A Modeling Protocol
For Assessment of the Ozone
National Ambient Air Quality Standards
in the Ozone Transport Region

DRAFT

The Modeling Committee of the
Ozone Transport Commission

December 2001

BACKGROUND

The Ozone Transport Commission (OTC) Modeling Committee has embarked on the task of exercising photochemical grid model(s) as a means to gain the necessary experience among its members. (OTC Modeling Work Plan, 2001) The background and additional details for this initiative are contained in the work plan, including a list of acronyms, and a planned implementation schedule. This modeling protocol (contemplated by the work plan) outlines the process and procedures for model assessment and execution. As part of this process, the period of June 11 to July 28, 1997 has been selected for evaluation. Meteorological data were developed for this period using MM5 at 12 km resolution (Zhang, 2000). Emissions data reflecting 1996/97 were assembled as part of the MARAMA RTC effort (Havens, 2000). Spatial surrogate data were developed for the unified grid (Sistla, 1999) at 12 km resolution. The emissions and met data require processing to suit the selected photochemical grid model(s).

METEOROLOGICAL DATA

The MM5 setup has been described by Zhang (2000) for generating meteorological fields that are based on the Gayno-Seaman scheme for the boundary layer calculations. The projection of the grid is in the Lambert conic conformal (LCC) system, covering the continental United States at 36 and 12 km spacing in the horizontal. In the vertical, the domain extends to about 17 km, with level 1 corresponding to about 10 m above the ground.

Given the design of the met data, transfer of these data to the photochemical model has been in some instances associated with interpolation due to non-matching of the grids or to an entirely different geographical projection. In this study, the transfer of the data to the photochemical model would be essentially on a one-to-one basis with limitations on the top layer being set in the 4 to 6 km range by collapsing some of the levels beyond the 2 km height. However, there is no reason to assume that the vertical structure based upon MM5 is the appropriate approach to simulate pollutant concentrations. As such, at least one or two sensitivity simulations consisting of modified vertical grid structures will be simulated as part of this effort. Based upon the relevant analysis of these simulations, a vertical grid structure will be selected for further applications. In all these cases, the photochemical model will be operated in the native mode of the meteorological data.

EMISSIONS DATA

Currently, emissions data can be processed using EMS95/EMS2000, and SMOKE. The surrogate data files for the unified grid have been developed and would be used in this study. For those pollutants that depend upon ambient temperature, MM5 layer-1 gridded temperature fields will be used. The intent is to process the emissions data for the entire period for which met data are available in the LCC system at the 12 km resolution.

The processing of the emissions data will be accompanied with several quality checks before the data are exercised in the simulations. Prior experience has shown that considerable time

and resources are often invested in developing the gridded emissions data. Currently, efforts are underway to make final corrections in the emissions inventory files that were created by MARAMA. The OTC Modeling Committee will set a date beyond which no further corrections will be made for purposes of this initial exercise. Any further changes or corrections will be archived and incorporated at a later date. In performing this work, close attention will be paid to the emissions within the OTR and, if necessary, corrections will be incorporated with the advice of the OTC Modeling Committee (or a sub-group).

PHOTOCHEMICAL MODEL

Currently several models are available for evaluation and assessment. Some of the models are UAM-V (1.34), CAMx, MODELS3/CMAQ, among others. While it is impossible to assess the merits of each one of these models, the goal is to exercise at least two models, which are capable of accepting input data in the native mode of LCC, and use the same chemical scheme such as CB-IV. In this way, inter-comparison of the models' ability to simulate pollutant concentration fields would not be considered as dependent upon transformation and interpolation schemes.

- Metrics and measure to assess model performance: surface layer

The modeling system will be exercised over a modeling domain that covers the eastern United States. Prior experience has shown that the model cannot be expected to have the same level of performance everywhere in the domain for every hour of simulation and for all days simulated. Thus, some guidelines or markers are to be defined to provide un-biased assessment of the model performance. Also, we need to investigate if it is meaningful to assess model performance based on the statistical measures identified in EPA (1991) or if other alternate metrics should be evaluated. Another issue that requires a consensus is the geographical grouping of the monitors or, in other words, the areal extent up to which the measured and predicted data can be compared. In prior studies, the definitions of CMSA or MSA were used; however, such a usage does not consider pollutant transport patterns but relies on demographics. While these definitions may be used as an initial screening of the data, a more refined approach is to be investigated and adopted for both the 1-hr and the upcoming 8-hr NAAQS for ozone.

Another important aspect is the requirement for the examination of the precursors. Since these data are becoming more readily available, it is important to establish performance measures for ozone precursors, including CO.

APPROACH

It is suggested that the model simulation and assessment be performed for the high ozone event of July 10 to 14, 1997. The specifics of the episode need to be agreed upon as well as the methods for the assessment. There are several other issues that would require reaching consensus, some of which are: initial and boundary conditions, number of spin-up days, etc. Each photochemical model requires pre-processing of the emissions and meteorological data for each episode day, as well the boundary conditions.

After performing the simulation, the data are extracted for the surface layer for selected pollutants for assessment of the model.

EMISSIONS

The emissions data are based upon EPA's 1996 NET inventory which has been updated and revised, taking into consideration state-level data as well revised off-road emissions data. These data are adjusted to reflect the episode period by appropriate corrections from 1996 to 1997 conditions. After growing the emissions from 1996 to 1997, the emissions are processed using the appropriate preprocessor. It is expected that the differences between 1996 and 1997 will be minor, with the exception of those sources that are temperature sensitive.

POINT SOURCES – ELECTRICAL GENERATION UNITS (EGUs)

Based upon the information available, the EGU database for 1997 will be adjusted to reflect the July 1997 period. Every effort will be made to match the 1996 data with the corresponding 1997 emissions data and in those instances that lack correspondence would be identified and resolved through contacting the responsible parties. It is expected that the number of such cases will be minimal as the emissions data are being adjusted to reflect the consecutive year.

POINT SOURCES – NON-EGUs (NEGUs)

To make the NEGU data to reflect 1997, these emissions will be adjusted by the growth rate information that is reported or available on a state-by-state basis. Again, since the change is reflective of consecutive years, the adjustment rate is expected to be at or near unity for most source categories.

AREA AND NON-ROAD SOURCES

Similar to the NEGUs, the area and non-road source emissions are adjusted using growth rate information.

MOBILE SOURCES

On-road mobile source emissions are day specific; therefore, vehicle miles traveled (VMT) data are adjusted from 1996 to 1997. The programs EMS95 and SMOKE have built-in MOBILE 5b input information to estimate the emissions, based upon VMT and temperature.

BIOGENIC EMISSIONS

Biogenic emissions will be prepared for each episode day using BEIS2 initially. Other models, such as BEIS3, BIOME, and GLOBEIS, will also be included as possible candidates. The temperature data from MM5 layer-1 will be used along with cloud cover information obtained

from MM5.

AIR QUALITY INPUTS

- **Initial conditions**

Prior experiences have shown that a 3-day ramp-up period is sufficient to build pollutant levels that are encountered in the beginning of the ozone episode. Thus, the initial conditions at the startup would be what are termed clean conditions, with emissions and meteorology data from the input database.

- **Boundary conditions**

Often, 3-dimensional pollutant concentrations are not available as required by the 3-dimensional grid models. Thus, the datum provided is the clean conditions similar to those assumed in the initial conditions. In prior studies attempts have been made to include any information that is available from ozonesonde and from monitors that are near the western and northern boundaries of the modeling domain. For this study, similar attempts will be made for ozone and other pollutants to obtain pollutant data at the boundaries.

- **Grid definition for the photochemical model**

While the horizontal grid spacing is fixed by default set forth by the meteorological model, in this case 12 km, the definition of the vertical structure is still a subject for discussion and reaching consensus. In general, as the number of levels in the vertical increases, so does the computational time as well as the storage requirements for the data. On the other hand, limiting the vertical resolution to a few layers would inherently discard the detailed information provided by the meteorological model. Thus a compromise solution would be to maintain the high resolution with a one-to-one design of the vertical layers up to approximately 1.5 km yielding about 7 levels, followed by a collapse of the MM5 upper layers from 1.5 km to about 6 km into another 7 levels. Thus, under this scenario there would be a total of 14 layers with 7 levels below 1.5 km and 7 levels above 1.5 km. This configuration may or may not be optimal; therefore, provisions will be made to have discussions within the OTC Modeling Committee.

- **Ambient air quality data**

Ambient air quality data will be extracted from the EPA AIRS archive for ozone, CO, NO, NO2, NOx, and for total and speciated hydrocarbons reported as part of the PAMS network. Since the modeling domain extends over two time zones, while the model simulations are reflective of a single time zone, EST, there will be a need to ”correct” the clock and assemble the ambient air quality database. Any special measurements conducted during the summer of 1997 will also be acquired, including upper air measurements.

MODEL SIMULATIONS

As stated above, the model simulations will be designed with a 3-day ramp-up followed by the episode days. The layer-1 concentrations will be extracted and archived for analysis. The

upper-level database will be processed on an as needed basis due to the voluminous nature of the information.

MODEL PERFORMANCE

This is an area that will likely require substantial dialog among member states. While there are many statistical tests that can be applied, it is important to define a priori some of the conditions of the analysis and the targets of evaluation. Also, it is quite critical to define the areal extent for which the assessment needs to be done to address the performance of the model. It is suggested that a "strawman" outlining the procedures be developed as a part of the application of this analysis. The strawman could have 3 or 4 levels of tests covering all or portions of the domain using daily maximum, daily average, daytime average, 8-h maxima, and other combinations of ozone concentrations to assess the model performance. Similarly, the statistical tests are to be applied to the precursor data as well, recognizing that all tests applied to the ozone data may or may not be valid. Thus, it is suggested that the statistical tests be geared appropriately based on the quality and quantity of the database.

As part of the model assessment, qualitative analysis will also be performed by comparing predicted and measured pollutant fields to establish if the spatial patterns are captured by the modeling system. This is a critical step, since the measured concentrations may fall into a neighboring grid cell (but not at the measured location itself) and may be found to be in good agreement.

Another area that is quite important is the predictive ability of the model with respect to height. Recognizing that the pollutants trapped above the mixed layer during the overnight hours would mix down during the daytime, comparison will be made between measurements and model predictions. Also, as noted in the work plan, special attention will be paid to elevated monitoring stations, such as the television tower near Durham, North Carolina; and the Sears Tower in Chicago, Illinois, and any other special monitors.

SUGGESTED PHOTOCHEMICAL MODELS

The work plan has indicated essentially three models -- Models3/CMAQ, UAM-V, and CAMx -- that are currently available for use by any interested party. Several of these models in recent years have included enhancements that facilitate sensitivity analysis and other assessments. As stated in the work plan, every attempt will be made to exercise these models for the July 1997 episode.

DOCUMENTATION

All analysis will be documented including any revisions to the protocol and the strawman on model performance. This technical document will be made available to interested parties.

REFERENCES

OTC Modeling Committee (2001). A Modeling and Analytical Work Plan for Assessment of the National Ambient Air Quality Standards for Ozone in the Ozone Transport Region (see http://www.sso.org/otc/Modeling/MOD.htm).

Wick Havens (2000): Development of an Emissions Inventory for Regional/Urban-scale Modeling, MARAMA-RTC (see http://www.marama.org/).

Gopal Sistla (1999): Development of a surrogate database for use in Regional/Urban-scale Modeling at 4 km spatial resolution (see http://envpro.ncsc.org/emcenter/).

Dalin Zhang (2000): Development of meteorological database for summer 1997 using MM5 at 12 km resolution in Photochemical Model Simulations.