Dynamic Evaluation of Regional Air Quality Model's Response to Emission Reductions in the Presence of Uncertain Emission Inventories

Sergey L. Napelenok¹*, Kristen M. Foley¹, Daiwen Kang², Rohit Mathur¹, Thomas Pierce¹, S. Trivikrama Rao¹

¹Atmospheric Modeling and Analysis Division, National Exposure Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, North Carolina
²Computer Sciences Corporation, Research Triangle Park, North Carolina

Abstract

A method is presented and applied for evaluating an air quality model's changes in pollutant concentrations stemming from changes in emissions while explicitly accounting for the uncertainties in the base emission inventory. Specifically, the Community Multiscale Air Quality (CMAQ) model is evaluated for its ability to simulate the change in ozone (O₃) levels in response to significant reductions in nitric oxide (NOₓ = NO + NO₂) emissions from the NOₓ State Implementation Plan (SIP) Call and vehicle fleet turnover between the years of 2002 and 2005. The dynamic model evaluation (i.e. the evaluation of a model's ability to predict changes in pollutant levels given changes in emissions) differs from previous approaches by explicitly accounting for known uncertainties in the NOₓ emissions inventories. Uncertainty in three sectors of NOₓ emissions is considered – area sources, mobile sources, and point sources – and is propagated using sensitivity coefficients calculated by the decoupled direct method in three dimensions (DDM-3D). The change in O₃ levels between 2002 and 2005 is estimated based on
differences in the empirical distributions of the modeled and observed data during the two years. Results indicate that the CMAQ model is able to reproduce the observed change in daily maximum 8-hr average $O_3$ levels at more than two-thirds of Air Quality System (AQS) monitoring locations when a relatively moderate amount of uncertainty (50%) is assumed in area and mobile emissions of $NO_x$ together with a low amount of uncertainty (3%) in the utility sector (elevated point sources) emissions. The impact of other sources of uncertainty in the model is also briefly explored.

Key words: dynamic model evaluation; CMAQ; direct decoupled method; DDM; air quality modeling; ozone; uncertainty; sensitivity

* Corresponding author. Email: napelenok.sergey@epa.gov; Phone: +1-919-541-1135; Fax: +1-919-541-1379; Address: EPA (MD-E243-04), RTP, NC 27711.

1. Introduction

Regional air quality models (RAQMs) are an integral part of air quality management. Their applications include developing national strategies for improving air quality (Moran, 2005), assessing the regional-scale transport of pollutants (Rao et al., 2008), and informing international treaties on the hemispheric transport of air pollutants (Holloway et al., 2003). When applied for these tasks, RAQMs typically employ “current” and “future” emission scenarios, with the commonly examined metric of interest being the change in the concentration or deposition of air pollutants attributable to the change in emissions.
The integrity of an RAQM is ascertained through comprehensive model evaluation by comparing model outputs with observations (Dennis et al., 2010). In the context of air quality management, where the model's response to varying inputs is of concern, it is essential to evaluate the model's ability to accurately predict changes in pollutant concentrations (i.e., model response) due to changes in emission inputs or meteorology. Evaluating the model's response is referred to here as dynamic model evaluation. In contrast to the operational evaluation, in which model outputs are usually paired in space and time with available observations, dynamic evaluation is more difficult to perform, because a fully controlled dynamic air quality scenario is nearly impossible to observe and quantify in nature. Rather, the model's ability to correctly respond to different input conditions can be evaluated by examining modeling periods separated by substantial and quantifiable differences in the input variables of interest, namely, emissions and/or meteorological conditions.

One frequently evaluated scenario satisfying the above prerequisites is the period between 2002 and 2005 in the eastern United States when substantial reductions of nitrogen oxide (NO\textsubscript{x} = NO + NO\textsubscript{2}) emissions were achieved through the U.S. Environmental Protection Agency's (US-EPA) NO\textsubscript{x} State Implementation Plan (SIP) Call (US-EPA, 2005). Additionally, reductions in mobile emissions due to vehicle fleet turnover and introduction of lower emitting automobiles also occurred during this period. Several previous dynamic evaluation studies that focused on this period (Gilliland et al., 2008; Godowitch et al., 2010; Pierce et al., 2010) found a smaller change in model predictions of peak surface ozone (O\textsubscript{3}) concentrations compared to the observed change. These studies reported that the model predicts around 40-80% of the observed change in the daily maximum 8-hr average O\textsubscript{3}, and hypothesized that a lower magnitude response in the model predictions may be attributable to errors in NO\textsubscript{x} emission inputs used in
the simulations. However, potential uncertainties in the NO\textsubscript{x} emission inventories were not explicitly accounted for in these dynamic evaluation studies.

The objective of the work presented here is to perform a dynamic evaluation of the Community Multiscale Air Quality (CMAQ) modeling system (Byun and Schere, 2006) covering the NO\textsubscript{x} SIP Call period while accounting for uncertainties in NO\textsubscript{x} emission inputs. Specifically, the change in the modeled daily maximum 8-hr average O\textsubscript{3} concentrations between the summers of 2002 and 2005 in the eastern United States is evaluated against the change seen in the observations. The proposed approach is used to investigate the extent to which the previously noted model's insufficient response to changes in NO\textsubscript{x} emissions is attributable to uncertainties in these model inputs.

Various methods have been developed to propagate uncertainty in input parameters through RAQMs. These typically include some variation of the "brute force" Monte Carlo simulations, in which the model is rerun many times with varying inputs (Hanna et al., 2001; Boynard et al., 2011), as well as studies where a reduced form of the model is developed based on calculated sensitivity coefficients (Digar and Cohan, 2010; Tian et al., 2010), which are then coupled with randomly sampled distributions of uncertainties. Each method culminates in an ensemble of model predictions from which inferences on the impact of input uncertainty can be ascertained. To dynamically evaluate CMAQ in light of uncertainties in NO\textsubscript{x} emissions inputs, the sensitivity-based approach to propagate uncertainty is used here as it is significantly more computationally efficient (Pinder et al., 2009), because the random sampling calculations are much less computationally demanding than running a regional photochemical simulation model many times. Thus, the sensitivity-based approach allows for the development of large member ensembles for two summer periods with a lower computational burden. Furthermore, the method
presented here extends previous work by combining the sensitivity-based uncertainty analysis with a dynamic evaluation.

2. Method

2.1. Modeling System and Observations

Model simulations were conducted with a 12 km horizontal grid covering the eastern United States nested within a 36 km grid over the entire continental United States, including parts of Canada and Mexico. The modeled period spanned the period from 1 June to 31 August in both 2002 and 2005. CMAQ version 4.7.1, with decoupled direct method in three dimensions (DDM-3D) (Napelenok et al., 2008), was used to calculate \( O_3 \) concentrations and sensitivities of \( O_3 \) to three emission sectors (area sources, mobile sources, and point sources) of \( NO_x \).

Atmospheric chemistry was simulated with the latest available version of the carbon bond mechanism (CB05) (Sarwar et al., 2008). Meteorological inputs were supplied by MM5 version 3.6.3 (Grell et al., 1994) configured with the standard physics options (see Godowitch et al., 2010). Boundary conditions for the larger domain were specified based on outputs of the global model GEOS-Chem (Bey et al., 2001). Emissions were developed using the Sparse Matrix Operator Kernel Emissions (SMOKE) processor (http://www.smoke-model.org) version 2.4 based on temporally and spatially resolved wildfire, electricity generating units, and mobile sources. Domain-wide \( NO_x \) emissions reflected reductions from the SIP Call as well as reductions in the mobile sector. On average, point source emissions of \( NO_x \) were reduced by 22% and mobile source emissions were reduced by 18% between the summer of 2002 and the
summer 2005 in the modeling domain; area sources, which included biogenic NO emissions,
remained relatively unchanged over these two years (0.2% reduction).

The observed O₃ concentrations for this study were obtained from the EPA’s Air Quality
System (AQS, http://www.epa.gov/air/data/aqsdb.html). Data from over 700 monitoring stations
were available within the 12 km modeling domain during the two modeling periods. AQS O₃
data was processed to calculate daily maximum 8-hour average mixing ratios according to
standard regulatory procedures (Office of the Federal Register, 1997). Monitoring sites with less
than 80 complete days of observations out of a total possible of 92 were excluded from the
analysis.

2.2 Uncertainty Propagation

Calculated DDM-3D sensitivity coefficients were used to estimate the O₃ response to
perturbations in the uncertain inputs of the three NOₓ emissions categories (area, mobile, point)
through Taylor series expansion (Morgan and Henrion, 1990; Hakami et al., 2003). Generally,
pollutant concentration as a function of any one perturbation can be reconstructed using the
following:

\[
C_j(x,t) = C_0(x,t) + \Delta \epsilon_j S_j^{(1)}(x,t) + \frac{1}{2} \Delta \epsilon_j^2 S_j^{(2)}(x,t) + \text{h.o.t.},
\]

where \(C_j(x,t)\) is the concentration due to a specific perturbation \(j\) at time \(t\) and location \(x\);
\(C_0(x,t)\) is base, unperturbed concentration; \(\Delta \epsilon_j\) is the fractional perturbation of the parameter \(j\);
\(S_j^{(1)}(x,t)\) and \(S_j^{(2)}(x,t)\) are the first and second order sensitivity coefficients, and \(h.o.t.\) are higher
order terms with little impact on the approximation.
To account for uncertainty in several parameters at once – area, mobile and point emissions – and dropping the higher order terms, the Taylor Series was expanded as follows:

\[
C_{a,m,p}(\bar{x},t) \approx C_0(\bar{x},t) \\
+ \Delta \varepsilon_a S_a^{(1)}(\bar{x},t) + \frac{1}{2} \Delta \varepsilon_a^2 S_a^{(2)}(\bar{x},t) \\
+ \Delta \varepsilon_m S_m^{(1)}(\bar{x},t) + \frac{1}{2} \Delta \varepsilon_m^2 S_m^{(2)}(\bar{x},t) \\
+ \Delta \varepsilon_p S_p^{(1)}(\bar{x},t) + \frac{1}{2} \Delta \varepsilon_p^2 S_p^{(2)}(\bar{x},t) \\
+ \Delta \varepsilon_a \Delta \varepsilon_m S_{a,m}^{(1)}(\bar{x},t) \\
+ \Delta \varepsilon_m \Delta \varepsilon_p S_{m,p}^{(1)}(\bar{x},t) \\
+ \Delta \varepsilon_p \Delta \varepsilon_a S_{p,a}^{(1)}(\bar{x},t)
\]

where the subscripts \(a, m,\) and \(p,\) represent area, mobile, and point NO\(_x\) emissions respectively.

The formulation in Equation 2 allows for efficient recalculation of O\(_3\) predictions based on perturbations in the three uncertain input parameters. An ensemble of model predictions was developed for the two modeling periods (summers of 2002 and 2005) by randomly sampling (with replacement) from a continous uniform distribution for each perturbation parameter, \(\Delta \varepsilon_a,\)

\(\Delta \varepsilon_m,\) and \(\Delta \varepsilon_p.\) Uniform distribution was chosen to not prescribe the shape of the distribution and specify only the lower and upper bounds (Cullen and Frey, 1999). One hundred such samples were found to be sufficient; higher sample sizes did not have an impact on the interpretation of results.

The uncertainty ranges used to sample \(\Delta \varepsilon_j\) were: \(\pm 3\%\) in point sources; \(\pm 50\%\) in mobile sources; and \(\pm 50\%\) in area sources. The low uncertainty in the emissions from point sources is due to the prevalence of the continuous emission monitoring systems (CEMS) on the majority of
these sources. Uncertainty in the other two sectors was based on the previous efforts that have attempted to quantify emissions uncertainty. For example, Hanna et al. (2001) suggested uncertainty ranges of approximately 20-50% for major point sources and 35-100% for other categories. Comparisons of different methods of calculating mobile emissions have shown a variability of approximately 35-47% (Parrish, 2006; Dallmann and Harley, 2010). NOx inventories constructed from inverse modeling studies based on satellite and ground-based observations coupled with regional chemical transport models also show a wide range of variability when compared to a priori estimates (Konovalov et al., 2006; Dequillaume et al., 2007; Napelenok et al., 2008). The Taylor Series approach (Equation 2) allowed for flexibility in selecting uncertainty ranges due to the trivial computational costs associated with the sampling. The sensitivity of the final results to the choice of these uncertainty ranges is discussed in more detail below.

2.3 Dynamic evaluation metrics

The model output data generated in this experiment consisted of two 100-member ensembles of daily maximum 8-hr average ozone concentrations: one for the 2002 summer season and one for the 2005 summer season, each with 100 Monte Carlo samples of the uncertainty ranges for the three NOx emission sectors (Figure 1). As noted before, a sample size of 100 was found to be sufficient in the ensemble analysis. These ensembles are compared to base model simulations and observational values. Observational data consisted of June through August time series data for each monitoring site and each year. All data were first processed to calculate the daily maximum 8-hr average O3 values. Since the two modeling seasons were separated by differing meteorology and mismatched temporal emission patterns (in addition to
different emission magnitudes), the dynamic signal should not be discerned by pairing model and
observational data in time. Instead, the data were paired in space only (at each observational
site.)

Figure 1. Base model, observations, and model ensemble empirical cumulative distributions of daily
maximum 8-hr average ozone concentrations for 2002 and 2005 at two AQS sites: a) Terre Haute, IN
(AQS#181670024); and b) Detroit, MI (AQS#261630019). All ensembles were constructed based on ±50%
uncertainty in emissions of area and mobile NOx, and ±3% uncertainty in emissions of point NOx. The wide
spread of the ensemble at the Terre Haute site indicates greater sensitivity to NOx emissions in comparison to
the site in Detroit.

The ability of the CMAQ model to correctly predict the change in daily maximum 8-hr
average O3 values was evaluated by estimating the difference between the empirical distributions
of ozone model predictions in the two years. The difference was quantified using two measures:
the root mean square deviation (RMSD) and mean deviation (MD) calculations. The RMSD was

175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190

9
\[ \text{RMSD}(2002,2005) = \sqrt{\frac{\sum_{p=0}^{n_s} (\hat{X}^{2002}(p) - \hat{X}^{2005}(p))^2}{N}}, \]  

(3)  

where \( N \) is the cardinality of \( \{p_l, \ldots, p_u\} \) and \( \hat{X}^{\text{new}}(p) \) is the \( p^{th} \) sample quantile of the data set. For example, \( \hat{X}^{2002}(0) \) equals the minimum value of the 2002 ozone predictions, \( \hat{X}^{2002}(50) \) equals the median, \( \hat{X}^{2002}(100) \) equals the maximum value, etc.

The RMSD measures the average distance between two distributions, but it does not indicate whether the difference tends to be positive or negative. Since the model simulation needs to capture the magnitude and direction of a change in ozone values due to changes in emissions, the mean deviation between the two distributions was also calculated:

\[ \text{MD}(2002,2005) = \frac{\sum_{p=0}^{n_s} (\hat{X}^{2002}(p) - \hat{X}^{2005}(p))}{N} \]  

(4)  

RMSD and MD were calculated for two cases: one where \( p_l = 0\% \) and \( p_u = 100\% \) to quantify the difference in the full range of \( \text{O}_3 \) concentrations; and one where \( p_l = 95\% \) and \( p_u = 100\% \) to quantify the difference in the more policy-relevant higher ozone values.

These metrics were calculated for the difference between the observational data sets in 2002 and 2005, and for the difference between the base modeled data sets in the same years.

Additionally, an ensemble of differences between modeled \( \text{O}_3 \) values for 2002 and 2005 was calculated by randomly sampling model ensemble members from each simulation year and then calculating the RMSD and MD for each simulation pair. The 95% confidence intervals based on the ensemble distributions for these metrics were then used to establish criteria for acceptable model performance. The model is said to have successfully captured the observed change in \( \text{O}_3 \) at any one site if the observed RMSD and MD metrics both fell within these estimated
confidence intervals. For example, at the AQS site near Terre Haute, IN the observed change
was estimated by RMSE=10.0 ppb and MD = 9.5 ppb (using $p_l=0\%$, $p_u=100\%$) (Figure 2a, b).
The 95% confidence interval based on the ensemble of model results was (3.34, 13.5) ppb for the
RMSE and (-0.6, 12.6) ppb for the MD. Thus, the model was able to capture the observed
change in the ozone distribution across these years, considering ±50% uncertainty in the inputs
of mobile and area sources, and ±3% uncertainty in point sources of NOx emissions. Use of both
the RMSE and the MD provides a stringent test for the model since each metric highlights a
different attribute of the difference between the two empirical distributions. This is illustrated by
an AQS site outside of Detroit, MI (Figure 2c, d). At this site, the observed RMSE of 6.6 ppb
was within the model range of (4.5, 6.8) ppb, but the observed MD of -3.3 ppb was well outside
the model range of (-0.7, 3.4) ppb. The behavior in MD at Detroit was caused by the fact that at
the lower part of the distributions of both modeled and observed values (10%-60% range), the O3
values were higher in 2005, and at the higher end of the distribution (60-100% range), the O3
values were higher in 2002 (Figure 1b). At the same time, the model underestimated the
difference at the low end of the distribution and overestimated the difference at the high end,
leading to a range of MDs much greater than the MD of observations (Figure 2d).
Figure 2. Root mean square deviation (RMSD) and mean deviation (MD) at two AQS sites: a,b) Terre Haute, IN (AQS#181670024); and c,d) Detroit, MI (AQS#261630019) for modeled and observed values indicated by solid and dotted vertical lines respectively, as well as the distribution of the two metrics for the ensemble of model simulations. Both the observation-based RMSD and MD fall within the 95% confidence interval of the ensemble distribution at Terre Haute (indicated by green shading), but the observation-based MD for Detroit is much lower than the model ensemble distribution. The model ensemble was developed based on ±50% uncertainty in area and mobile emissions of NOx, and ±3% uncertainty in point emissions of NOx.

3. Results
3.1 Base model performance

Before examining the model’s response to changes in emissions between 2002 and 2005, model performance for the base case simulations for the two years was assessed (Table 1). In 2002, normalized mean error (NME) was 16.6%, and normalized mean bias (NMB) was 0.8%. The results for 2005 showed NME of 17.6% and NMB of 2.6%. These results are similar to the model performance metrics reported in other studies for this domain (Eder and Yu, 2006; Appel et al., 2007).

<table>
<thead>
<tr>
<th>N</th>
<th>Mean_{obs}</th>
<th>Mean_{model}</th>
<th>RMSE (ppb)</th>
<th>NME (%)</th>
<th>MB (ppb)</th>
<th>NMB (%)</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002 All</td>
<td>61379</td>
<td>54.6</td>
<td>55.1</td>
<td>11.9</td>
<td>16.6</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Green sites</td>
<td>39634</td>
<td>54.8</td>
<td>55.2</td>
<td>11.2</td>
<td>15.8</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Red sites</td>
<td>21745</td>
<td>54.4</td>
<td>54.9</td>
<td>13.1</td>
<td>18.3</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>2005 All</td>
<td>61126</td>
<td>49.9</td>
<td>52.5</td>
<td>11.4</td>
<td>17.6</td>
<td>2.6</td>
<td>5.2</td>
</tr>
<tr>
<td>Green sites</td>
<td>39456</td>
<td>49.9</td>
<td>52.4</td>
<td>10.9</td>
<td>16.8</td>
<td>2.5</td>
<td>4.9</td>
</tr>
<tr>
<td>Red sites</td>
<td>21670</td>
<td>50.0</td>
<td>52.9</td>
<td>12.4</td>
<td>19.0</td>
<td>2.9</td>
<td>5.8</td>
</tr>
</tbody>
</table>

3.2 Spatial Distribution of the Dynamic Signal in the Presence of NOx Emissions Uncertainty

Under a moderate amount of emission uncertainty (±50% in area and mobile sources of NOx and ±3% in point sources of NOx), the model ensemble was able to capture the observed change in the ozone distribution at a majority of AQS sites (Figure 3). The results were consistent when the RMSD and MD were calculated across all percentiles, p = 0 – 100%, and when using only the upper end of the distribution, p = 95 - 100%. This similarity suggests that the model performs equally well in predicting both high ozone values and overall O3 when
accounting for uncertainties in NO\textsubscript{x} emissions inputs. No discernable pattern was found in the
spatial distribution of sites for which the model ensemble did not encompass the observed
change in the O\textsubscript{3} distribution. Therefore, it is likely that several factors contributed to the poor
performance at these sites that are not related to errors in the NO\textsubscript{x} emissions. For example, the
model has been shown to have errors associated with the transport of O\textsubscript{3} and its precursors in the
area of the I-95 corridor in the Northeast (Godowitch et al., 2010). Some of the sites where
model performance was poor also included various urban areas where O\textsubscript{3} is frequently less
sensitive to NO\textsubscript{x} emissions. In fact, the base model performance was relatively poor at the sites
where the observed change was not captured by the model ensemble (Table 1).

The spread of the ensemble of predicted RSMD and MD values was evaluated using the
Talagrand diagram (Hamill, 2001). This evaluation shows that the observed RMSD and MD
values tend to fall outside of the range of model predicted values (Figure 4.) The Talagrand
diagrams are characterized by a U-shape, indicating a statistically overconfident model (i.e. the
range of model values is too narrow). In addition, the observed values are often larger than the
largest ensemble member shown by the far right bin of the histogram. This supports the earlier
findings that the model response to the changes in NO\textsubscript{x} emissions is too low compared to
observations (Gilliland et al., 2008). This also provides evidence that the uncertainty ranges for
the emission inputs of NO\textsubscript{x} were too narrow and/or that other sources of model uncertainty (e.g.
boundary conditions, VOC emissions inventory, chemistry, and meteorology), as well as those
associated with capturing the inherent variability in point measurements with grid-average model
predictions, are still present. These sources of uncertainties have been identified in the past and
work has been progressing to address them (McKeen et al., 1991; Hanna et al., 2005; Vautard et
al., 2006).
Figure 3. Geographical extent of the ability of CMAQ to capture the observed change in daily maximum 8-hr average ozone between the summers of 2002 and 2005. Green points represent AQS sites that fell within the 95% confidence interval of the model ensemble distributions of RMSD and MD based on ±50% uncertainty in area and mobile source emissions of NOx and ±3% point source emissions of NOx. Red points are AQS sites where the observed change was not captured by the model ensemble.
Figure 4. Talagrand diagrams for a) root mean square deviations (RMSD) and b) mean deviations (MD) of the change in daily maximum 8-hr average ozone concentrations between 2002 and 2005. Plots show the percentile of the observed metric with respect to the ensemble of modeled values based on ±50% uncertainty in area and mobile source emissions of NOx and ±3% point source emissions of NOx. Each bar represents 2.5% and the green bars show the 95% confidence interval.

3.3 Impact of Uncertainty Assumptions

The presented methodology for evaluating a regional air quality model was designed to allow for flexibility in the assumptions made about the range of uncertainty in the model input parameters. The results presented above were based on the assumption of relatively moderate amount of uncertainty in emissions of area and mobile source NOx (±50%) and a low amount of uncertainty in point source emissions (±3%). As discussed previously, higher levels of emissions uncertainty have been suggested in the literature. Therefore, the dynamic evaluation presented in section 3.2 was repeated using a range of different uncertainty estimates (Figure 5). As expected, accounting for uncertainty in area and mobile sources of emissions of NOx,
improved the model’s ability to capture the observed change in daily maximum 8-hr average O$_3$.

At 100% uncertainty in area and mobile NO$_x$ emissions, the observed change in the ozone
distribution was captured at more than 90% of the sites.

Figure 5. Ability of the model ensemble to capture the observed change in daily maximum 8-hr average O$_3$
values between the summers of 2002 and 2005 shown as a percentage of all sites where the observations fall
within the 95$^{th}$ percentile of the model ensemble (y-axis) as a function of the uncertainty in area and mobile
sources of emissions of NO$_x$ (x-axis). The uncertainty in area and mobile sources are assumed equal for
simplicity of presentation. In each case, the uncertainty in point emissions of NO$_x$ was assumed to be ±3%.
The shaded square indicates the ±50% uncertainty case.

3.3 Sensitivity to VOC Emissions and Boundary Conditions

To investigate possible causes for the narrow range of model predicted values, indicated
by the Talagrand diagram (Figure 4), two additional sources of uncertainty that may have
influenced the model’s ability to accurately predict a change in ozone during this time period
were considered. DDM-3D sensitivity of O$_3$ to emissions of Volatile Organic Compounds
(VOC) and to boundary conditions of all modeled species was calculated for a sample two week
period in July 2005 (Figure 6). The modeling domain showed several large areas where ozone sensitivity to VOC emissions was greater than the sensitivity to NOx emissions (Figure 6d) during this period. Higher VOC sensitivity would suggest a VOC-limited ozone formation regime where perturbations in NOx emissions would have little impact.

Many of the sites with observed RMSD and MD metrics that fell outside the estimated 95% confidence intervals exhibited higher relative sensitivity to these two additional parameters than to emissions of NOx. For example, at the poor performing site in Detroit, MI, sensitivity to VOC emissions accounted for a higher contribution to ozone formation than the sensitivity to NOx emissions (Figure 7) on most days. At the site at Terre Haute, IN, for which the model was able to capture the observed change in O3 well, the sensitivity to VOC emissions was lower with one notable exception on July 14th. This day was characterized by low predicted O3 concentrations. At both of these sites, as well as all others in the domain, model results were also highly sensitive to boundary conditions. These results help explain the narrower range of the ensemble distributions seen at the Detroit site (Figure 1b), assuming these findings can be extended to the full summer episodes of the two years. Furthermore, since NOx emissions were greatly reduced in 2005 from their 2002 levels, it is likely that VOC sensitivity became relatively more important in 2005, leading to possible O3 formation regime changes and transitioning some NOx-limited regions to become VOC-limited (e.g. the northeastern urban corridor along the interstate 95).
Figure 6. Average sensitivities of daily maximum 8-hr average ozone to a) emissions of NOx, b) emissions of VOCs, c) boundary conditions, and d) the ratio of VOC to NOx sensitivity between 3 – 15, July 2005. Regions where VOC sensitivity is higher approximate a VOC-limited ozone formation regime.
4. Summary

An analysis of the change in O$_3$ concentrations due to large reductions in NO$_x$ emissions in the Eastern U.S. between the summers of 2002 and 2005 was performed while accounting for uncertainty in these emissions. The full distribution of observed and modeled O$_3$ concentrations was analyzed, as well as the more policy-relevant higher end of the distributions. The difference between the two study years was estimated by comparing modeled and observed O$_3$ distributions at AQS sites without pairing the data in time. This approach was used to account for the fact that weekly emissions patterns and meteorological drivers are different between the two years. A model performance criterion was developed based on the model’s ability to capture the observed
O3 change (estimated by the RMSD and MD metrics), while accounting for uncertainties in NOx emission inputs. Many standard statistical tests exist for testing for differences between two empirical distributions (e.g. the nonparametric Kolmogorov-Smirnov test). In contrast, the interest in this application is in testing for the difference of differences between two pairs of distributions. Specifically, it is important to test not only whether or not the air quality model can simulate that there is some change in the ozone distribution as a response to changes in emissions, but whether the model can capture the magnitude and direction of this change. This complication motivates the novelty of the proposed test metrics.

Assuming ±50% uncertainty in NOx emissions from area and mobile sources, and ±3% uncertainty in point sources, the ensemble of model predictions was able to capture the observed change in O3 levels at about two-thirds of AQS sites based on the 95% confidence interval chosen here. Several factors can explain the poor ensemble performance. The main contributor was likely the fact that errors in the dynamic evaluation cannot be attributed to NOx emissions uncertainty alone. Evidence was found for significant presence of VOC-limited regions in the domain, and VOC emissions are likely to have the same, if not higher, level of uncertainty. Furthermore, boundary conditions were found to have a substantial impact on O3 formation, and, although untested in this experiment, uncertainty in meteorological inputs was also likely present. Since uncertainty in the emission inventory is difficult to quantify explicitly, it was shown that ensemble performance is sensitive to the selection of the uncertainty levels in emissions of mobile and area sector NOx.

In summary, the methodology presented in this study illustrates the impact of propagating assumed levels of uncertainty in one set (emissions of NOx) of many uncertain model input variables. It shows that accounting for even one uncertain parameter influences the
interpretation of model's responsiveness to changes in emissions as well as base case model
performance evaluation. Ideally, a more comprehensive analysis of uncertainties in all model
input variables would be extremely useful, and our work is progressing towards this purpose.

Disclaimer and Acknowledgements

Although this manuscript has been reviewed and approved for publication, it does not
necessarily reflect the policy or views of the US Environmental Protection Agency. The authors
would like to thank Jenise Swall, Brian Eder, Rob Pinder, as well the anonymous reviewers for
their constructive comments on the manuscript.
References


Moran, M. D. (2005) "Current and proposed emission control programs: How will acid deposition be affected?" Environment Canada. 99-162.


