

## A Modelling System for Mechanism Evaluation Using Chamber Data

Paul A. Makar

Ping Y. Li

Atmospheric Environment Service

4905 Dufferin Street

Downsview, Ontario

Canada M3H 5T4

paul.makar @ ec.gc.ca

Gas-phase chemical mechanisms for use in regional or global tropospheric atmospheric chemistry applications require validation against measurement data in order to assure the modeller and policy-makers that the mechanism adequately represents the chemical processes of the real atmosphere. One means of mechanism validation is through the use of smog-chamber data, in which atmospheric chemistry experiments are performed under controlled conditions, and eliminate the potential confounding effects of non-chemical processes in the ambient atmosphere. The absence of the ambient atmosphere's confounding factors (in which some of the chemical changes may be due to advection, diffusion, emissions, deposition, convection, rainout and washout) allow a worthwhile comparison of chemical effects alone.

The need for a means of mechanism verification was identified as one of the components of the Atmospheric Environment Service Unified Regional Air-Quality Modelling System (AURAMS). This is a regional gas and particulate matter modelling system which is the subject of current research at the AES Laboratory in Downsview, Ontario. One of the key components of this modelling system is the gas-phase chemistry module, which includes an operational simulation system used within the regional model itself, and a mechanism validation module to be used in determining the validity and expected performance of a mechanism when used within the regional model. A description of the validation module its first applications are the subject of the current work.

The research requirements for the validation module were such that the system would have to be capable of performing the following tasks:

- (1) Assessing the accuracy of gas-phase reaction mechanisms.
- (2) The use of the mechanisms in the module would have to be consistent with their use in the regional model (with regards to lumping of emissions and calculation of reaction rates).
- (3) Provide a means of immediate verification/assessment of updates to gas-phase chemistry, so that the regional model's chemical mechanism is always current and up-to-date.
- (4) Use recognized statistical measures of mechanism performance.
- (5) Create and store input and output in a format allowing simple and direct comparison of model and measurements – the i/o must be portable to the greatest extent possible.
- (6) The validation of a mechanism and/or its components must be performed in a small amount of processing time, preferably on a vector supercomputer.

Any system of mechanism validation will have five major components. These include: a process for reading in measurement data, a conversion system whereby the raw data is converted into a format suitable for modelling, a simulation system which numerically predicts concentrations based on the initial conditions (starting concentrations in the chamber) and boundary conditions (temperature, pressure, humidity, light intensity, wall effects) of each experiment, a package for comparing the model results to data, and the creation of summary statistics as the final product of the system.

In the AURAMS mechanism validation system, the measurement data is read in by a suite of chamber-specific programs (UNCREAD, SAPRCREAD, DTCCREAD) which also convert the measurements to a common format, in which all data appear on the same time interval. The concentrations of all measured species are linearly interpolated to the same time interval as the  $\text{NO}_x$  and  $\text{O}_3$  measurements. The  $\text{NO}_x$  and  $\text{O}_3$  values usually have the finest time resolution of the measurement record in a typical chamber experiment – the linear interpolation results in no loss of information. The use of linear, rather than higher order, interpolation prevents the potential creation of spurious maxima or minima in the measurement record. The interpolation allows a more direct comparison of measurements to model results, since the latter are constrained to operate on a single time step. The time-interpolated data from this stage are stored as spreadsheet readable time series, each row in a record being the measurements (or interpolated measurements) specific for a given time. The speciation for this intermediate output is the same as that of the measurement record itself. Time series records of chamber boundary conditions at the measurement times of  $\text{NO}_x$  and  $\text{O}_3$  are also created at this stage.

The second stage of data conversion (program CONREAD) has several subcomponents. Any measurements reported as total concentrations of pre-set mixtures of hydrocarbons are converted to the detailed speciation through the use of a fractionation input file. The hydrocarbon measurements are then lumped in the same fashion as detailed emissions in a regional model (Middleton et al., 1990). That is, the speciation employed in the regional mechanism will in general be simpler than the speciation in the measurement record; the measurements must therefore be lumped in a fashion consistent with the lumping of the highly speciated emissions data available for regional modelling. The detailed, measured species are first assigned to an “emitted” speciation. This assignment is done on a mass-weighted basis, and provision is also made for additional weighting factors. The intermediate, “emitted” species are then lumped into a “model” speciation. This second stage of lumping is performed using reactivity weights designed to preserve the total organic peroxy radical formation rate, of importance in accurately simulating ozone formation. Again, provision is made for additional or other weighting factors, to allow for other means of species compression. The final result of these operations is a second data set, time interpolated to the  $\text{NO}_x$  and  $\text{O}_3$  measurement interval, with all hydrocarbon measurements in the speciation of the mechanism for which validation is desired.

Following the creation of the mechanism-specific set of chamber data, simulations are performed using the initial concentrations from these files and the chamber boundary conditions for a set of experiments specified by the user. These are done using a third code (CHAMRUN), which makes use of a vectorized Gear-type solver usually employed in simulations of regional or global atmospheric chemistry (SMVGEAR, Jacobson and Turco, 1990). As part of the processing, three important modifications are made to the data to allow for the use of this very accurate and

very fast numerical solver. The constraints posed by the solver (due to its intended use in a regional model) are a single number of integrations for all experiments, and the use of a single time-step between updates to reaction rates for all experiments.

First, the chamber boundary conditions data, after being read in by the simulation model, are appended so that the conditions for each experiment all have the same total number of measurements. This is done in two stages – the current set of simulations is searched to find the simulation with the greatest total number of measurements, and the last set of measurements in the remaining experiments is repeated until all boundary conditions have the same total number of measurements. Simulations are performed for this “maximum” number of steps, but only the actual number of steps for any given experiment is retained on output.

Second, a default time step between updates to chamber boundary conditions is set (these boundary conditions give rise to the reaction rates). The reaction rates for each experiment are then multiplied by the ratio of the actual time between updates to this default value. The simulation model thus operates on a default timestep, while the scaling of the rates ensures that the same model produces the same results as if different time steps were employed for each simulation to be performed.

Both of the above modifications can be classified as mathematical tricks which allow the use of a regional model simulation package on chamber data with a varying number of measurements and time-steps between data. The modifications have no effect on the accuracy of the simulations, but allow the processing of hundreds of chamber experiments on a vector supercomputer in a few minutes of clock time.

CHAMRUN also automatically adds variables that track the mass through every reaction in the mechanism under study. This allows easy process analysis of the mechanism, and comparison to other mechanisms analyzed using the system.

The output from the simulation code takes the form of columns of model variables on the same time interval as the interpolated measurements, in the same format as the input data. Any spreadsheet or graphics package may be used to compare time series of model and measured variables for any individual experiment or set of experiments. For the purposes of providing objective evaluations of mechanism performance, a final stage of processing takes place (program CHAMSTAT). This program reads the model-speciated measurement and model simulation time series, and retrieves the maxima of ozone, NO<sub>2</sub>, formaldehyde, and PAN. The time taken to reach these maxima, the time to the NO/NO<sub>2</sub> crossover, and the average value of the change in (O<sub>3</sub> – NO) per step as a function of time, are also retrieved. These are outputted in a final table, again as columns readable by spreadsheet codes or graphics packages.

During the course of the creation of the system, a few minor problems were encountered with a small number of the measurement files, which prevented their further use. These are mentioned here for the purposes of providing feedback for future chamber experiments. The problems encountered included repeat time stamps (same time being repeated with different measurements), different names being used for the same chemical species, and multiple

measurements of the same species, with no information being given as to the expected relative accuracy of the different measurements.

As an example use of the system, a set of 14 Ethene experiments from the UNC database were selected for processing with the AURAMS mechanism. The conversion of the whole UNC database to the model speciation using CONREAD required about 3 minutes of processing time on a workstation. The corresponding simulations for the set of experiments (each of duration of about 10 hours) required less than a minute of processing time using CHAMRUN on a vector processor. In the next few months, the system will be used to validate the AURAMS mechanism, and compare its performance to that of the ADOM-II mechanism which it supercedes and the Master Chemical Mechanism.

#### References:

Jacobson, M.Z. and R.P. Turco, SMVGEAR: A sparse-matrix, vectorized Gear Code for Atmospheric Models, *Atm. Environ.*, 28, 273-284, 1994.

Middleton, P., W.R. Stockwell and W.P.L. Carter, Aggregation and analysis of volatile organic compound emissions for regional modelling, *Atm. Environ.*, 24, 1107-1123, 1990.