Evaluation of EPA’s Modeled Attainment Test Software (MATS) for Modeled Future Design Value ($DV_F$) Calculation in HGB

Jim Smith
Southeast Texas Photochemical Modeling Technical Committee
February 18, 2009
CAMx Ozone Modeling in SIP Development

The Big Picture

- **Base Case**: Day-specific meteorology and emissions; replicate what actually happened
- **Baseline Case**: Day-specific meteorology and typical emissions; used in RRF to predict future design values
- **Future Base Case**: Apply future growth + on-the-books controls to estimate future ozone
- **Control Strategy Testing**: Determine control strategies that will effectively reduce ozone
- **SIP**: Document modeling procedures
Background

• EPA developed Model Attainment Test Software (MATS) for calculation of future design values for attainment demonstrations.
• Prior to the release of the software, TCEQ modelers developed in-house procedures to implement the EPA guidance for performing this calculation.
• Some stakeholders and modeling researchers contend TCEQ should be using MATS because
  – EPA has devoted considerable resources in providing the software for states to use
  – Other states are using MATS
  – There appear to be some minor inconsistencies between MATS’ and TCEQ’s calculations.
Background (Continued)

• The question - Is TCEQ justified in continuing to use its own PERL script after evaluation of MATS?

• To address this question, we used MATS to process output from the 1c baseline and the CS02 2018 future case model runs, and compared MATS’ $D_{VF}$ values with our own.
**DV_F and RRF Primer**

- **Regulatory Design Value (DV_R):**
  - Average of three consecutive years’ fourth highest 8-hour ozone concentrations measured at an individual monitor
  - Highest DV_R in an area determines attainment status, classification

- **Baseline Design Value (DV_B):**
  - Average of three years’ DV_R values, as per EPA Guidance
  - Basis for modeled attainment test

- **Relative Response Factor (RRF):**
  - A ratio estimating the model’s response at a monitoring site
  - Based on modeled baseline and future ozone concentrations in a neighborhood near the monitor

- **Future Design Value (DV_F):**
  - The product of the RRF and the DV_B: \( DV_F = RRF \times DV_B \)
  - Used to demonstrate attainment of the ozone NAAQS
Baseline Design Value Example

- To calculate the $DV_B$ for 2005 at Deer Park, we need the three $DV_R$ values which include 2005:
  - The 2005 $DV_R$ is the average 4th highest ozone concentration for 2003, 2004, and 2005. For Deer Park,
    The 2005 $DV_R$ is $(113 + 97 + 92) / 3 = 100$ (truncated)
    The 2006 $DV_R$ is $(97 + 92 + 101) / 3 = 96$ (truncated)
    The 2007 $DV_R$ is $(92 + 101 + 86) / 3 = 93$

- The $DV_B$ for 2005 is then the average of the three $DV_R$ values: $(100 + 96 + 93) / 3 = 96.3$
2005 Base Year

Average of 2005 DV, 2006 DV, and 2007 DV - weights the 2005 4th high 8-hour ozone value as most influential
How RRF is calculated at a monitor

- Select a suitable area surrounding the monitor, usually 3X3, 5X5, or 7X7 grid cells, depending on grid cell size.
- For each day modeled, find the maximum modeled baseline 8-hour ozone concentration in the selected area.
- Select days to use in the RRF calculation. EPA recommended method is:
  1. Select days with max modeled baseline 8-hour ozone concentration ≥ a threshold value $T_1$ (default 85 ppb)
  2. If < 10 days selected in Step 1, then reduce threshold progressively until either:
     a) Ten days are selected, or
     b) a lower threshold $T_2$ (default 70 ppb) is reached.
  c) If $T_2$ is reached before 10 days are selected for a monitor, Guidance recommends states discuss with regional office.
How RRF is calculated at a monitor (cont.)

- After days are selected, the baseline and future case modeled ozone concentrations* are averaged across days for each monitor.

\[
\text{Average modeled future case concentration} = \frac{1}{n} \sum_{i=1}^{n} \text{Concentration}_i
\]

- RRF = \[
\frac{\text{Average modeled future case concentration}}{\text{Average modeled baseline concentration}}
\]

* maximum baseline and future case modeled concentrations within nearby grid cells
Modeled Attainment Test Software (MATS)

• MATS is provided by EPA to help states use model output in their attainment demonstrations
  – Performs RRF and DV$_F$ calculations
  – Performs an “unmonitored area analysis” (see Dave Westenbarger’s SIM presentation from March 20, 2008).

• MATS is a Windows-based interactive program which supports a limited number of choices in performing RRF/DV$_F$ calculations:
  – Lets you choose the size of the area around the monitor from which to pick 8-hour ozone maximum concentration
  – Lets you try different thresholds for selecting days (T$_1$ and T$_2$).
MATS

Choose Desired Output

Scenario Name: MATS Test

Point Estimates
Forecast
- Temporarily adjust ozone levels at monitors

Spatial Field
Baseline
- Interpolate monitor data to spatial field
- Interpolate gradient-adjusted monitor data to spatial field
POSTOCCUS
- Interpolate monitor data to spatial field; Temporarily adjust ozone levels
- Interpolate gradient-adjusted monitor data to spatial field; Temporarily adjust

Calculate DVF Values
Monitor locations (lat/long) and \( D_{V_R} \) values for at least three years

Modeled ozone concentrations by grid cell, with lat/long coordinates

Use maximum concentration in 7X7 grid cell neighborhood
Use 2005-2007 DVₚ values to calculate DVₜ
MATS

Thresholds T1 and T2 for selecting days to use in RRF calculation (set to 80 & 0 for test)
<table>
<thead>
<tr>
<th>Monitor*</th>
<th>MATS 2005 DVb</th>
<th>RRf</th>
<th>DVf</th>
<th>TCEQ 2005 DVb</th>
<th>RRf</th>
<th>DVf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danciger</td>
<td>81.5</td>
<td>0.8504</td>
<td>69.3</td>
<td>81.5</td>
<td>0.85</td>
<td>69.305</td>
</tr>
<tr>
<td>Houston Aldine</td>
<td>88</td>
<td>0.8835</td>
<td>77.7</td>
<td>88</td>
<td>0.884</td>
<td>77.751</td>
</tr>
<tr>
<td>Channelview</td>
<td>85.7</td>
<td>0.9391</td>
<td>80.4</td>
<td>85.67</td>
<td>0.939</td>
<td>80.45</td>
</tr>
<tr>
<td>Northwest Harris Co.</td>
<td>91.7</td>
<td>0.844</td>
<td>77.3</td>
<td>91.67</td>
<td>0.841</td>
<td>77.105</td>
</tr>
<tr>
<td>Houston Bayland Park</td>
<td>100.7</td>
<td>0.8619</td>
<td>86.7</td>
<td>100.67</td>
<td>0.862</td>
<td>86.768</td>
</tr>
<tr>
<td>Houston Monroe</td>
<td>95.3</td>
<td>0.8944</td>
<td>85.2</td>
<td>95.33</td>
<td>0.894</td>
<td>85.261</td>
</tr>
<tr>
<td>Wallisville Road</td>
<td>94</td>
<td>0.9254</td>
<td>86.9</td>
<td>94</td>
<td>0.925</td>
<td>86.983</td>
</tr>
<tr>
<td>HRM-3</td>
<td>88</td>
<td>0.9332</td>
<td>82.1</td>
<td>88</td>
<td>0.933</td>
<td>82.124</td>
</tr>
<tr>
<td>Lynchburg Ferry</td>
<td>89</td>
<td>0.9482</td>
<td>84.3</td>
<td>89</td>
<td>0.947</td>
<td>84.252</td>
</tr>
<tr>
<td>Houston East</td>
<td>82.7</td>
<td>0.9223</td>
<td>76.2</td>
<td>82.67</td>
<td>0.922</td>
<td>76.217</td>
</tr>
<tr>
<td>Clinton</td>
<td>86.3</td>
<td>0.9252</td>
<td>79.8</td>
<td>86.33</td>
<td>0.925</td>
<td>79.869</td>
</tr>
<tr>
<td>Hou.DeerPrk</td>
<td>96.3</td>
<td>0.921</td>
<td>88.6</td>
<td>96.33</td>
<td>0.921</td>
<td>88.716</td>
</tr>
<tr>
<td>Conroe Relocated</td>
<td>85</td>
<td>0.8554</td>
<td>72.7</td>
<td>85</td>
<td>0.855</td>
<td>72.706</td>
</tr>
</tbody>
</table>

* Selected HGB Monitors
Why the Difference?

• While MATS and PERL script results are very close, the difference of .003 in the RRF at NW Harris County is too big to attribute solely to round-off error.
• Further investigation showed MATS used 22 days in its RRF calculation, while PERL script used 21 (21 is correct). MATS actually showed other monitors in the wrong grid cells.
## Calculation Differences

<table>
<thead>
<tr>
<th>Monitor</th>
<th>MATS</th>
<th>TCEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2005 DVb</td>
<td>RRf</td>
</tr>
<tr>
<td>Danciger</td>
<td>81.5</td>
<td>0.8504</td>
</tr>
<tr>
<td>Houston Aldine</td>
<td>88</td>
<td>0.8835</td>
</tr>
<tr>
<td>Channelview</td>
<td>85.7</td>
<td>0.9391</td>
</tr>
<tr>
<td>Northwest Harris Co.</td>
<td>91.7</td>
<td>0.844</td>
</tr>
<tr>
<td>Houston Bayland Park</td>
<td>100.7</td>
<td>0.8619</td>
</tr>
<tr>
<td>Houston Monroe</td>
<td>95.3</td>
<td>0.8944</td>
</tr>
<tr>
<td>Wallisville Road</td>
<td>94</td>
<td>0.9254</td>
</tr>
<tr>
<td>HRM-3</td>
<td>88</td>
<td>0.9332</td>
</tr>
<tr>
<td>Lynchburg Ferry</td>
<td>89</td>
<td>0.9482</td>
</tr>
<tr>
<td>Houston East</td>
<td>82.7</td>
<td>0.9223</td>
</tr>
<tr>
<td>Clinton</td>
<td>86.3</td>
<td>0.9252</td>
</tr>
<tr>
<td>Hou.DeerPrk</td>
<td>96.3</td>
<td>0.921</td>
</tr>
<tr>
<td>Conroe Relocated</td>
<td>85</td>
<td>0.8554</td>
</tr>
</tbody>
</table>
Analysis and Possible Explanation

• Since MATS is a “Black Box”, it’s difficult to know for sure why the RRF/DVF calculations differ subtly from TCEQ’s.
• Best guess is that discrepancy is due to using different map projections:
  – We conduct our modeling on a Lambert Conformal Grid, but
  – MATS requires input in Latitude/Longitude.
• Since the map projections are tilted a few degrees with respect to each other, it seems likely that some monitors might “move” when converting from LCP to Lat/Long, causing different sets of model output to be used in MATS’ calculations.
**MATS Advantages**

- MATS is easy to use and runs quickly.
- It is EPA’s preferred tool for performing the unmonitored area analysis.
- MATS threshold values can be easily manipulated to test various combinations of values.
MATS Disadvantages

- MATS is Windows-based and cannot easily be incorporated into our LINUX-based runstream; it has to be run separately on a Windows-based PC.
- MATS requires input in latitude and longitude which means extra processing of model output.
- MATS is not set up to handle baseline modeling for multiple years (2005 and 2006 in our case). It had to be “tricked” to work for us.
- MATS is a “black box” – it’s impossible to know what it’s really doing without some serious detective work.
- MATS gives close (but not exact) approximations of the true RRF and DV_F values.