exaggerated. Ozone monitors are underrepresented at coastal locations, so it is difficult to judge. Biases offshore must be interpreted cautiously as offshore measurements were not available for this comparison.

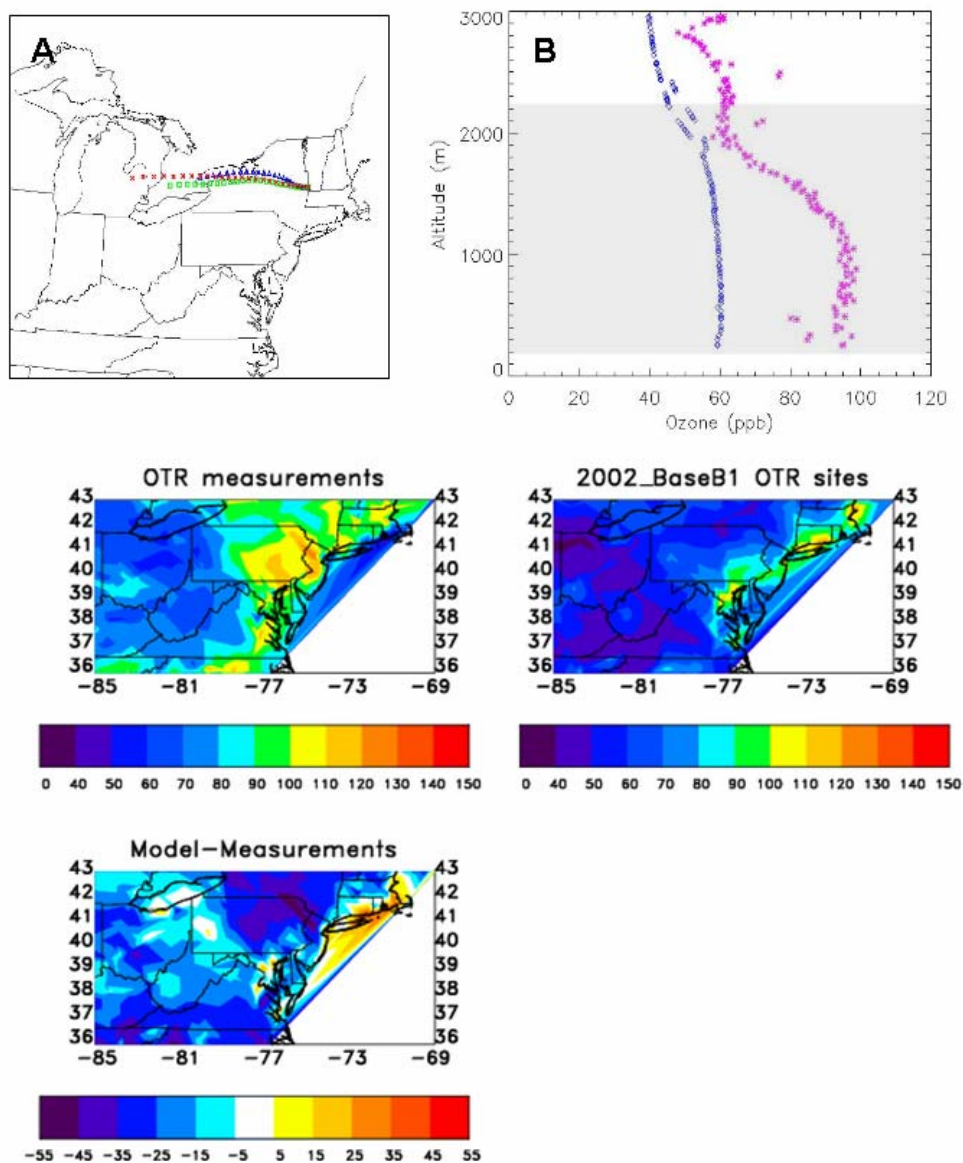


Figure 16. Monday, August 12, 2002, 14:00 EST, Bennington, VT. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Bennington, VT.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Tuesday, 08/13/2002, 14:00 UTC Case Study

August 13 was classified as a type 3 episode (high ozone throughout the OTR). A 24-hour back trajectory shows west-southwest flow ending over Morrisville, VT on the 13th of August, passing over Buffalo, NY and southern Ontario, Canada (Figure 26a). Figure 26b is a late morning (10:00 AM EST) aircraft profile over Morrisville, VT. As on the 12th, CMAQ underpredicts ozone in the lower boundary layer. Aircraft observations show a sharp ozone gradient with near-surface values of ~ 40 ppbv rapidly increasing to over 100 ppbv by ~400 m. The peak ozone levels persist until ~1100 m and then gradually fall off to a constant value of 60 ppbv by 1500 m. The CMAQ simulation is marginally improved from the previous day. Even though CMAQ underpredicts ozone throughout the entire profile it does mimic the vertical morphology of the observations. Unlike August 12th, this profile occurs at ~ 10:00 AM EST, so it is less likely that premature convection in MM5 / CMAQ is the cause of the underprediction. CMAQ performance on August 13 is similar to the 12th in that it is generally characterized by widespread underpredictions (Figure 27). Underpredictions ranged from 15-25 ppbv over much of the domain with extreme excursions of over 50 ppbv occurring in isolated areas of central PA, eastern NY, and portions of the New England region.

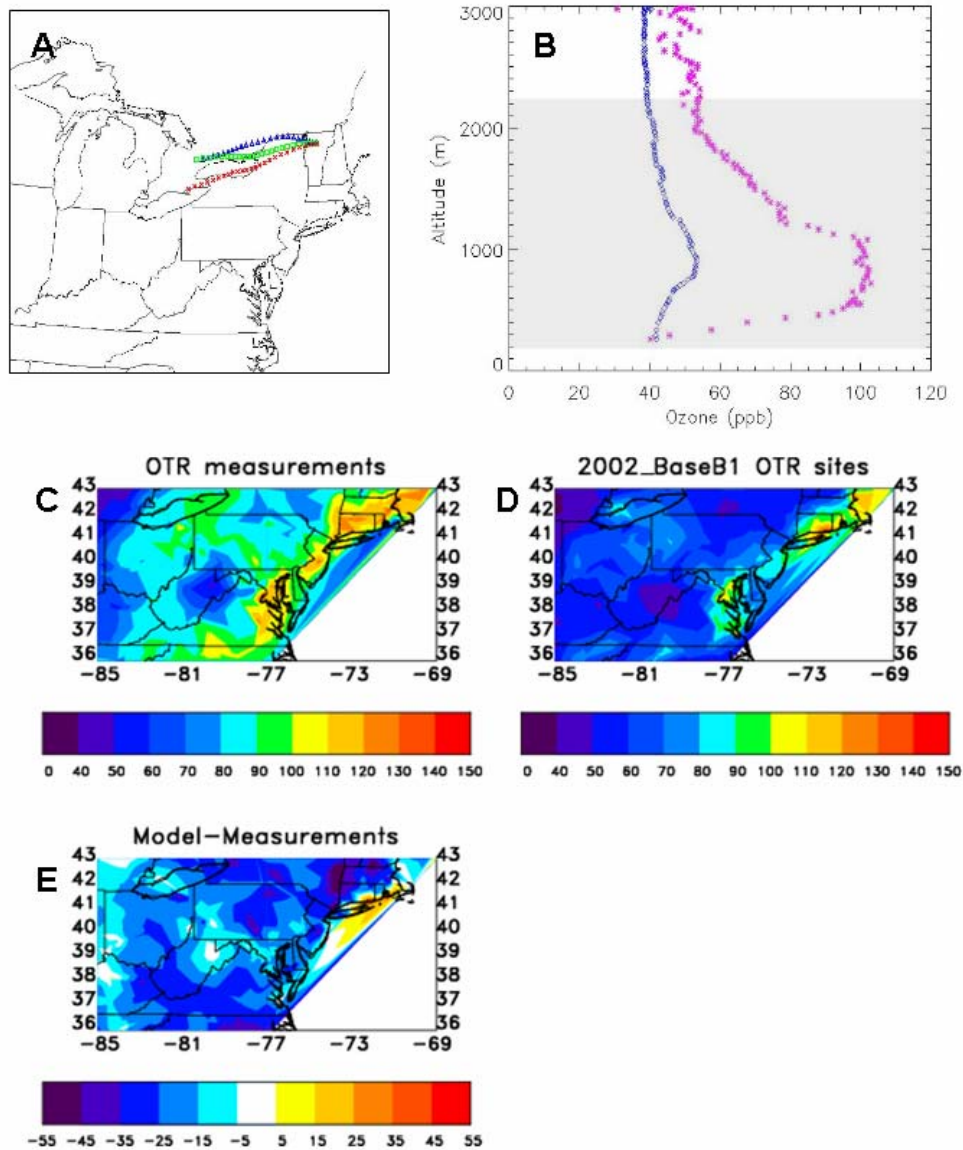


Figure 17. Tuesday, August 13, 2002, 09:00 EST, Morrisville, VT. Case Study
 A. 24-hour HYSPLIT back trajectories terminating over Morrisville, VT.
 Blue = 500m, Green = 1000m, Red = 1500m
 B. Aircraft (pink stars) and CMAQ (blue diamonds) ozone profiles.
 C. OTR surface ozone monitor data.
 D. 2002 base B1 CMAQ simulation averaged for OTR monitor locations.
 E. Difference plot. Negative values indicate model under-prediction.

Conclusions

CMAQ model performance was evaluated for the base year 2002 by comparing surface and aloft CMAQ simulations to corresponding observations from the summer of 2002. Based on these comparisons, it is concluded that CMAQ replicates the spatial patterns of high ozone events. The ability of CMAQ to capture the spatial morphology of ozone events suggests that from a meteorological perspective certain key processes such as synoptic-scale winds and temperature fields are being modeled correctly. In most cases CMAQ does replicate the spatial patterns of an ozone event, but does not fully capture the magnitude or spatial extent of peak ozone values. A comparison of CMAQ to observations from selected surface monitors shows that CMAQ has a low bias of approximately 7 ppbv during pollution events but only ~1 ppbv when averaged over the entire ozone season. Clearly, high-biases during low ozone periods compensate for low-biases during pollution events. As discussed later, in chapter 9, the compensating biases are part of a general tendency of CMAQ to capture the mean ozone concentrations and miss the extreme ozone concentrations. Or to put it another way, CMAQ is adroit at capturing temporal fluctuations in 8-hour ozone (correlation coefficients of 0.8-0.9 are common); however, the amplitude of these fluctuations is generally underestimated. Biases also vary between urban and suburban/rural locations. CMAQ performs has smaller biases at urban locations than it does at more rural locations.

The comparisons of CMAQ with aircraft observations taken on 136 spirals flown in the summer of 2002 (May-September), show CMAQ underpredicts ozone aloft (i.e. between 600-2200 m) by ~10 %. As seen in case studies performed in this report, comparisons between aircraft observations and model simulations can be problematic because of rapidly changing temporal and spatial scales. Comparisons of individual cases highlight the complex nature of ozone predictions and measurements. More generally, the underprediction of aloft ozone by CMAQ suggests that CMAQ is not fully capturing all of the transport and/or boundary layer processes associated with ground level ozone and the formation, transport and destruction of its precursors. This may help explain some of the ozone underprediction at some surface monitors. Some possible reasons as to why CMAQ is underpredicting ozone in the lower boundary layer aloft are as follows:

- *Errors in mesoscale aerometric data (MM5 wind fields) which could compromise the advection of ozone and precursors by the mean wind components*
- *Errors in CMAQ's handling of turbulent diffusion (gradient transport using eddy diffusivity)*
- *Incomplete chemical reaction schemes (e.g. see Chapter 4 of this Appendix)*
- *Improper accounting of actinic flux aloft where actinic flux is defined as the quantity of light available to molecules at a particular point in the atmosphere and which, on absorption, drives photochemical processes in the atmosphere (See Chapters 4 and 9 of this Appendix for details)*
- *Unaccounted sources of ozone*

The relatively poor performance of CMAQ at rural/suburban sites during pollution events provides information on the shortcomings of CMAQ. Rural/suburban sites are more power-plant dominated than motor vehicle dominated. The underestimation of ozone fluctuations at these sites suggests that variations in power plant emissions are not fully felt at these sites. Consequently, the full benefit of power

plant emission reductions are unlikely to be realized at these sites and also sites downwind from these sites. These biases also lessen our confidence in future ozone predictions from CMAQ (see also Chapter 9).

Future Work

Future efforts to compare CMAQ simulations (or other deterministic photochemical models) to observations will remain vital to better understanding the complex nature of ground-level ozone. Some of the methods that could facilitate future efforts include:

- Simulations with higher vertical resolution might do a better job of capturing the initial dispersion and transport of ozone plumes from power plants. A more dense network of observations would be useful for evaluating the representativeness of individual ozone monitors. Increased, more-coordinated measurements aloft (e.g. hourly ozonesonde launches during high-ozone periods coupled with nighttime aircraft flights) would help explain the diurnal cycle of ozone in rural, urban and elevated surface sites.
- Aircraft observation remains one of the best tools to collect aloft data because of the ability to cover vast distances (both vertically and horizontally) in a relatively short period of time and, perhaps more importantly, the ability to collect ancillary measurements (i.e. VOC's, NO_x, SO₂, CO, PM_{2.5}).
- Ozonesondes provide another measurement tool that over time, if collected regularly, can produce an ozone climatology in the lower boundary layer, stable nocturnal boundary layer, and lower free troposphere.
- A rigorous evaluation of input parameters from the MM5 would also be useful. For example, comparison of MM5 meteorological parameters with balloon soundings, FAA on-board aircraft data, and/or satellite cloud fields.
- How about having an equally rigorous evaluation of MM5 using climatologies of meteorological parameters from balloon soundings and FAA on-board aircraft meteorological parameters?

Increased use of all the tools detailed in this analysis will help produce a better comparative product which will increase confidence in the robustness of future ozone simulations.

References

Draxler, R. A., Stunder, B., Rolph, G. and Taylor, A., 2006. HYSPLIT4 USER's GUIDE Version 4.8 - Last Revision: June 2006. NOAA Air Resources Laboratory, 327 E West Hwy, Silver Spring, 20910.

http://www.arl.noaa.gov/data/web/models/hysplit4/win95/user_guide.pdf

Hains, J.C., Taubman, B.F., Thompson, A.M., Marufu, L.T., Doddridge, B.G. and Stehr, J.W., Dickerson, R.R., 2007. A clustering methodology to identify distinct chemical events in the troposphere using aircraft trace gas and aerosol vertical profiles, manuscript.

Stoeckenius, T. and Kemball-Cook, S., 2005. Determination of representativeness of 2002 ozone season for ozone transport region SIP modeling. *Environ International Corporation*.

Taubman, B.F., Hains, J.C., Thompson, A.M., Marufu, L.T., Doddridge, B.G., Stehr, J.W., Piety, C.A. and Dickerson, R.R., 2006. Aircraft vertical profiles of trace gas and aerosol pollution over the mid-Atlantic United States: Statistics and meteorological cluster analysis, *J. Geo. Res.*, 111, D10S07.

Acronym List

CMAQ- Community Multi-scale Air Quality
RAMMPP- Regional Atmospheric Measurement Modeling and Prediction Program
AGL- Above Ground Level
OTR- Ozone Transport Region
UMD- University of Maryland
EPA- Environmental Protection Agency
NAA- Non Attainment Area
R- Rural
S- Suburban
U- Urban
HYSPLIT- HYbrid Single-Particle Lagrangian Integrated Trajectory
MM5- The PSU/NCAR mesoscale model (5th generation)