Comments to
EPA Clean Air Scientific Advisory Committee, Ozone Review Panel

Health Risk and Exposure Assessment for Ozone (HREA), 2nd External Review Draft
Welfare Risk and Exposure Assessment for Ozone (WREA), 2nd External Review Draft Health Risk
Policy Assessment for Ozone (PA), 2nd External Review Draft

Model-Based Rollback in the Assessment of Risk

Christopher Emery
ENVIRON International Corporation, Novato, CA

My comments focus on EPA's model-based projections of ozone air quality in 2006-2010 to meet current and alternative primary and secondary standards that serve as input to the various health and welfare risk and exposure assessments. I specifically address EPA's photochemical modeling of 2007 conducted with the CMAQ instrumented with the High Order Decoupled Direct Method (HDDM), their development of regression-based models from HDDM results and their application to adjust 2006-2010 measurement data, and EPA's new analyses of modeled background ozone specific to the year 2007.

I have several concerns about the approach and sources of potentially large uncertainty, which I enumerate below. My comments relate to the following EPA charge questions to CASAC: #4 and #5 of the HREA and WREA; #2 of the PA, Chapter 2. My remaining comments provide support for these concerns based on HDDM modeling and assessment of regression-based rollback that ENVIRON has conducted over the past two years.

SUMMARY

1. The REA and PA documents need to emphasize that the methodology to project observation data in 2006-2010, using sensitivity regression models formed from CMAQ/HDDM photochemical modeling of only 8 months of 2007, is equivalent to a ***"screening-level" exercise that is subordinate to year-specific photochemical modeling***. The projection approach does not explicitly account for inter-annual variability in measured ozone, emissions, and meteorology. Furthermore, using a regression approach to project conditions in other years cannot account for inter-annual variability of background ozone. I strongly disagree that five years represent a "small window" (as stated in the HREA) relative to 8 months of modeling. Ideally, CMAQ/HDDM modeling should be run for every year being analyzed in the risk assessment.

2. Recognizing that CMAQ/HDDM modeling of every year is not realizable, ***uncertainty ranges in the regression-based projections (e.g., standard error or similar) should be propagated into the risk models to quantify the range of risk in meeting the current and each alternative standard.*** EPA analyzes the uncertainty of the HDDM technique relative to "brute force" simulation results with reduced emissions, and report that mean errors and related uncertainty are "small" (3-4 ppb). They perform no similar comparison for the

regression-model projections, which is the technique actually applied to 2006-2010 measurement data. Nevertheless, just 5 ppb error can result in large (>10%) differences in emission reductions needed to meet alternative standards because of shallow ozone response slopes. ***Comparisons of regression projections to brute force should be determined and resulting error ranges added to the discussion of uncertainty.***

3. The HREA and associated Chapter 4 appendices state that use of regression models to characterize the central tendency of ozone sensitivity to emission reductions will tend to dampen variability in the ozone response. Our modeling analyses also show this effect – it alters the resulting frequency distributions (relative to HDDM) and reduces the level of control needed to meet alternative standards. ***This needs to be highlighted and emphasized up front as a major uncertainty and disadvantage of the approach.***

4. Overall, CMAQ performance in replicating observed conditions in 2007 is quite good in most cities and seasons, but significant disagreements occur for certain cities and seasons (Cleveland, Dallas, Houston in summer; Sacramento and Los Angeles in winter). This elevates (and compounds) uncertainty in the HDDM sensitivities, the regression models from which they are derived, the projections of ozone, and the quantification of risk. ***This calls into question the reliability of risk estimates in those circumstances.*** Los Angeles is a particularly difficult environment to simulate with models, and ***error metrics of the magnitude shown in Chapter 4 appendices, along with the reported spatial performance issues (Figure 92 in Appendix 4-B) lead to very large uncertainty in projected spatial fields for exposure and "composite monitor" calculations used in BenMAP.***

5. Certain problematic issues associated with the regression-based projections are raised in the HREA for New York and Los Angeles, and the HREA describes ad hoc solutions for these cities to circumvent unrealistic projections. If resulting uncertainty is so much higher in these two cities because of these adjustments, how much (if any) faith can be put into the resulting risk and exposure calculations at those locations? ***Keeping these cities in the analysis undermines the overall results and they should be removed.***

6. EPA mostly applied NOx-only emission reductions to project 2006-2010 measurement data to just meet current and alternative standards. In correctly stating that equivalent NOx and VOC reductions do not optimize the costs or reductions needed to meet the standard, ***it is also quite imperative to also make clear that different NOx and VOC combinations, while able to meet the same standards, will result in a variety of ozone frequency distributions and spatial ozone response patterns, and this adds an important source of uncertainty to the risk and exposure calculations***.

7. In the WREA, ***additional analyses should be presented to calculate projections of W126 fields across the US when meeting any of the alternative primary ozone standards*** (not just the current 75 ppb standard and alternative secondary standards). This would allow insight into whether the primary alternative standards are equivalently or more protective than the secondary alternative standards.

8. The PA describes new modeling analyses of background ozone during the 2007 modeling year. ***It is important to clearly reiterate in the PA (as opposed to the in-depth discussion in the PA appendix) that background ozone is not static as emissions change.*** Background is destroyed by chemical interactions with US anthropogenic emissions (Lefohn et al., 2014) and so the frequency distribution for background ozone will shift to higher concentrations as emissions are reduced. Therefore, "zero-out" brute force modeling (from which US background is determined) represents the upper end of the background estimate, and thus the actual minimum achievable risk, while apportionment-based modeling (which accounts for current influences of US emissions) represents the low end of the background estimate.

9. The PA describes analyses for season mean background ozone. Analyses should be extended to include the high tail (98^{th} percentile or 4^{th} high maximum daily 8-hour) of modeled background from the zero-out modeling, to identify locations/seasons/years where background is estimated to be above current and alternative standards. Figures should be added that show the fraction of anthropogenic contributions when USB is at 98^{th} percentile from source apportionment modeling.

GENERAL COMMENTS AND SUPPORTING INFORMATION

Overall, the techniques and analyses performed to characterize air quality and its temporal-spatial response to emission reductions to meet current and alternative standards are significant technological improvements over the 2008 ozone review and the first draft HREA in the current review. It is clear that the EPA has attempted to address all of the issues raised in previous reviews with respect to non-linear ozone response to emission reductions, and in most cases they include novel and thorough analyses of responses to controls and associated uncertainties. An impressive quantity and quality of work has been accomplished in a very short timeframe, and for this EPA staff should be commended.

Here I present an assessment of model-based rollback techniques by summarizing results from HDDM modeling that ENVIRON has conducted over the past two years (Yarwood et al., 2013; Downey et al., 2014) within the context of the approach described in the HREA and by Simon et al. (2012). These results have been shared periodically with EPA's modeling group since the September 2012 CASAC meeting.

Assessment of the HDDM/Regression Approach

Yarwood et al. (2013) have run the CAMx/HDDM photochemical model for the entire year of 2006 on a 12 km grid system covering North America. HDDM sensitivity output was extracted for all AQS monitoring sites in 23 US cities. In many respects, the general HDDM modeling approach of Yarwood et al. is similar to the approach of Simon et al. in that multiple HDDM simulations were used to quantify ozone sensitivity to US anthropogenic NOx and VOC emissions spanning the full range of emissions from 100% (replication of actual conditions) to 0% (representing US background ozone). We refer the reader to the respective publications for additional information on the respective methodologies.

From the 2007 CMAQ/HDDM results, Simon et al. developed linear regression models that describe the "central tendency" of first and second-order ozone sensitivity to a given ozone input. Separate regressions were developed for each monitoring site, season, and hour of the day. The purpose of the regression models is to extend HDDM sensitivity results in 2007 to hourly monitoring data in years ranging from 2006 to 2010. Simon et al. quantify uncertainty resulting from the regression technique for two example cities (Charlotte and Detroit) in terms of standard error and correlation, both of which quantify the degree of data point scatter around the regression line.

To further evaluate uncertainties in the regression technique, ENVIRON developed and evaluated a sensitivity regression model from our 2006 CAMx/HDDM results. Figure 1 displays first and second order sensitivity regressions for an urban site in Dallas during summer afternoon hours and during winter night hours. Despite some wide scatter and very low correlation values, especially during cold/night conditions, the vast majority of slopes were statistically significant to within 99% confidence limits for any season, time period, or environment (urban or rural). We found that standard errors for slope were typically within a few percent, but could range well over 10% in complex/variable conditions (e.g., night, winter, NOx-heavy environments). Correlation, as given by the coefficient of determination (R^2), indicated a wide amount of scatter but was typically better for rural than urban sites, summer than winter, and day than night. The worst correlation cases were associated with larger levels of fresh NOx emissions available for ozone titration.

Poor correlation is a result of large variations in sensitivity response calculated by HDDM due to numerous complex interactions involving emissions, vertical mixing, transport, etc., and this variability is replaced with a "central tendency" as determined by the linear regression model. An example of this variability-reducing effect is shown in Figure 2, which plots projections of hourly ozone time series at the urban Dallas site over 2006 for zero US anthropogenic NOx and VOC emissions (US background ozone). The results using HDDM sensitivities directly are shown at the bottom and the results using regression model sensitivities are shown at the top. While the HDDM sensitivities retain more realistic hourly variability, the regression sensitivities result in some obvious reduction in variability and "clipping" effects for extended night/cool periods when high NOx/low ozone conditions resulted in zero-slope regressions (differences among the seasonal regressions are also apparent among these "clipping" periods).

Evidence of similar non-physical effects in the EPA's regression-based ozone projections are noted in Figures 61 (Denver) and 63 (Houston) in the REA, Appendix 4-D (replicated below as Figures 3 and 4). Figure 3 appears to be a classic case of "clipping" as a result of regressions; there is an artificial capping at 15 ppb ozone response in the 15-60 ppb range when ozone is adjusted to meet the 65 ppb standard. Figure 4 shows strange maximum positive responses in a very small range of observed MDA8 ozone around 20 ppb. ***Rogue patterns such as these suggest poor regression for certain sites/hours/ seasons.***

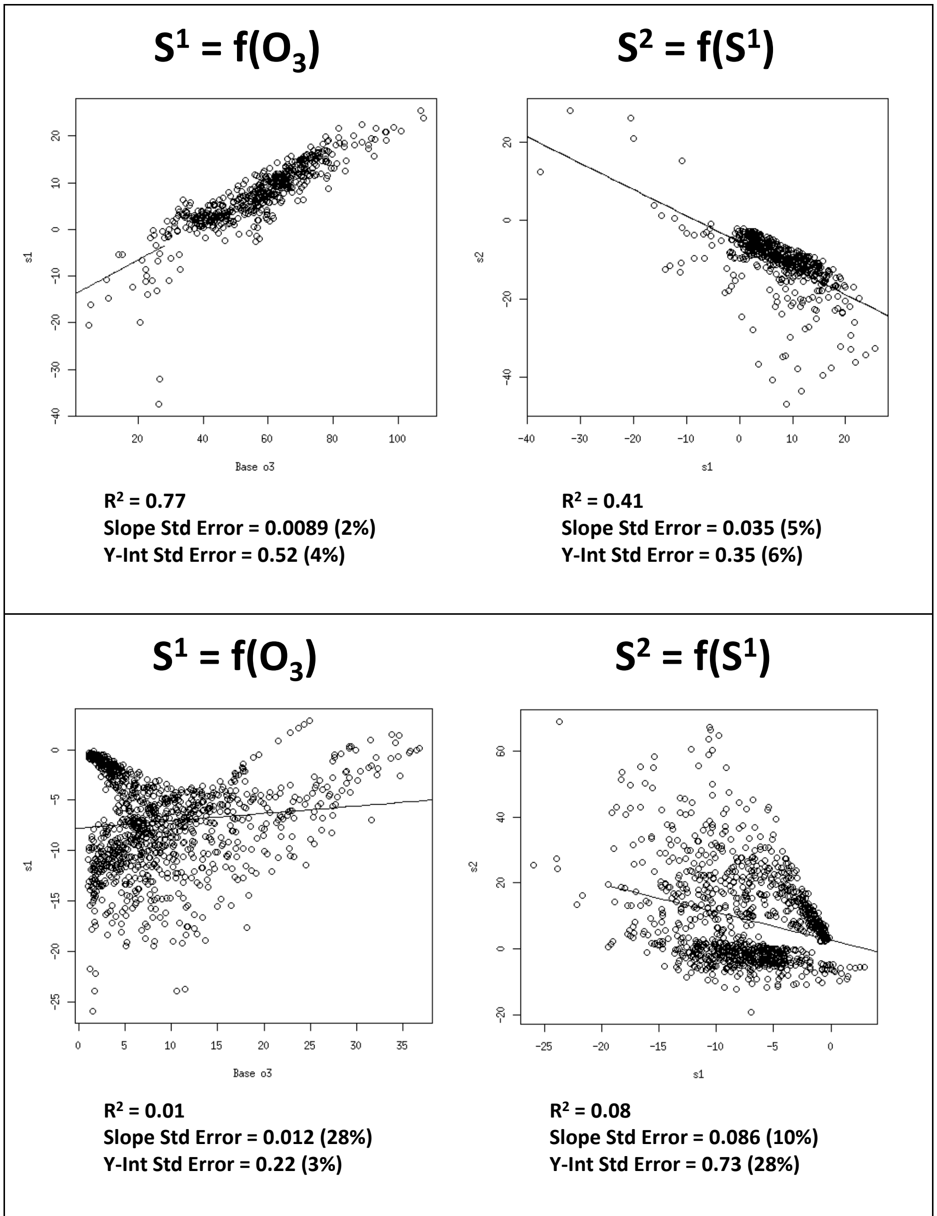
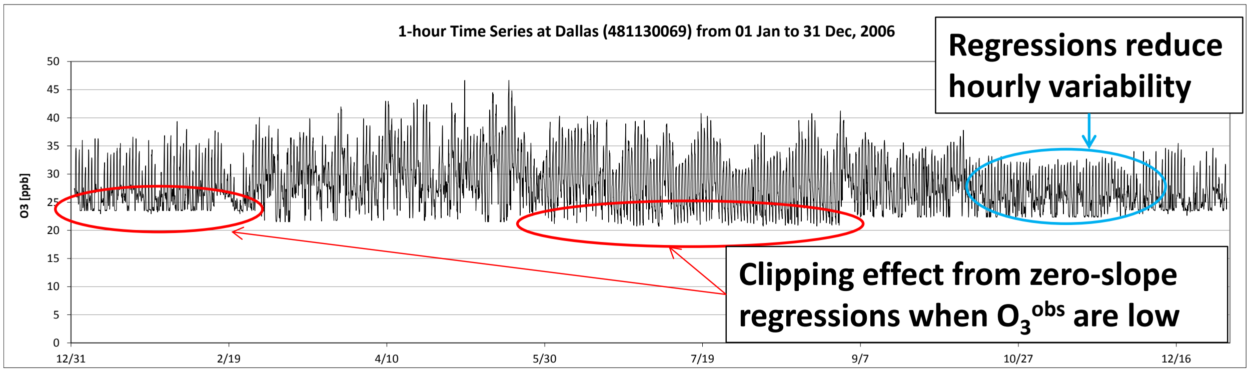


Figure 1. Regressions for first (S^1) and second (S^2) order ozone sensitivity from CAMx HDDM simulations at an urban Dallas monitoring site for summer afternoons (top) and winter nights (bottom). Statistical parameters for each regression model are noted below each figure.



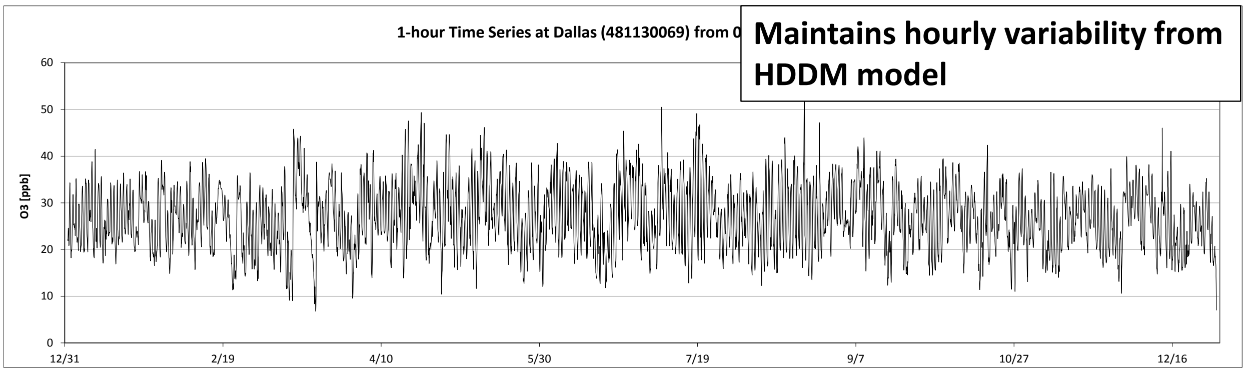


Figure 2. Projections of 1-hour ozone over 2006 at an urban Dallas monitoring site resulting from zero US anthropogenic NOx and VOC emissions (US background ozone). Top: ozone resulting from sensitivity regression model. Bottom: ozone resulting directly from HDDM sensitivities.

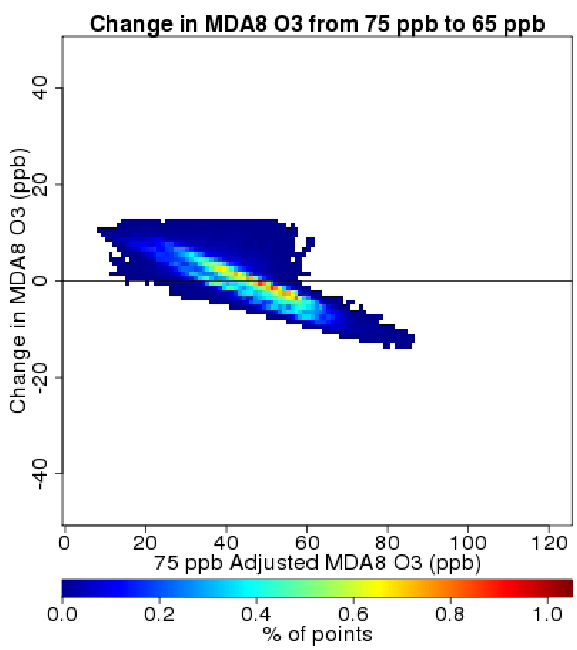


Figure 3 (Figure 61 in REA Appendix 4-D). Change in VNA estimates of the daily maximum 8-hour average (MDA8) O_3 concentrations based on HDDM adjustments in Denver.

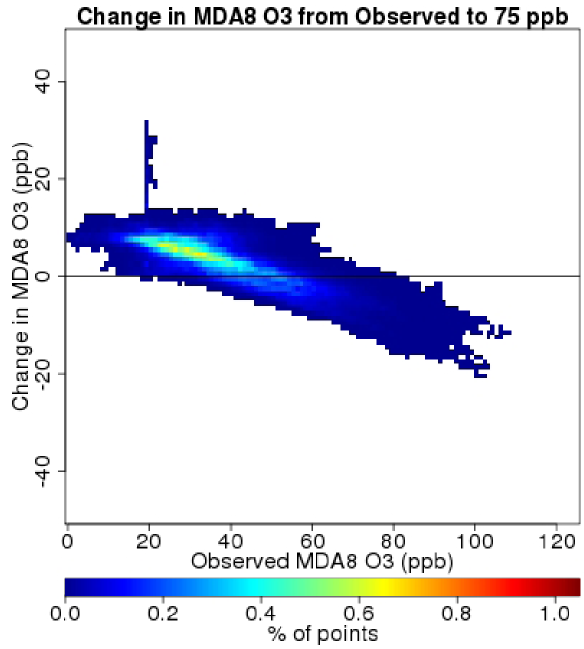


Figure 4 (Figure 63 in REA Appendix 4-D). Change in VNA estimates of the daily maximum 8-hour average (MDA8) O_3 concentrations based on HDDM adjustments in Houston.

We compared ozone projections generated from our HDDM and regression models to CAMx brute force runs in which NOx and VOC emissions were reduced by 75% and by 100% (US background ozone). Figure 5 displays projected ozone during July 2006 at the urban Dallas site from the regression model (right) and the HDDM model (left) compared to the brute force results; both scatter plots and 1-hour frequency distributions are presented. Ozone projections using HDDM sensitivities directly performed better in replicating brute force results than the regression sensitivities, which deteriorated toward the zero emissions case. However, HDDM and regression results were more consistent at rural sites (not shown).

A final analysis was conducted to investigate how the amount of emission reductions needed to meet current and alternative ozone standards change using HDDM sensitivities directly versus regression sensitivities. Figure 6 presents these results for three cities where the CAMx model performed well in replicating the observed 2006 hourly ozone frequency distributions. The regression approach consistently allowed for more emissions (less reduction) to meet proposed standards, except in Los Angeles for the lowest standards. We found that this effect was driven by the fact that the regression approach produces a combination of larger sensitivity and less variability at the upper tail of the frequency distribution. Therefore, peak ozone (i.e., 4^{th} highest MDA8) is more affected by these artificial tendencies than mean ozone.

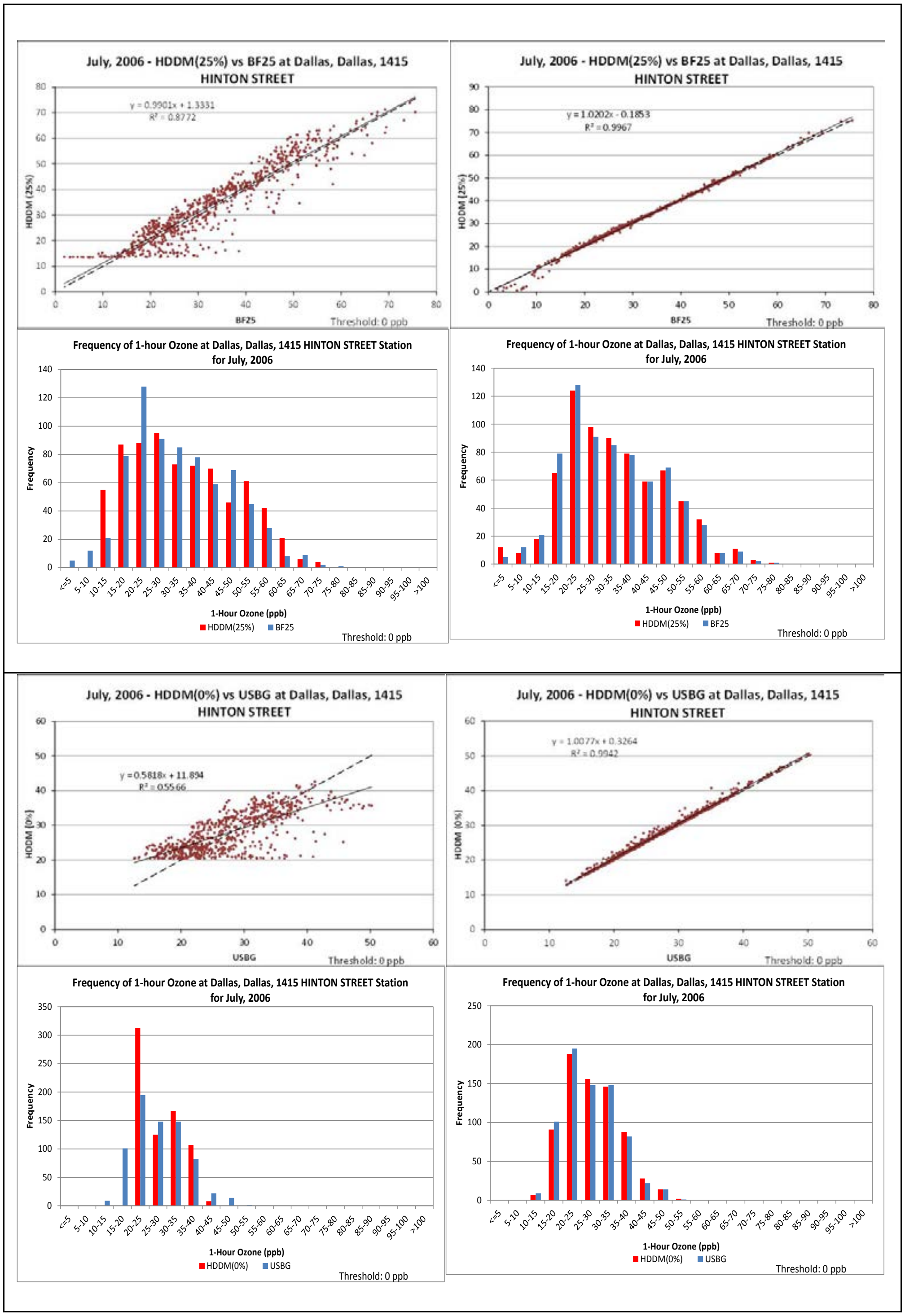


Figure 5. July ozone projections at a Dallas urban monitoring site using sensitivity regressions (left) and HDDM sensitivities directly (right) compared to ozone projections from brute force emission reduction simulations (labeled USBG). Top panel is for 25% NOx and VOC US anthropogenic emissions, bottom panel is for 0% emissions. Frequency distributions show results from the brute force run (blue) and the respective projections using HDDM and regression sensitivities (red).

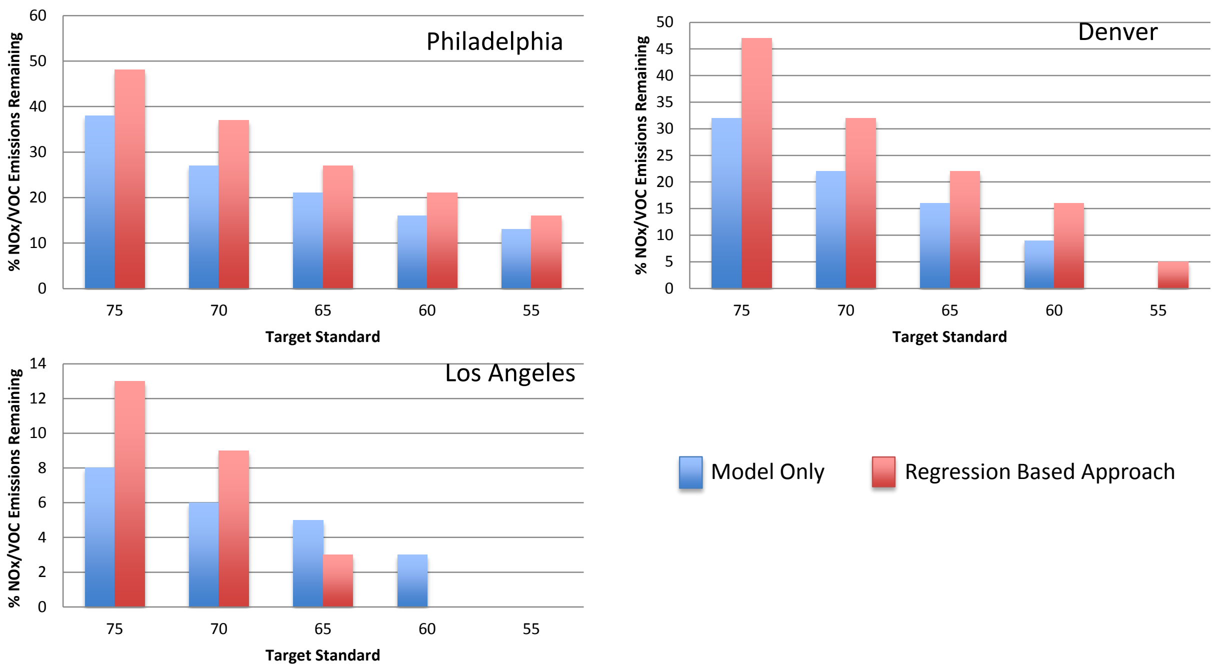


Figure 6. Estimated emissions necessary to meet current and alternative ozone standards in three well-performing cities, expressed as percent of total emissions in 2006. Projection estimates are shown using HDDM sensitivities directly ("Model Only", blue bars), and regression sensitivities (red bars). Figure contributed by Nicole Downey, Earth Systems Sciences, LLC.

Implications

HDDM modeling, especially employing high resolution over the entire US, is very time-consuming and requires significant computer resources. Nevertheless, given the strong non-linear response of ozone to NOx, EPA found that multiple CMAQ/HDDM runs were necessary to adequately address the full range (0-100%) of emission reductions. We arrived at a similar conclusion based on our CAMx/HDDM modeling.

High ozone days at a fixed location can occur for many different reasons, which contribute to variability in ozone sensitivity and thus result in different ozone responses to precursor reductions. As a simplistic thought experiment, consider a monitoring site where many 3 PM hours during a particular season reported the same ozone concentration (say 60 ppb). On some days ozone increased to that value, on others ozone decreased to that value; some days were cloudy with shallow vertical mixing, others were clear with deep vertical mixing; some days were dominated by long-range transport, others were stagnant and/or controlled by various local emission mixtures depending on wind direction; some days were weekdays with heavy traffic, others were weekends or holidays with different traffic patterns; and so on. To the extent that these effects are adequately replicated in photochemical models, HDDM ozone sensitivity will appropriately vary among these hours, ***but the "central tendency" of the regression approach will apply exactly the same sensitivity response for each of these 60 ppb hours***.

Based on this example, it is not clear that EPA's current regression approach can adequately address hour/day-specific responses to emission reductions across a 5-year period involving large variability in meteorology, emissions, and economic activity. Our sensitivity regression model leads to unrealistic responses to emission reductions, which are worse for deep cuts and complex conditions – when HDDM sensitivity variability is particularly relevant and important. Furthermore, our regression model has difficulty replicating brute force simulations in the same season in which they are derived. ***It is therefore a logical stretch to assume they can be extended to project ozone in other years when environmental and emission conditions are very different***.

CAMx/HDDM Estimates of Emission Reductions Needed to Meet Alternative Standards

ENVIRON and Earth System Sciences have recently developed a paper (Downey et al., 2014; in review) presenting the results of CAMx/HDDM simulations for 2006, following the technique described by Yarwood et al. (2013). Our analysis includes estimates of the response of annual 4th high (H4) maximum daily 8-hour (MDA8) ozone in four cities to reductions in 2006 US anthropogenic NOx and VOC emissions from zero to 100%. The four cities (Los Angeles, Sacramento, St. Louis and Philadelphia) were selected because they met strict model performance criteria; no ad hoc techniques (e.g., scaling or regression relationships) were applied to adjust for model error.

Figure 7 shows the response of city-wide peak H4 MDA8 ozone over the full 2006 emissions range. A key attribute in this figure is the shallow response of peak ozone to large reductions in emissions, where ozone response is weak down to 25% of 2006 emissions, and then steeper below 25%. Not only are deep cuts of more than 80% necessary to meet alternative standards down to 60 ppb in all four cities, but model error in projected H4 MDA8 of just 5 ppb would alter emission reductions by 10% or more because of the shallow slopes. ***Uncertainty introduced by HDDM relative to "brute force" emission reduction scenarios should compound with uncertainty introduced by employing "central tendency" linear regression sensitivity models to project observational data in 2006-2010***.

Effect of Emissions on Background Ozone

"Background" ozone is not static with respect to local anthropogenic emissions, rather it evolves as emissions change. A simplistic schematic example of this effect is shown in the Figure 8. At current emissions, total ozone is high along with the anthropogenic contribution, while background ozone is at its lowest contribution because it is chemically scavenged by local emissions. This is the background simulated by source apportionment methods as EPA has presented and which is described by Lefohn et al. (2014), who have labeled this component as "Emissions Influenced Background" (EIB) ozone. As anthropogenic emissions are reduced to zero, total ozone decreases with the anthropogenic component, but the background increases toward a maximum at background due to the diminishing influence of chemical scavenging. This is the background simulated by "zero out" simulations. As a result, the amount of ozone reduced by anthropogenic emission reductions is non-linearly replaced by increases in background, leading to a shallower ozone slope than would occur in the absence of any background contribution. This effect is much larger for cities than for rural areas, that is,

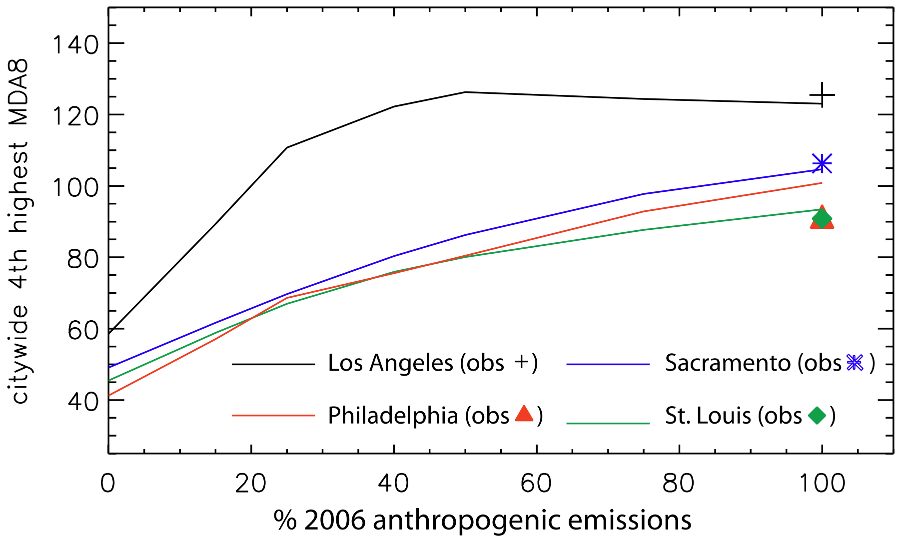


Figure 7. Response of peak H4 MDA8 ozone in four cities over the range of US anthropogenic NOx and VOC emissions (100% is 2006 baseline emissions, 0% is US background). The symbols at 100% emissions indicate the 2006 observed peak H4 MDA8 ozone for reference. From Downey et al. (2014).

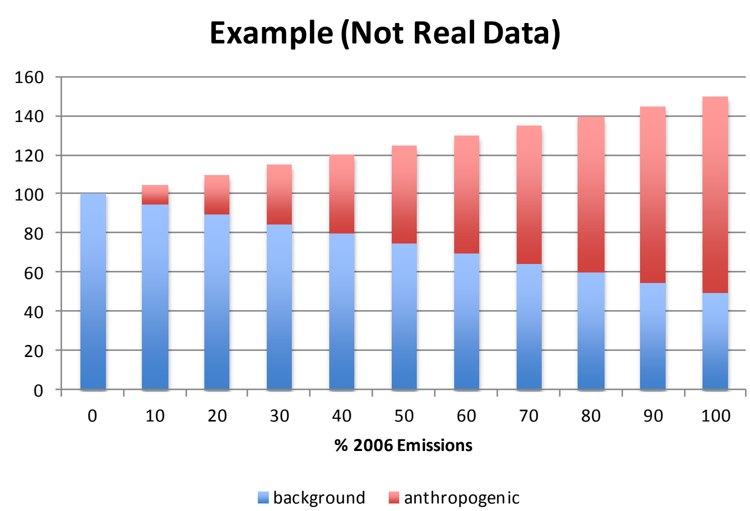


Figure 8. Example schematic of the relative contributions to total ozone from background (blue) and local anthropogenic (red) sources. The vertical axis represents percent of unaffected background (e.g., NAB or USB), the horizontal axis represents percent of current (2006) US anthropogenic emissions. Figure contributed by Nicole Downey, Earth Systems Sciences, LLC.

background will increase most rapidly in areas where ozone suppression is lifted as NOx emissions decrease. Therefore, it is important to understand that zero-out modeling represents the upper end of the background estimate, while apportionment-based background represents the low end (or current conditions).

As a final note, it would not be appropriate to estimate the risk associated with background ozone by simply scaling ozone data going into the risk models by the seasonal mean fractional values comprising background, as EPA suggests in the PA. A better approach would be to explicitly use hourly, site-specific background estimates from source apportionment runs for a given scenario (say, an alternative standard) in the risk models, and then subtract the resulting risk estimates from the risk derived using full ozone estimates for the same scenario. The ultimate minimum achievable risk would be represented by the upper limit of background ozone resulting from a US zero-out scenario.

REFERENCES

Downey, N., Emery, C., Jung, J., Sakulyanontvittaya, T., Hebert, L., Blewitt, D., Yarwood, G.: Emission reductions and urban ozone responses under more stringent US standards, submitted to Environmental Science & Technology, 2014.

Lefohn, A.S., Emery, C., Shadwick, D., Wernli, H., Jung, J., Oltmans, S.J.: Estimates of Background Surface Ozone Concentrations in the United States Based on Model-Derived Source Apportionment, Atmospheric Environment, doi:10.1016/j.atmosenv.2013.11.033, 2014.

Simon, H., Baker, K. R., Akhtar, F., Napelenok, S. L., Possiel, N., Wells, B., and Timin, B.: A Direct sensitivity approach to predict hourly ozone resulting from compliance with the National Ambient Air Quality Standard, Environmental Science & Technology, doi:10.1021/es30674e, 2012.

Yarwood G., C. Emery, J. Jung, U. Nopmongcol, and T. Sakulyanontvittaya: A method to represent ozone response to large changes in precursor emissions using high-order sensitivity analysis in photochemical models, Geosci. Model Dev., 6, 1601–1608, doi:10.5194/gmd-6-1601-2013, 2013.