# EMS2001

Gwendolyn Judson, Wisconsin Department of Natural Resources
Mark Janssen, LADCO

**Abstract**: EMS95, released in 1995, was designed to meet the needs of OTAG, the Ozone Transport Assessment Group. Now, six years later, additional enhancements to this emissions model warrant the release of a new version: EMS2001. The model has been updated to include Biome3; include Speciate 3.0 profiles; eliminate the need for ARCINFO to grid point sources; convert between UTM (Universal Transverse Mercator), latitude/longitude and Lambert Conformal Conical projection systems; improve QA/QC reporting; and include a full version that runs on an NT platform. The release date of EMS2001 is May 1, 2001.

When EMS95 was released in 1995, the main concern of those responsible for its development was to create an emissions model that could be used by OTAG, the Ozone Transport Assessment Group, to apply growth and control factors to a large domain for the purpose of projecting a base emission inventory to a future year as part of the effort to evaluate ozone control strategies. Now, six years later, the focus has shifted; emissions models need to encompass larger and larger domain areas, be flexible enough to include more pollutants and data submitted in different formats, produce results in a reasonable amount of time, and be affordable and simple enough for the more casual user to run. EMS2001 has many new features designed to address these issues including a biogenics module, BIOME3; SPECIATE 3.0 profiles; processors which eliminate the need for ARCINFO to grid point sources; the ability to convert between UTM (Universal Transverse Mercator), latitude/longitude, and Lambert Conformal Conical projection systems; improved QA/QC reporting; and a full version that runs on an NT platform.

BIOME3, produced for LADCO (Lake Michigan Air Directors Consortium) by Jim Wilkinson of Alpine Geophysics, heralds the return of biogenic emissions modeling to EMS2001 and, like the rest of EMS2001, it is written in SAS. Use of this module provides flexibility; both the BEIS2 and BEIS3 emissions algorithms have been included, allowing the user to choose which algorithm to use in the SAS command line. BIOME3 takes advantage of the BELD3 (Biogenics Emissions Landuse Database) data coverage. This database covers the entire United States and parts of Mexico and Canada and provides 1 km by 1 km resolution of 26 land use types and 230 plant species. Processors in BIOME3 allow the user to extract only those tiles needed from the BELD3 coverage to cover the desired modeling domain, which reduces the size of the data set and improves run time. When run in BEIS3 mode, the user can also define the number of layers to model in the canopy. BIOME3 processes the BELD3 data coverage, meteorological data, and the BIES3 emission factor table and using either the BEIS2 or BEIS3 emission algorithms creates a SAS data set that identifies the FIPS state ID, FIPS county ID, icell (x-direction index of modeling grid), jcell (y-direction index of modeling grid), BEIS3 species code, pollutant identifier, and hourly emission estimates (ug/hr). The addition of BIOME3 allows the EMS2001 user to create episode specific biogenic data sets for any grid domain within the boundaries of the BELD3 data.

Six years ago the focus of regional emissions modeling was ozone precursors; now more states and other users are concerned with haze and want to be able to model particulates on a regional scale which requires a better picture of the elements of particulates included in the inventory. Even for use with ozone precursor inventories, the speciation profiles in EMS95 were very out of date and did not reflect current practices or formulations. Incorporating the speciation profiles from SPECIATE 3.0 updates EMS2001 to the most recent readily available data. SPECIATE is EPA's repository of total organic and PM speciated profiles for a wide range

of sources. In 1999, USEPA released SPECIATE 3.0, a database with more than 300 source types, 262 new TOC (total organic compound) profiles and 13 new PM (particulate matter) profiles. EMS2001 includes input processors that can directly read the database version of the SPECIATE 3.0 files. These new speciation profiles provide a much more accurate picture of industry practices and formulations currently in use.

Some improvements are designed to make modeling with EMS2001 easier. Many times an emissions modeler is called upon to model a new grid or to run a strategy based on proposed controls. In the past, anytime the number or position of point sources was changed, new point source grid data files needed to be created using ARCINFO. This required the user to own the ARCINFO software, maintain an updated license and be trained to use ARCINFO. With EMS2001, this is no longer the case. A SAS program included in the point source module now does the work of creating the point source grid data files. When working with "theoretical" strategies that add and delete point sources to check for sensitivity in the photochemical model, this new element will make that job easier.

In order to locate point sources in the modeling grid, coordinates for the facility and stack location are submitted as part of the point source data. During OTAG, it was assumed that everyone would submit data using UTM coordinates. Data now comes from many sources and control over its content is not possible. In response to this predicament a new variable to track coordinate type was added to the point source facility file. Processors written into the point source module automatically convert between UTM, latitude/longitude and Lambert Conformal Conical projection systems based on the coordinate type indicated and translate all the data to UTM.

The subgrid.sas processor is capable of creating spatial surrogate files for mobile and area sources based on nationally available pre-gridded surrogate files. For example, this processor will create surrogates for any sub domain nested within the Eastern Unified Domain. This is important because for most modelers this will eliminate the need to use ARCINFO to build these spatial surrogates.

While additions and updates greatly improve the way that EMS2001 processes data, modeling results will only be as good as the data itself. Because of this, EMS2001 includes new as well as improved processors written to assist the inventory developer to QA (quality assure) the data submitted for modeling. For point and area sources, QA occurs almost immediately in the modeling process. There is a separate QA processor for each of the five point source input files (facility, stack, device, process, emission) and each of the two area source input files (area, areatprl). Any reports generated by this QA process were written to one large point or area source file respectively. Moreover, there was no analysis to make sure that the reports were consistent and useful to the inventory developer. These seven QA processors, which are run automatically during data formatting, have now been rewritten to include more valuable reports and to provide better output. The "Errors" and "Warnings" from each QA processor are classified and written to four tier report files. Each tier represents a different level of concern that the input data may not produce the results expected by the inventory developer. The four tiers are defined as follows:

Tier 1 - Directly will cause lost emissions
Tier 2 - Indirectly could cause lost emissions or reporting problems.
Tier 3 - Will result in a system default being applied
Tier 4 - Problems that are not critical; correct only if you have the time.

This structure allows the inventory developer to better use their time addressing problems that will directly affect modeling results. Many of the reports have been summarized so that only the information needed to attend to the issue is provided. This has eliminated hundreds of pages of reports previously generated by EMS95 making it easier to glean pertinent information. As part of this effort, a full analysis of the QA checks was performed and the processors were altered where needed to improve consistency and thoroughness.

These reports are generated at the beginning of the modeling process. EMS2001 also has report generators that are run on the final emissions data sets to assist in both QA and presentation of the inventory data to interested parties. These reports for point, area, motor vehicles, and biogenics provide both text reports and graphics summarizing the data in a variety of ways. These processors are written to report ozone precursors (VOC, NOX, and CO). Versions for other pollutants can be easily derived. To improve flexibility, the processors determine the states in the modeling domain on the fly, so that different domains can be evaluated. In addition the area source report generator was modified to improve run time by 60%.

But what do the monitors say? The PAMS (Photochemical Assessment Monitoring Stations) tool can help to answer that question. Ambient data from PAMS monitoring sites can be used to evaluate the modeling emissions inventory. The tool is designed for use in situations where the local ozone precursor concentrations are dominated by local ozone precursor emissions. This requires high local emissions and analysis during the time of day when meteorology and chemical reactions are less significant. Analysis using this tool can help modelers understand if the right amount of pollutants is present in the inventory and if the emissions are being properly speciated. Presentations on the use of the PAMS tool in the Lake Michigan region have been given at previous conferences and is available at the LADCO website www.ladco.org. This tool has now been integrated into EMS2001 so it is available to all EMS2001 users.

Regardless of features, EMS95 was cost prohibitive to use for some. It runs on a UNIX platform only. Most users spent thousands of dollars on high-end workstations. It required licenses for ARCINFO and SAS that for a UNIX workstation were several thousand dollars each year. Since 1995, the power of personal computers has increased exponentially. EMS2001NT takes advantage of that power, providing a full version of EMS2001 that can run on an NT workstation at a fraction of the cost and just as fast as its UNIX cousin.

EMS2001NT is written primarily in SAS with a structure familiar to those who have used EMS95 in UNIX. Fortran is used to compile the Mobile5b executable used to calculate motor vehicle emissions data sets. Purchasing these programs and maintaining the license for an NT machine is much less expensive than for a UNIX workstation: up to 10 times less expensive. Moreover, a very powerful NT workstation can be purchased for a fraction of the cost of a high-end UNIX workstation. These factors combined make EMS2001NT much more affordable.

EMS2001NT was written specifically to look and feel very similar to the UNIX version so that those already familiar with EMS95 would have no problem switching between platforms. The entire core model of EMS2001NT fits on one CD with the EMS tree structure zipped. To install EMS2001NT the tree structure must be extracted (unzipped.) This will load the core of the model on the machine. The core programs and supporting files take approximately 1 Gigabyte of space, but you will want to install the model on a drive with at least 20-30 Gigabytes of available space to allow for the reports and data sets that will be written when running the model. The SAS data sets supplied in the core installation assume that you are running SAS with a V7 or V8 engine (data set extension of ‘sas7bdat’.) If this is not the case, you will need to

import the SAS data export files supplied so that they are compatible with your set up. Examples of import file programs are also supplied on the EMS2001NT CD.

You will also need to set up the "SCENARIO" file structure in the following manner:

%SCEN_HOME%\%EMS_PROJECT%\scenario\%EMS_SCENARIO%\sas

Where:

- %SCEN_HOME% is the drive and path to the location where scenarios are stored,
- %EMS_PROJECT% is the name of the current project (e.g. gridm, strat1, otag),
- %EMS_SCENARIO% is the name of the current scenario (typically the days of the strategy: 910713, 910714, etc.).

Now that the file structure and EMS2001NT code and support files have been installed, it is time to set up the system environment to run the model. EMS95 relied on environment variables and aliases added to a users ".cshrc" file that would be active when they logged on to assist in initial navigation through the model. Users must also log onto an NT machine but only path names and environment variables can be set. The system environment parameters that must be set for each user are:

1. Put SAS directory in the PATH name
2. PUT Fortran directory in the PATH name
3. Put <drive:>\ems\ems2001NT\batchsys in the PATH name, where <drive:> refers to the drive on which EMS2001NT has been installed.
4. Set environment variables SASHOME, SCEN_HOME, and EMS_HOME where:
    a) SASHOME is the drive and path to the location of the "SAS" program directory.
    b) SCEN_HOME is the drive and path to the location where scenarios are stored.
    c) EMS_HOME is the drive and path to the location of the "EMS2001NT" directory

Application and examples of these settings can be found in the *User's Guide to EMS2001NT* included with the model.

EMS2001NT is run from a DOS shell. This was done primarily so that SAS could be run in the background using command lines in batch files but it also makes using the NT version much more like the UNIX version. In order to allow the user to use familiar aliases, such as 'gobin', a system of batch files was created and stored in the %EMS_HOME%\batchsys directory. One of these batch files, modelme.bat, sets up the basic parameters needed to begin moving around EMS2001NT. It defines specific locations of directories and will need to be edited to reflect the set up on your system. It should be the first thing run after opening the DOS shell. More information about running and creating batch files in the DOS environment is in the *User's Guide to EMS2001NT* included with the model.

A listing of the *.bat files in the %BATCHSYS% directory shows most of the aliases familiar to the EMS95 user. Most of these files run only the lines within the batch file itself. There are some exceptions. EMS95 users will be familiar with "locme", "editme", "viewme", and "insime". These batch programs in the UNIX environment allowed the user to set the location of the %EMS_LOC% directory, run "proc edit" on a user specified SAS data set, run "proc fsview" on a user specified SAS data set, and run "proc insight" on a user specified SAS data set, respectively. These important tools could not be written using batch files because of differences in the way NT batch files can receive data from the user. Therefore, executables for each of these tools were written using Visual Basic® 5.0. The batch files for each tool creates the text files needed by the executable, runs the executable, and then deletes the text files created. The executables must be run using the batch files; otherwise, an error message will be given. Each of these tools will open a Windows dialog box with appropriate options for choosing a directory

or SAS data set. Just like in UNIX, a log file will be created in the directory from which you called the batch file.

The last tool included in the %BATCHSYS% directory is a DOS UNIX / UNIX DOS converter: convertcrlf.exe. In addition to the possibility of sharing text files between users of EMS2001 and EMS2001NT, many times input data files received are made on a different platform than the one you are using to run EMS2001. The differences in return characters used to mark the end of a line will not function in EMS2001/NT. You can use any format converter you wish. This particular one is included in the EMS2001NT release because it is easy to install (you need to run the executable once to load the software), it can convert whole directories of files at a time, it provides flexibility for file management, and it is shareware (it can be distributed for free but not sold.)

There is one more executable file you will need to check before running the motor vehicle module: mobile5b.exe. Depending on your system, you may need to recompile this executable using Fortran. The Fortran files needed to create this executable are included with the model. Further instructions are available in the *User's Guide to EMS2001NT.*

At this point EMS2001NT runs in a very similar manner to the UNIX version. Batch files can be written to run all the processors for a module. Differences between UNIX and DOS batch files are described in the *User's Guide to EMS2001NT* to help the user more familiar with UNIX write batch files in DOS. The SAS programs will run in the background, allowing the user to perform other tasks while the machine is running EMS2001NT.

EMS2001NT has never been tested to check for "minimum" system requirements. The Wisconsin Department of Natural Resources (WDNR) is currently running EMS2001NT on a Dell 610 Workstation with dual 550 Mhz processors, 1 Gigabyte of RAM, and 72 Gigabytes of hard drive space (4-18 Gigabyte drives.) The workstation runs Windows NT 4.0, SAS 7.0, and Lahey Fortran95. This set up has been benchmarked against a SUN Sparc Ultra and has produced identical results. Better yet, it has had comparable run times for all modules.

Like most models, the improvements do not stop here. Planned updates to EMS2001/NT include adding the ability to read the National Emissions Inventory Format (NIF), adding CMAQ compatible output processors, increasing the processing speed of the motor vehicle module, incorporating the new Non-Road Model and MOBILE6 (when they are available.)

EMS2001 provides many updates from its predecessor, EMS95. The inclusion of BIOME3, the SPECIATE 3.0 profiles, and the PAMS analysis tool assure that EMS2001 uses the latest information available for estimating emissions. But the true strength of EMS2001 lies in the flexibility of its open code structure. Processors to convert between projection systems or improve QA reporting can be added with relative ease to respond to the ever-changing needs and imagination of the end user.