Discussion of EPA Attainment Modeling Guidance for the 8-Hour Ozone NAAQS for Houston Area modeling

Erik Snyder

EPA Region 6

May 25, 2006

@ UT- Pickle Center Commons

Purpose of Meeting

- Meeting is being held at request of local HGB stakeholders.
- HGB is unique in many ways compared to rest of 8-hr nonattainment areas. Specific questions arise that require case-by-case analysis and input on appropriate modeling guidance.

GENERAL CAVEAT

- This meeting is being held by EPA to answer general questions on EPA's modeling guidance in light of HGB specific questions raised by stakeholders.
- Stakeholder groups are conducting modeling of additional episodes/sensitivities. This meeting is not meant to judge the quality or potential applicability of stakeholder modeling.

GENERAL CAVEAT (continued)

- TCEQ has the authority and responsibility for development of the 8-Hr attainment demonstration SIP (including modeling).
- How any stakeholder modeling is utilized in development of the 8-Hr SIP for HGB will be up to TCEQ.

Status of 8-Hr Ozone Modeling Guidance

- "Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hour Ozone NAAQS"
  - Original draft released in 1999
  - Released "Draft Final" version on February 17th, 2005
    - Comments received in March 2005
  - Final version released November 5, 2005
- Available on EPA's website at:
  - http://www.epa.gov/scram001/guidance_sip.htm

Final Ozone Modeling Guidance-Key Updates

- Where is the attainment test applied?
- Screening test/unmonitored areas
- Calculation of current/baseline design values
- Language concerning transport
- Which future year to model?
- Weight of evidence/supplemental analyses
- Minimum RRF threshold
- Minimum number of days to model

Overview

- Some of this presentation will highlight specific areas of EPA guidance that indicates what analyses should be covered by SIP level modeling that TCEQ is conducting.
- Where possible, it is helpful in weighing and evaluating stakeholder modeling to have the same level of detailed analyses.
- Other parts of this presentation will highlight EPA guidance and Region 6 thoughts on specific questions that have arisen.

Episode Selection and Conceptual Model

Episode selection

- EPA guidance has 4 primary factors to consider and 6 secondary factors to consider in selecting episodes.
- Understanding how episodes were selected and fit the conceptual model help weigh modeling outputs.

Conceptual model and Episode Selection

- Episodes selected should be analyzed and documented how they fit with the conceptual model for the area.
- Episode selection should target selections to cover primary meteorological regimes and then cover other regimes as possible.

Episode Selection Criteria (Primary)

- Choose a mix of episodes which represents a variety of meteorological conditions which frequently correspond with observed 8-hour daily maxima exceeding 84 ppb
- Choose episodes having days with monitored 8-hour daily maxima close to observed average 4th high daily maximum ozone concentrations.

Episode Selection Criteria (Primary) (Cont.)

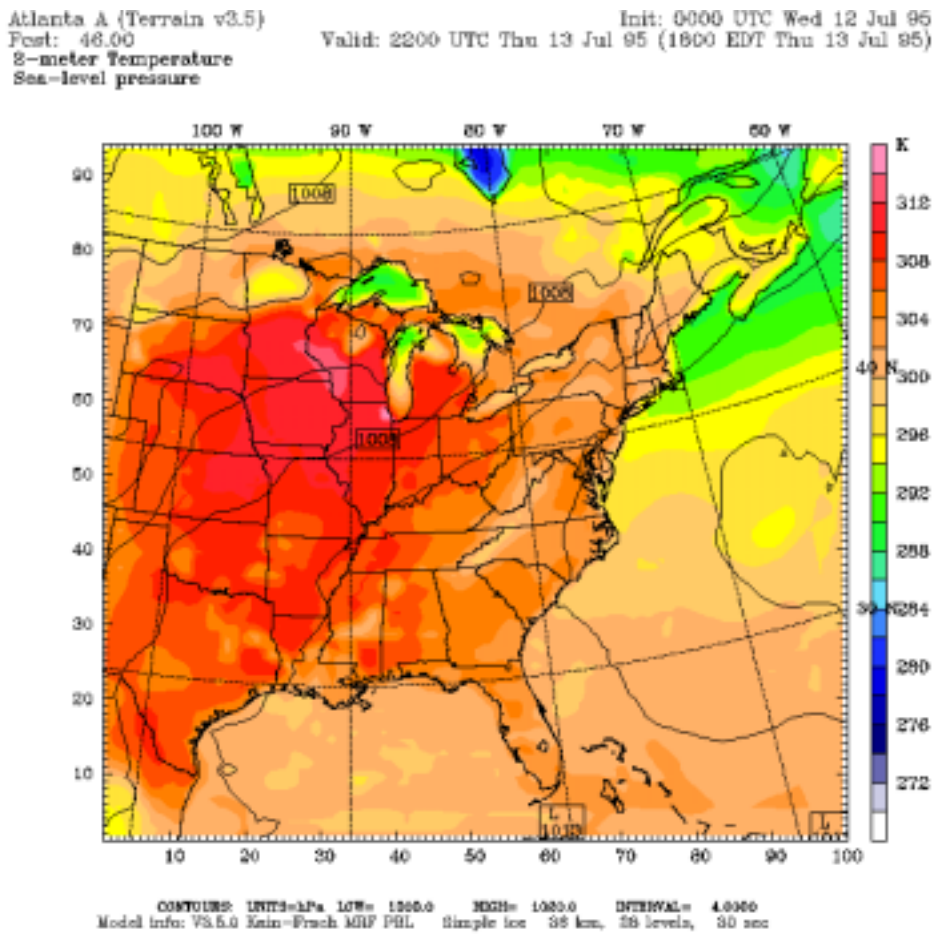
- Choose days with intensive data bases.
- Choose a sufficient number of days to enable the monitored attainment test to be based on multiple days at each monitoring site violating the NAAQS.

Episode Selection Criteria (Secondary)

- Choose episodes which have already been modeled.
- Choose episodes which are drawn from the period upon which the baseline design value is based.
- Choose episodes having observed concentrations “close to” the NAAQS on as many days and at as many sites as possible.
- It is desirable to include weekend days among those chosen, especially if concentrations greater than 84 ppb are observed on weekends.

Met Model Performance

Meteorological Modeling: Model Performance Evaluation



- Importance of meteorological modeling
- Evaluating meteorological modeling input data
  - Operational evaluation
  - Phenomenological evaluation
  - Blending the meteorological evaluation results w/ the AQM results

Importance of Meteorological Modeling

- Generally less energy goes into meteorological modeling, relative to emissions & the model itself
  - Three-legged stool analogy (Met, Emissions, Photochemical)

- Meteorological effects are increasingly important in Ozone/PM/RH applications

Evaluation of Meteorological Modeling

- Guidance rewritten to incorporate many of the latest findings re: evaluation of meteorological modeling. Influenced by several documents:
  - **Tesche et al (2002): "Operational evaluation of the MM5 meteorological model over the continental United States: Protocol for annual and episodic evaluation."**
  - **Emery and Tai (2001): "Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes."**
  - **Numerous MM5 evaluations: (e.g., Olerud et al., 2000; Doty et al., 2001; Johnson, 2003; Baker, 2004; Gilliam, 2004).**
- AQMs used in attainment demonstrations have consistently been subjected to a rigorous performance assessment, but in many cases the meteorological inputs to these models are accepted as is, even though this component of the modeling arguably contains a higher quantity of potential errors that could affect the results of the analysis (Tesche, 2002).

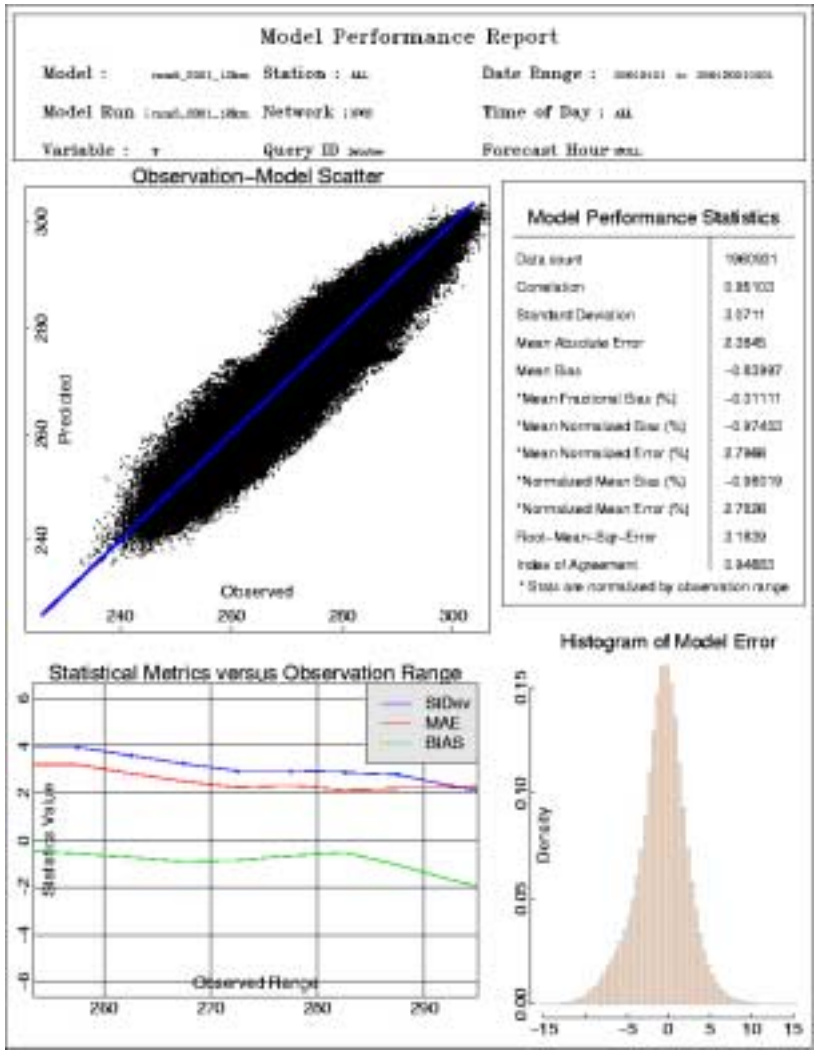
Evaluation of Meteorological Modeling

- Evaluation goal is to move away from “as is” acceptance of meteorological inputs and move toward an understanding of how the bias & error of the meteorological data affect the resultant AQM simulations.
- Two specific evaluation objectives:
  - **1) Determine if the meteorological model output fields represent a reasonable approximation of the actual meteorology that occurred during the modeling period. (Operational evaluation)**
  - **2) Identify and quantify the existing biases and errors in the meteorological predictions in order to allow for an downstream assessment of how the air quality modeling results are affected by issues associated with the meteorological data.**

    (Phenomenological evaluation)

Operational Evaluation:

- **Statistical Measures:**
  - comparisons of the means,
  - mean bias, mean normalized bias,
  - mean absolute error, mean absolute normalized error, RMSE (s & u), and
  - index of agreement.
- **Met Parameters:**
  - temperature
  - water vapor mixing ratios / humidity
  - wind speed / wind direction
  - clouds / radiation
  - Precipitation
  - PBL (max, time series evolution)



Operational Evaluation

- For large modeling exercises, it is recommended that the operational evaluation be broken down into individual segments to allow for a more comprehensive assessment of the meteorological strengths and weaknesses.
  - Geographic subregions.
  - Individual episodes.
  - Aloft vs. surface.
  - Diurnal cycle.
  - As a function of synoptic regimes.

Operational Evaluation

What about the use of performance benchmarks?

- Can help with the initial objective of assessing general confidence in the meteorological model data, however …
- EPA has concerns about potentially misleading comparisons of model performance findings across different analyses w/ differing model configurations and FDDA strengths, therefore …
- while we plan to compare the statistical evaluation outputs back against a set of performance benchmarks, we do not recommend the results of these comparisons be used in a pass/fail mode.

Phenomenological Evaluation

- Components of the evaluation should be tied to the conceptual model of the AQ issue being considered:
  - For example, if a sea breeze is an important regulator of where and when ozone is formed, then the phenomenological evaluation should assess the ability of the meteorological model to reproduce this feature in space/time.
- Other possible elements of a phenomenological evaluation include:
  - Comparison of model trajectories vs. “actual” trajectories (transport).
  - Existence of a low-level jet in model / ambient.
  - Ability of the model to capture frontal passages, airmass residence time.
- Because of the event-oriented nature of the phenomenological evaluation, one will need to summarize model performance in terms of a different suite of statistical metrics: e.g., probability of detection and false alarm rate.

Met Analysis Wrap-up

- Meteorological performance analysis (Operational and Phenomenological) are important components to understanding confidence level of ozone modeling results. i.e. Are you getting the right answer for the right reasons?

Blending the Meteorological Evaluation Results w/ the AQM results

- Armed with the results of the operational and phenomenological evaluation results, one should have an idea of the strengths and weakness of the model.

- At this stage, a decision should be made as to whether to proceed with the meteorological inputs, or whether it is necessary to rerun the met model.
  - Generally, manipulation of existing MM5/RAMS/WRF datasets is not recommended, due to interdependencies of data

- The evaluation results can be used to guide AQ modelers as to those regions/periods/regimes where model performance is optimized, yielding more certain AQM projections.

Ozone Model Performance

Ozone Model Performance

- Model Performance Evaluation (MPE) is the process of comparing model results to measurements to ensure that the model successfully replicates the fundamental processes leading to observed pollutant concentrations.

- The purpose of MPE is to develop confidence that the model will respond to changes in its inputs the same way the airshed would respond.

1-Hr or 8-Hr metrics??

- Eight-hour averaging smoothes both measured & modeled concentrations, and can hide features that are critical to understanding the processes that cause high pollutant concentrations.
- TX EAC areas, ENVIRON, and Region 6 looked at many 8-Hr metrics and in comparison to 1-Hr metrics they were smoother and didn't give as much of an understanding to bias/performance issues. Result: Region 6 gave guidance to use mostly 1-hr ozone metrics in conjunction with the new 8-hr metrics that did not overlap.

Components of Ozone MPE

- One-Hour Ozone Statistics (MNB, MNGE, Peak Prediction Bias & Error, NMB, NMGE, Fractional Bias & Error, RMSE, correlation coefficients, etc.)
- One-Hour Time Series
  - Ozone and Precursors
  - Local Area and Regional ("Background") Monitors
- One-Hour Ozone Daily Peak Isopleth Plots
- One-Hour Scatter Plots
  - Ozone and Precursors
  - By day
  - By Monitor
  - All Data Pairs
  - Q-Q plots
- One-Hour Ozone Videos

Components of Ozone MPE (cont.)

- Eight-Hour Statistics (see 1-Hr metrics list and guidance for complete list)
  - Mean Normalized Bias
  - Mean Normalized Gross Error
  - Correlation Coefficient
  - Calculated for daily peak O3 at monitors
- Eight-Hour Ozone Daily Peak Isopleth Plots
- Eight-Hour Ozone Scatter Plots
  - By Day
  - By Monitor
  - All data pairs
  - Q-Q plots
- If the statistical metric or scatter plot is done for 1-Hr data, then 8-Hr analysis of that metric/scatter plot may not be necessary.

Other MPE considerations

- Diagnostic tests
  - Testing of EI sensitivities 25, 50%, and even higher increases/decreases in different EI components (mobile, biogenic, point source, etc.) In HGB testing of imputed and non imputed HRVOC EI is a good test.
    - Alternate meteorology (e.g. vertical mixing)
    - Alternate Chemical solvers (SAPRC, CB-IV)
    - Alternate Models (CMAQ, CAMx)
    - Alternate boundary conditions
- Diagnostic tests are important for identifying if the modeling is responding as expected to EI changes.

MPE – Refined/Other analysis

- Observational Methods (VOC vs. NOX limitation):
  - Indicator ratios (e.g. VOC/NOX, O3/NOY, O3/HNO3).
  - Smog Production (SP) algorithm.
  - Weekday/weekend analysis.
- Retrospective Analyses:
  - Compare model predictions with observed changes in concentrations. May be complicated by differing meteorologies.
- Probing Tools:
  - Process Analysis.
  - Direct-Decoupled Method.
  - APCA, OSAT.
- Use of non Reference method monitors? Significant Auto GC data is available as well for more recent episodes. Both of these groups of data will help understand MPE.

MPE Summary

- Eight-hour guidance avoids specific performance criteria.
- Requires modelers to assess overall package, rather than relying on overly-simplistic pass-fail tests.
- Key questions:
  - Does the model behave in accordance with the area's conceptual model?
  - Overall, does the model produce unreasonably high or low ozone concentrations?
  - Are there unexplained spatial or temporal patterns in the model output that differ from measurements?

MPE Summary (Cont.)

- Eight-hour guidance introduces the concept of evaluating how well the model responds to changes in emissions, compared with how the airshed responds. This is a very important component of getting the right answer for the right reasons in the development of a control strategy.

Additional MPE considerations

- Aircraft data, Auto GC data, other intensive data sets
- Intermediate reaction products (e.g. HCHO)
- Conceptual models
  - Generic
  - Episode-day specific
- Grid cell arrays to be used, values to be used from arrays. … To be discussed a little later.

AD issue – Base Year/Base EI

Base Year EI– EPA guidance

- Project to a common year (2002) for the Current Year after basecase modeling performance is determined to be acceptable. (EPA Final Guidance pages 24-25)
- For the 2000 episode EPA R6 has not indicated that TCEQ will need to project to a 2002, as that would decouple the intensive field study data from the EI used in the attainment test and would only be a 2-yr change.

Base Year EI– EPA R6 guidance

- For other episodes (non field study) it would be appropriate to project EIs to 2002 or maybe 2000.
- Currently either a 2002 or 2003 EI would be the most recent that would contain the appropriate 5-yr period for the attainment procedures.
- Selection of projected year may inadvertently drop some ozone monitors. Selection of projected year may also yield a below average period for the base DV, that may be decoupled from meteorology that drives the DV over time (i.e. a lower than average DV could be chosen). These issues should be weighed in selecting year to project emissions in a modeling protocol.

Base Year EI– EPA R6 guidance (continued)

- For HGB modeling, a very high resolution EI is best. Previous Houston modeling and field study data confirm the need for detailed EI.
- If RPO or other monthly/yearly average EIs are used, then this should be compared and weighted accordingly in the analysis conducted with these modeled episodes.

Base Year EI– EPA R6 guidance (continued)

- Most recent forecasting of ozone events in 2005 have been evaluated and it appears necessary to still impute HRVOC EI to get decent model performance.
- This seems to indicate that for historical episodes prior to 2006, that an imputed HRVOC EI is necessary to get adequate model performance. Many HRVOC rules didn't become effective until 2005/6 timeframe.

AD issue – Monitors to be Used

Monitors to be used

- All monitors should be utilized.
- 8-Hr values are less likely to be strongly influenced by THOE type events (short large increases and decreases in ozone due to unusual emissions).
- Separating THOE events that are met driven vs. emission driven and how they may affect a DV at a monitor is also very difficult to quantify.
- Discussion?

AD issue – Calculation of Base DV

EPA 8-Hr Guidance (pages 13-15)

- Several possible methodologies to calculate baseline design values are:
  - 1) The designation design value period (i.e. 2001-2003).
  - 2) The design value period that straddles the baseline inventory year (e.g., the 2001-2003 design value period for a 2002 baseline inventory year).
  - 3) The highest (of the three) design value periods which include the baseline inventory year (e.g. the 2000-2002, 2001-2003, and 2002-2004 design value periods for a 2002 baseline inventory year).
  - 4) The average (of the three) design value periods which straddle the baseline inventory year.

EPA recommended Calculation of Current/Baseline Design Values

- Continue to use the 5 year "weighted" average design values as the basis for future year projections.
  - The 3 design value periods which straddle the base emissions year are averaged (at each monitor) (option 4).
- Supporting analysis of 10 years of DV data shows that the weighted DV is much more stable than a single DV.
- At a minimum a monitor must have 3 years of data.
- The calculation is consistent with CAIR.
- The other three methodologies could potentially be done as a supplemental/WOE analysis.

AD issue – Specific Day/Monitor Model Performance Issues

Specific Day/Monitor Model Performance Issue

- Keep in mind – Detailed Case-by-Case specific analysis needs to be done and all parties need to be conferred and in agreement (including EPA) before doing a WOE type analysis with dropping a monitor on a given day for model performance issues.
- Issues/Discussion??

AD issue #5 – Appropriate size Arrays/ RRF cut-offs and # of days

EPA Default Guidance (page 16)

Table 3.2. Default Recommendations For Nearby Grid Cells Used To Calculate RRF's

Size of Individual Cell, km	Size of the Array of Nearby Cells, unitless
4 - 5[13]	7 x 7
>5 - 8	5 x 5
>8 - 15	3 x 3
>15	1 x 1

Example of Dense monitoring network in Houston

HGB area is “Special”

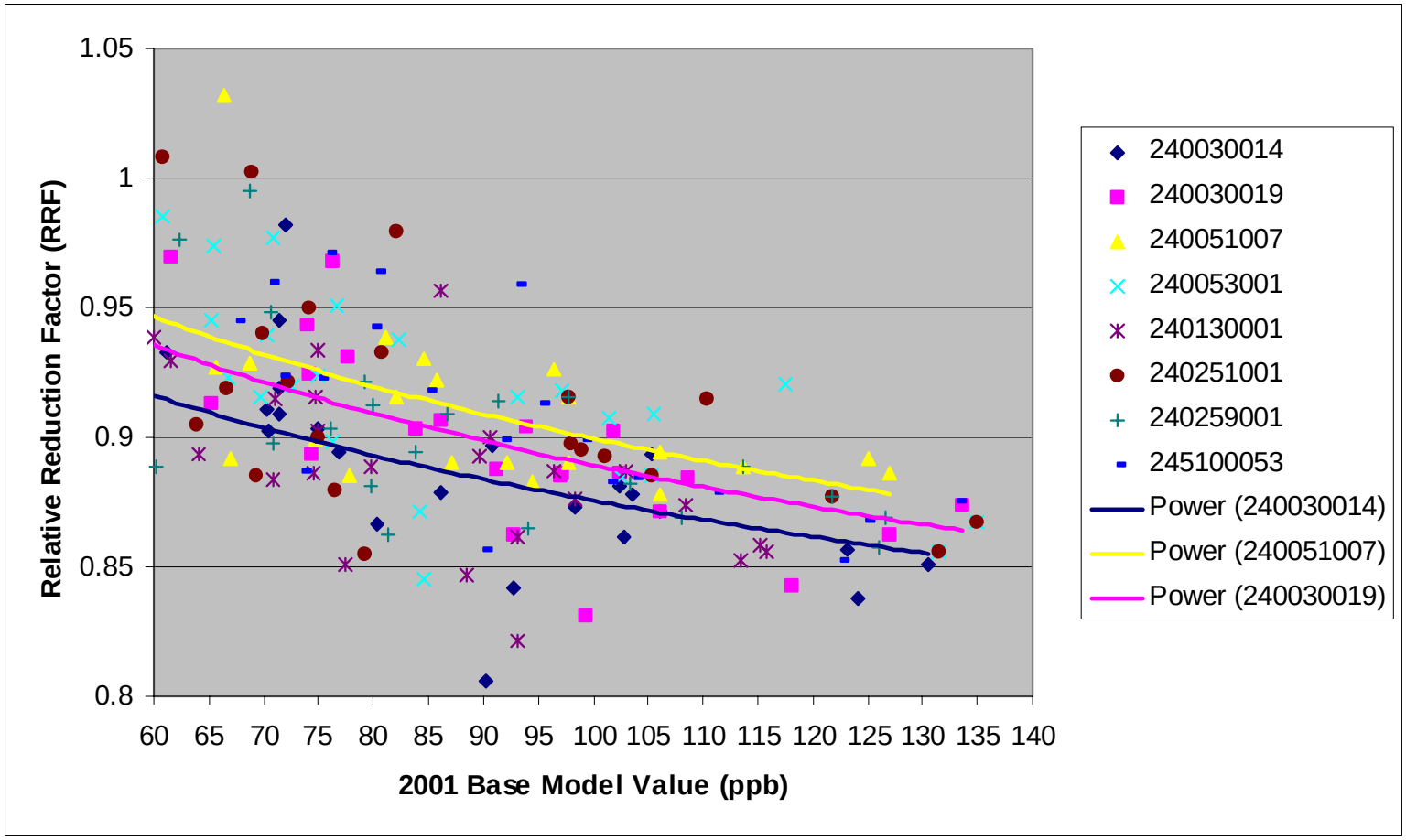
- Many areas of Houston have a high monitor density.
- EPA concept on 7x7 is that a monitor is representative of the 15 km radius circle around the monitor.
- Not necessarily true for Houston.
- EPA is discussing with TCEQ on the appropriate array sizes for each monitor with potentially varying sized arrays throughout the domain.
- It may be interesting to utilize the same arrays from the attainment test as an alternate approach to basecase model performance. Although this test would be an easier test than EPA’s guidance method to use just the grid cell containing the monitor or the bilinear interpolation of 4 grid cells around the monitor.

Minimum RRF Threshold

- Analysis found that the 70 ppb threshold was too low because the model was less responsive at predicted concentrations as low as 70 ppb (the old threshold).
- Recommend increasing the minimum threshold to 85 ppb.
- This removes some bias from the results and also makes the resulting RRF more stable.

Minimum RRF Threshold

Example: Daily Ozone RRF vs. threshold at a monitor in Baltimore

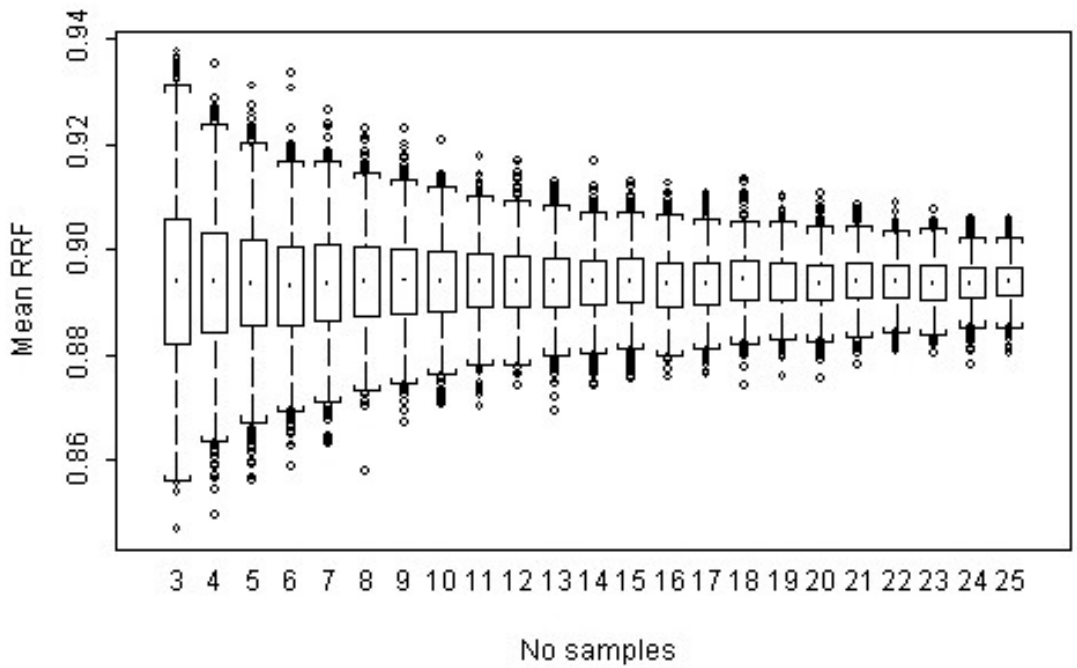


Number of Days to Model

- The guidance said to model “several days” at each monitor
- The recommended minimum is 10 days (with an absolute minimum of 5 days)
- The number of days is based on an analysis which examined the variability of relative reduction factors (RRF) as a function of the number of days in the calculation.
  - The results show that 10 or more days provides for a relatively stable RRF.

Number of Days to Model

Example: Variability of RRF as a function of the number of days in the calculation



As the number of days in the calculation increases, the potential variability of the RRF decreases

Threshold/Number of Days

- The minimum threshold and the number of days are combined to get a hierarchy of the number of days and thresholds used in any possible situation.
  - This will handle all cases where few days are being modeled and/or the modeled peaks are often below 85 ppb.
- The minimum threshold can be as low as 70 ppb and the minimum number of days at each monitor can be as few as 5. Although more days and higher threshold is preferred.

What value to chose from an array?

- High, Average, Best Fit, etc.. ??
- Tom's slides

- EPA guidance is to use the highest 8-Hr value for each day from each array.

AD issue #6 – Unmonitored Area Test

Outside the network attainment test

- Why is this necessary and status of proposed tools.

Where is the Attainment Test Applied?

- The issue is specifically addressed within the unmonitored area analysis.
  - It is assumed that the monitored attainment test will be applied within the nonattainment area.
  - Language concerning areas outside of the NAA:
    "…The unmonitored area analysis for a particular nonattainment area is intended to address potential problems within or near that nonattainment area. The analysis should include, at a minimum, all nonattainment counties and counties surrounding the nonattainment area (located within the State)…. In large States, it is possible that unmonitored area violations may appear in counties far upwind or downwind of the local area of interest. In those cases, the distance to the nonattainment area and ability of the modeling to represent far downwind areas should be evaluated on a case by case basis.""

Why do we need an unmonitored area test?

- 40 CFR Part 50.1 (e) which defines ambient air as ". . . that portion of the atmosphere, external to buildings, to which the general public has access." A letter dated December 19, 1980, from Douglas Costle to Senator Jennings Randolph, reaffirmed and clarified this definition by stating the exemption from ambient air is available only for the atmosphere over land owned or controlled by the source and to which public access is precluded by a fence or other physical barriers. The codified definition plus the 1980 (letter from EPA Administrator Douglas Costle to Senator Randolph) and 1986 clarifications (letter from Director Emison of OAQPS to API) essentially constitute the national policy on ambient air.
- Short/Practical answer – Ambient air NAAQS for attainment demonstrations is everywhere within U.S. jurisdiction
- Other EPA memorandums exist declaring waterways as ambient air as well. These are available through OAQPS's Model Clearinghouse database.
  - From 1985 memorandum (point source modeling):
    - "For example, receptors should be included over bodies of water, over unfenced plant property, on buildings, over roadways, and over property owned by other sources."

Why do we need an unmonitored area test? (Continued)

- 1-hr Ozone guidance –
  - From "Guidance on use of Modeled Results to demonstrate attainment of the NAAQS", June 1996.
    - From the deterministic test discussion…"After a proposed strategy is modeled, the benchmark in this test is to compare the predicted daily maximum ozone concentration in each surface grid cell with 124 ppb on each of the primary episode days."
    - This section continues … "It may happen that the benchmark in this test is very nearly passed, except for isolated modeled day maxima which exceed 124 ppb. In such cases, attainment can still be satisfactorily demonstrated using this approach. First, let us define what we mean by 'isolated'. This is subjective, but generally we mean no more than ~2-3 surface grid cells on any primary episode day.
    - If the predicted exceedances are isolated, then one needs to make the argument that each instance likely reflects the inability of the model to properly characterize the physical/chemical processes accompanying the modeled exceedance.

Attainment Test in Unmonitored Areas

- We eliminated the “screening test” from the Draft guidance and replaced it with an “unmonitored area analysis”.
- Use “model adjusted” spatial fields to estimate ozone concentrations in unmonitored areas.
  - The spatial fields can be created using software provided by EPA. EPA has a work assignment in place to create the software and plan to perform an example analysis. First drafts of tool available in a couple of months.

AD issue #7 – Considerations in utilizing other modeling efforts

Considerations in using other modeling

- Grid Cell size, EI Cell size, etc.
- Meteorological and Ozone model Performance over fine grid and coarse grid domains
- Adequacy of emission inventory to replicate real world fluctuations in basecase/current year.
- How does Future year EI compare with main SIP modeling.
- Level of Documentation/Analyses in comparison to the main SIP modeling. (See Table 6.1 in Final EPA guidance, pages 37-40)

Table 6.1 Recommended Documentation for Demonstrating Attainment of the 8-hour NAAQS for Ozone

Subject Area	Purpose of Documentation	Issues Included
Conceptual Description	Characterization (qualitative and quantitative) of the area's nonattainment problem; used to guide the development of the modeling analysis.	Emissions and air quality assessment; Processes, conditions, and influences for ozone formation.
Modeling/Analysis Protocol	Communicate scope of the analysis and document stakeholder involvement.	Names of stakeholders participating in preparing and implementing the protocol; Types of analyses performed; Steps followed in each type of analyses; Rational for choice of the modeling system and model configurations.
Emissions Preparations and Results	Assurance of valid, consistent emissions data base. Appropriate procedures are used to derive emission estimates needed for air quality modeling.	Data base used and quality assurance methods applied; Data processing used to convert data base to model-compatible inputs; Deviations from existing guidance and underlying rationale; VOC, NOx, CO emissions by State/County for major source categories. Quality assurance/quality control procedures

Table 6.1 Recommended Documentation for Demonstrating Attainment of the 8-hour NAAQS for Ozone (continued)

Subject Area	Purpose of Documentation	Issues Included
Air Quality/Meteorology Preparations and Results	Assurance that representative air quality and meteorological inputs are used in analyses	Description of data base and procedures used to derive and quality assure inputs for modeling; Departures from guidance and their underlying rationale. Performance of meteorological model used to generate meteorological inputs to the air quality model.
Performance Evaluation for Air Quality Model (and Other Analyses)	Show decision makers and the public how well the model (or other analyses) reproduced observations on the days selected for analysis for each nonattainment area and appropriate sub-regions.	Summary of observational data base available for comparison; Identification of performance tests used and their results; Ability to reproduce observed temporal and spatial patterns; Overall assessment of what the performance evaluation implies.
Diagnostic Tests	Ensure rationale used to adjust model inputs or to discount certain results is physically justified and the remaining results make sense.	Results from application prior to adjustments; Consistency with scientific understanding and expectations; Tests performed, changes made and accompanying justification; Short summary of final predictions.

Table 6.1 Recommended Documentation for Demonstrating Attainment of the 8-hour NAAQS for Ozone (continued)

Subject Area	Purpose of Documentation	Issues Included
Description of the Strategy Demonstrating Attainment	Provide the EPA and the public an overview of the plan selected in the attainment demonstration.	Qualitative description of the attainment strategy; Reductions in VOC, NOx, and/or CO emissions from each major source category for each State/county/Tribal land from current (identify) emission levels; Clean Air Act mandated reductions and other reductions; Show predicted 8-hour future design values for the selected control scenario and identify any location(s) which fails the unmonitored area test described in Section 3; Identification of authority for implementing emission reductions in the attainment strategy. Evidence that emissions remain at or below projected levels throughout the 3-year period used to determine future attainment.
Data Access	Enable the EPA or other interested parties to replicate model performance and attainment simulation results, as well as results obtained with other analyses.	Assurance that data files are archived and that provision has been made to maintain them; Technical procedures for accessing input and output files; Identify computer on which files were generated and can be read, as well as software necessary to process model outputs; Identification of contact person, means for downloading files and administrative procedures which need to be satisfied to access the files.

Table 6.1 Recommended Documentation for Demonstrating Attainment of the 8-hour NAAQS for Ozone (concluded)

Subject Area	Purpose of Documentation	Issues Included
Weight of Evidence Determination	Assure the EPA and the public that the strategy meets applicable attainment tests and is likely to produce attainment of the NAAQS by the required time.	Description of the modeled attainment test and observational data base used; Identification of air quality model used; Identification of other analyses performed; Outcome of each analysis, including the modeled attainment test; Assessment of the credibility associated with each type of analysis in this application; Narrative describing process used to conclude the overall weight of available evidence supports a hypothesis that the selected strategy is adequate to attain the NAAQS.
Review Procedures Used	Provide assurance to the EPA and the public that analyses performed in the attainment demonstration reflect sound practice	Scope of technical review performed by those implementing the protocol; Assurance that methods used for analysis were peer reviewed by outside experts; Conclusions reached in the reviews and the response.

Supplemental Analyses and Weight of Evidence

Weight of Evidence/Supplemental Analyses

- The guidance recommends supplemental analyses for **all** attainment demonstrations.
- Less emphasis on modeling (CAAAC recommendations)
  - Increased weight to emissions and ambient data trends and analyses
- Weight of evidence range redefined as 82-87 ppb
  - 85 ppb becomes less of a bright line

Questions to be answered

- Why are supplemental analyses and weight of evidence needed for attainment demonstration purposes?

- What does the final ozone guidance say about supplemental analyses and weight of evidence?
  - Do attainment demonstrations have to include supplemental analyses?
  - How are supplemental analyses different from weight of evidence?
  - Are all analyses eligible for a weight of evidence attainment determination?
  - Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
  - Do you have any examples of what a weight of evidence analysis might look like?

- What changes are being considered regarding weight of evidence resulting from comments on the draft final ozone guidance?

Why are supplemental analyses and weight of evidence needed for attainment demonstration purposes?

- By definition, models are reasonable but simplistic approximations of complex phenomena.
- Model inputs (emissions, meteorological, IC/BC, landuse, etc.) and model code are subject to uncertainty.
- In recognition of these uncertainties, the regional modeling community has evolved over the past decade:
  - from using models in an absolute sense w/ attainment “bright lines” (90s),
  - to using the model in a relative sense to project an expected change in ozone which can be used to assess whether attainment will be reached in the future via a “bright line” determination (recent national rules),
  - to using the modeling as the primary element in a suite of tools that assist in an assessment of present and future air quality trends (next round of attainment demonstrations).

Why are supplemental analyses and weight of evidence needed for attainment demonstration purposes?

- As part of their recommendations to transform the SIP process into one that is more performance-oriented, the CAAAC AQM workgroup recommended increased use of weight of evidence within State/Local attainment demonstrations.

- *"EPA, in conjunction with S/L/T and affected stakeholders, should modify its guidance to promote weight-of-evidence (WOE) demonstrations for both planning and implementation efforts. In particular, these demonstrations should reduce reliance on modeling data as the centerpiece for SIP/TIP planning, and should increase use of monitoring data and analyses of monitoring data, especially for tracking progress. Enhanced tracking and ambient monitoring data is a better use of available resources than intensive local modeling."*

- Guidance notes that AQ models still represent best tool for integrating emissions and meteorological data with known chemistry.

What does the final ozone guidance say about supplemental analyses and weight of evidence?

- All attainment demonstrations should include supplemental analyses to assess the validity of the modeled attainment test projection.
  - Guidance suggests amount of supporting information should be greatest for those areas with modeled projections near the NAAQS.
- Three separate types of supplemental analyses should be considered within an attainment demonstration.
  - Additional air quality modeling evidence
  - Trends in ambient air quality and emissions
  - Other air quality analyses that can identify potential control targets

What does the final ozone guidance say about supplemental analyses and weight of evidence?

- "Weight of evidence" differs from "supplemental analyses" in that:
  - 1) WOE is a set of supplemental analyses for areas whose attainment test results indicate future AQ levels near the NAAQS.
  - 2) WOE combines and weights the various supplemental analyses with the results of the attainment test … the end result being an aggregate, weighted, and ultimately subjective conclusion as to whether a set of control strategies will yield attainment by the relevant future year.

What does the final ozone guidance say about supplemental analyses and weight of evidence?

- Can all attainment demonstrations use weight of evidence aggregations to "reverse" the conclusion from the modeled attainment test?
  - Yes, in both directions
    - model projects attainment, yet WOE suggests nonattainment; or,
    - model projects nonattainment, yet WOE suggests attainment
  - However, the further the attainment projection is from the NAAQS, the more compelling the contrary evidence produced by corroboratory analyses must be to draw a conclusion differing from that implied by the modeled attainment test results.

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
Element 1: Additional AQ Modeling

- In some cases, there is still a role for the absolute model predictions to show the degree of expected AQ improvement in the future.
  - Limited to modeling applications that feature solid performance.
  - Several metrics could be considered:
    - % change in total ozone >= 85 ppb w/in nonattainment area
    - % change in grid cells >= 85 ppb w/in nonattainment area
    - % change in grid cell hours >= 85 ppb w/in nonattainment area
    - % change in maximum modeled 8-hr ozone w/in nonattainment area
  - There are no thresholds in these metrics that are necessarily indicative of attainment.

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration? (Region 6 thoughts)

Element 1: Additional AQ Modeling

- Evaluate daily RRF type calculations at each monitor and develop a more complete understanding of what is happening on an individual day basis at each monitor.
- Research day/monitor specific phenomena using process analysis, source apportionment, Ox production, etc.
- Evaluate higher cut-offs for RRF as long as minimum of five days is used.
- Source apportionment of future year 2009 model estimations with different levels of cut-off for the source apportionment grid cells (apportionment for grid cells above 80, 85, and or 90 ppb).

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration? Element 1: Additional AQ Modeling

- Example of Ozone WOE discussion:
  1) Absolute modeling metrics indicate substantial improvement in future-year O3 >= 85 ppb;

  2) Yet, there remains a small set of monitors/cells that the attainment test shows will remain nonattainment;

  3) However, there is demonstrable reason to believe that the model response is less certain in this area (e.g., poor performance).

8-hour Ozone Projections (2010)	Base	Controlling cell	All cells
Projected Design Value	93.0	**86.7**	N/A
Average change in 2002-2010 projected ozone DV	N/A	6.3	11.3
% Reduction in 2010 total nonattainment	N/A	84.2%	97.5%
% reduction in pop-weighted 2010 total NA	N/A	84.2%	97.8%
% reduction in NA area exceedance counts	N/A	73.1%	86.9%
% reduction in NA area exceedance days	N/A	75.0%	87.5%
Absolute maximum modeled ozone in NA areas	91.0	90.3	86.1

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
Element 1: Additional AQ Modeling

- Use of available regional (or other local) AQ modeling projections
  - In order for SA/WOE usage, one would need to make the case that the regional modeling was appropriate for use in the local area (e.g., episodes, base/future emissions, performance, etc.)
- Multiple AQ Models / Input Data Sets (aka, alternative basecases)
  - There may be different, technically-plausible combinations of models, model physics options, and model input data sets that yield acceptable base year model performance.
  - Focus on the sensitivity of estimated relative reduction factors to variations in inputs or model formulations
- Use of same modeled attainment demonstration but w/ DV_F values calculated in an alternative manner than in Guidance
  - Requires strong justification for why approach is equally valid

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration? Element 1: Additional AQ Modeling

- Example of Ozone WOE discussion:
  1) CAIR modeling for the area projects a 2010 value of 83.5.

  2) There is evidence that the regional modeling is an equally valid future projection.

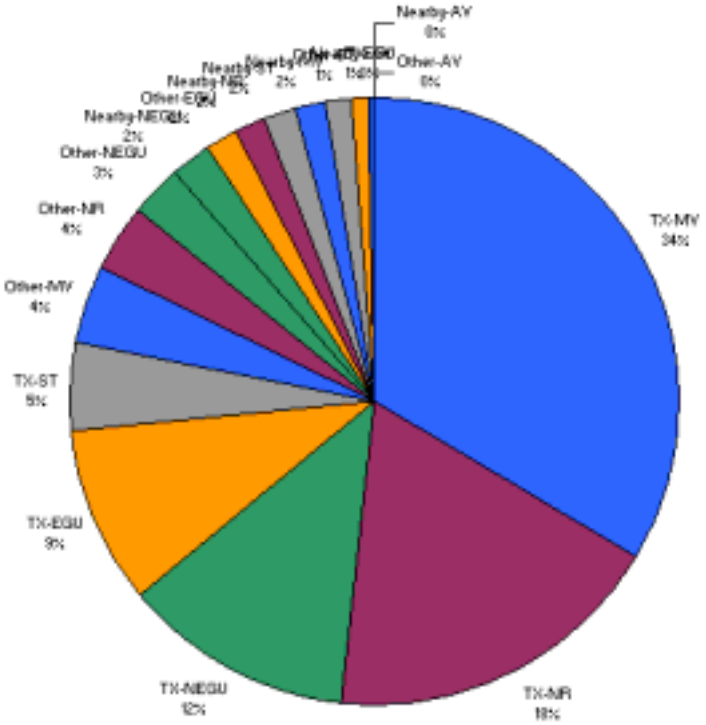
  3) Sensitivity analyses have shown that CB4 (used in base case) and SAPRC mechanisms are equally valid and perform equally well, but can cause small variations in future DV projections.

8-hour Ozone Projections (2010)	**Base (CB4)**	**Base (SAPRC)**	**CAIR**
Projected 2010 Design Value	86.7	84.4	83.5
Appropriate Episodes	Yes	Yes	somewhat
Appropriate Emissions (base/future)	Yes	Yes	mostly
Average model bias in base year	3.2%	-4.7%	-2.1%
Average model error in base year	18.4%	19.6%	21.2%
Subjective confidence in results	high	high	medium

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?

Element 1: Additional AQ Modeling

- Use of alternate AQ modeling techniques to help explain why attainment should (or should not) have been demonstrated.
  - Source apportionment
  - Response surface modeling
- There may be other technically defensible ways to utilize modeling results to show attainment beyond those shown here.



Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
Element 2: Emissions and AQ Trends

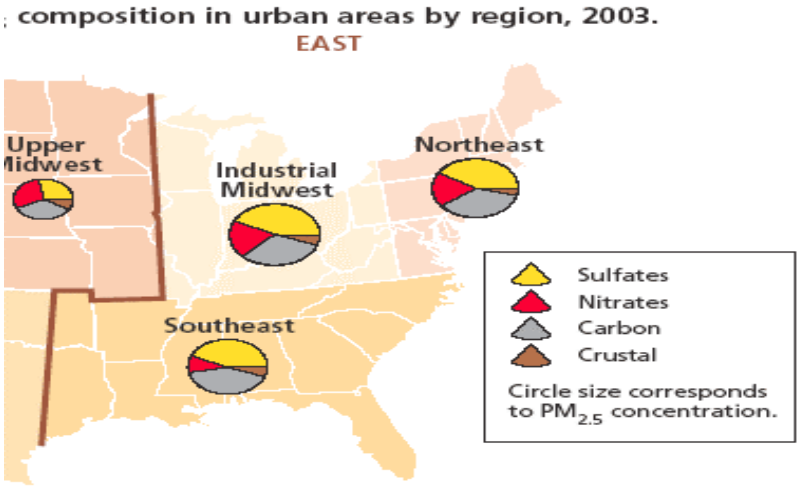
- May be possible to extrapolate future trends in 8-hour ozone based on measured historical trends of air quality and emissions.

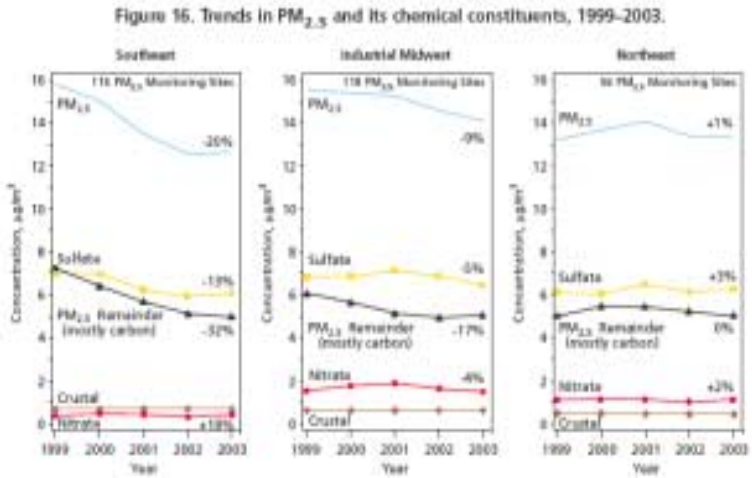
- Several complicated elements to these analyses
  - Must normalize historical AQ data trends to account for year-to-year meteorological variations.
  - Must have a solid conceptual model of PM/O3 formation in the area of interest (e.g., which PM species dominate, how are they formed)
  - Must have an accurate accounting of year-to-year changes in actual emissions that lead to high PM/O3 over a given area.

- Goal is to develop a curve that relates past emissions changes to differences between historical and current AQ.
  - Curve can then be extrapolated to account for expected changes in relevant emissions by the attainment year.

Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
Element 2: Emissions and AQ Trends

- Example of PM2.5 WOE discussion:
  1) For hypothetical area in question, sulfates and carbonaceous mass are known comprise the majority of the PM.
  2) Meteorologically-adjusted AQ trends indicate that sulfate and carbon trends are downward in a period of declining EC and SOx emissions (not shown).
  - 3) Assume that a linear extrapolation of observed past correspondence between monitored data and estimated emission changes will describe future air quality.
- This could be done with Auto GC data.





Specifically, what types of supplemental analyses should be completed as part of an attainment demonstration?
Element 3: Other AQ analyses

- Observational models can take advantage of monitored data to draw conclusions about the relative importance of different types of PM2.5 emissions or ozone precursors emissions.
  - Strength of evidence is proportional to the completeness of the ambient data base

- Possible AQ analyses for ozone attainment demonstrations
  - Indicator species (VOC/NOx limitation)
  - Trajectory modeling
  - Receptor modeling (CMB, PMF)
  - Hybrid models

Weight of Evidence/Supplemental Analyses: Summary

- All attainment demonstrations should include supplemental analyses to corroborate the modeling results. Supplemental analyses may include:
  - Emissions trends
  - Ambient data analysis and trends (including meteorologically adjusted trends)
  - Receptor-based and/or observational model analyses

- For those areas where attainment test projections are close to the NAAQS, an aggregate weight of evidence determination should be made subjectively based on the results of the supplemental analyses
  - For ozone, “close to the NAAQS” is defined as 83-87 ppb.
  - For PM2.5, “close to the NAAQS” will be defined in the guidance.

- For those areas w/ projections further removed from the NAAQS, WOE can still be used, but it is less likely to “reverse” the modeling test results.

Weight of Evidence/Supplemental Analyses: Summary

- The weighting factors associated w/ each of the individual supplemental analyses should weigh each type of analysis according to:
  - the credibility of the analysis, as well as
  - its ability to address the question being posed (i.e., is the strategy adequate for meeting the ozone NAAQS by a defined deadline?).

- The conclusions derived from the weighted supplemental analyses are combined to make an overall assessment as to whether meeting the air quality goal is likely. This last step is an unavoidably qualitative one involving some subjectivity.

EPA's Next Steps

- Continue to work with TCEQ in finalizing the HGB protocol.
- Provide follow-up to unresolved questions.