

GLOSSARY

Accumulation mode. Aerosol particles with diameters nominally between 0.1 and 1.0 micrometers. The term is based upon the concept that particles in this size range accumulate in the atmosphere from the growth of much smaller particles. (See also “Aitken mode.”) The growth may be from new mass condensing on the smaller particles, or from the coagulation of the smaller particles.

Adaptive grid. A grid structure that varies during model execution according to the value(s) of some model parameter(s). For example, in a photochemistry model, grid resolution may automatically increase in areas of high spatial gradients of NO_x . This allows more accurate determination of plume-to-background concentration ratios, which greatly influence photochemical ozone production. In a meteorological model, an adaptive grid may automatically increase the grid resolution in an area of modeling where there is a large atmospheric pressure change across a grid cell.

Air quality modeling system. A computational system environment that combines a set of physical and chemical models that describe atmospheric processes, which are important to trace-gas and particulate matter distributions in the atmosphere. These systems typically include meteorological models, emissions models, chemistry-transport models, and the analysis and visualization tools necessary for supporting decisions related to air quality.

Aitken mode. Aerosol particles with diameters nominally between 0.01 and 0.1 micrometers (μm). Such particles are formed in the atmosphere by nucleation of gas-phase precursors, or by direct emissions from sources. The most common source is combustion (e.g., diesel soot).

Arakawa-B. Horizontal grid staggering system described by Arakawa and Lamb (1977) and used by MM5. Mass variables and momentum variables are defined on separate horizontal grids of spacing equal to Δx . The two grids are offset by $0.5 \Delta x$ in the north-south and east-west directions.

Arakawa-C. Horizontal grid staggering system described by Arakawa and Lamb (1977) and used by WRF-ARW and CCTM. Mass variables and each horizontal component of the momentum are defined using separate horizontal grids of spacing equal to Δx .

ArcInfo. A high-end geographical information system (GIS) with capabilities for the automation, modification, management, analysis, and display of geographical information.

Automatic quality control (QC). QC correction that is accomplished automatically without user intervention.

Bookmark. In on-line help, a bookmark marks an entry of a help document so it can be quickly accessed later. A list of bookmarks appears in the Bookmarks Menu on the help browser window and also in the bookmark's window. Each item on the list is a hypertext link to a marked entry.

Chemistry-transport model (CTM). A model that simulates various physical and chemical processes that are important for understanding atmospheric trace-gas and particles distributions. The processes include atmospheric transport; vertical mixing; atmospheric chemistry in the gas

phase, in the aqueous phase, and in aerosols; cloud mixing and precipitation scavenging; and surface removal processes. Generally, a chemistry-transport model relies on a meteorological model for the description of atmospheric states and motions, and depends on an emissions model for the description of anthropogenic and biogenic emissions that are injected into the atmosphere.

Class. A collection of software modules associated with a science process.

Conforming datasets. Conforming datasets are in I/O API format. Most programs (models, visualization and analysis routines) in the CMAQ system are able to read and write conforming datasets. This capability eliminates the need for data conversion, even within a distributed computing environment.

Conforming programs. Conforming programs generally use the I/O API library routines for reading and writing data. Key data structures are defined by globally shared information. You define this critical data structure information once, and it is automatically made available for use in conforming code throughout the system. This globally shared information is permanently stored as objects in an object data base. In CMAQ, these objects are a stored collection of related information, such as grid dimensions and resolution, coordinate system definitions, chemical mechanism species, and reactions.

'''Cross point. '''Grid point where all scalars are defined.

Daemon. A process that runs in the background independently and performs a function. An example is a printer daemon that controls the job queue for the printer.

Decision support system. An automated system that provides

information for decision making.

Domain. The domain for a CMAQ calculation is the horizontal geographic area (e.g., the eastern United States) that is of interest. CMAQ calculations are performed for physical and chemical processes occurring in the atmosphere above this horizontal geographic area.

'''Dot point. '''Grid point where wind components are defined

Emission model. This type of model calculates the emission rates of trace gas and particulate matter into the atmosphere under ambient meteorological conditions and socioeconomic activities. Emissions from both natural and manmade sources are calculated. One example of an emissions model is SMOKE.

Environmental modeling system. A set of computational models that mathematically represents a simplified version of real-world phenomena, along with the necessary mechanisms for managing the processing, the data produced by it, and the analysis and visualization tools necessary for making related decisions. For example, this could be the creation and behavior of environmental pollutants as a function of weather and chemical interactions. Researchers use these “scaled-down” versions of the world to perform experiments that would be too dangerous, too costly, or simply impossible in the real world.

Eulerian. Fluid motion specification where one concentrates on what happens at a spatial point, 'x, 'so that independent variables are taken as a function of time at that point.

Evasion. Loss of material from the land surface (e.g. the upward flux of mercury from the land surface).

Exploratory models. Test-bed models for developing scientific

algorithms that are not thoroughly reviewed and evaluated. (See also “operational models” and “screening models.”)

Fortran. Formula translator (computer programming language).

Framework. A system of mechanisms to manage the scheduling and execution of computational models, the data produced by them, the analysis and visualization tools necessary for understanding their results for decision making, and the interfaces to all these capabilities.

Generalized coordinate system. A scheme for constructing coordinate systems that allows for choices of horizontal map projection type (e.g., latitude/longitude, Lambert, Mercator, polar stereographic, Universal Transverse Mercator, or matrix) and projection parameters, as well as for choices of various vertical coordinate types (e.g., pressure coordinate, height above ground, height above sea level, sigma-p hydrostatic, sigma-p nonhydrostatic, and sigma-z). The advantage of a generalized coordinate system is that a CMAQ model can adapt to a variety of different possibilities in horizontal and vertical parameters.

Generic grid system. A scheme for constructing grid systems that allows for choices of origin, orientation, and spatial resolution. It allows a model to be expressed in a grid system that is optimal for representing the governing equations. For a regular, rectangular grid system, mapping gridded data to the earth’s surface can be achieved by defining the number of rows and columns, cell size, and location and extent. For an irregular grid system, grid intersections (nodes) are described by coordinates from a reference position.

Geocoded. An entity with an identifier associated with geographic boundaries (e.g., state or county code).

Geographic information system (GIS). A computer-based system for

managing, analyzing, manipulating, and displaying geographic information in both numeric and graphical ways.

Geospatial. Refers to the spatial extent of a geographic boundary.

Grid cell. The smallest subdivision of a grid.

Grid size. Length of shortest side of a rectangular grid cell.

Grid. A network of conceptual locations at which numerical calculations are performed. This network extends horizontally to cover the domain of interest, and vertically through the troposphere.

Growth factor. An adjustment factor used to estimate the growth in a source's activity level between the inventory base year and a projected year. Valid values are 0.00 - 99.99.

Heterogeneous, distributed computing environment. A

heterogeneous computing environment consists of multiple machines of different types (e.g., supercomputers, graphics workstations, and workstation file servers). A distributed computing environment permits the execution of a large computational task to be shared across multiple machines linked together by a network. Thus, a heterogeneous, distributed computing environment consists of many different kinds of machines networked together.

Hydrostatic. Used to indicate to a model that it is to assume that the atmosphere is in hydrostatic equilibrium (i.e., surfaces of constant pressure and constant mass coincide and are horizontal throughout). Complete balance exists between the force of gravity and the pressure force.

Hypertext link. A hypertext link is a specially designated word or phrase displayed in text that has been saved in Hypertext Markup Language

(HTML) format. A hypertext link provides nonsequential access to other entries in a document set. In CMAQ, the Help facility is done using hypertext. Hypertext linking is done to all help entries.

Input/Output Applications Programming Interface (I/O API). A software library that reads and writes files. It uses an internal data format that is machine independent and that conforms to the widely used University Corporation for Atmospheric Research Network Common Data Format (netCDF). The I/O API files contain self-describing headers with complete information that is necessary to use and interpret the data contained in the file. The I/O API format is slightly more restrictive than the standard netCDF format regarding how the header information must be written. The I/O API library provides a variety of data structure types and a set of reusable access routines that offer selective direct access to the data in terms that are meaningful to the environmental modeler. Supported data types include gridded, boundary, vertical profile, grid nest, time series, and event-driven. For additional information on the I/O API, see Chapter 4.

'''Internal. '''With respect to CMAQ data, internal means that the data are available within the software; the user does not have to provide them. Examples include look-up tables, ranges, and lists of state/county names.

Inventory Data Analyzer (IDA). Program used for input and quality control checks of emission inventories.

Inventory. With respect to an emission processing system, inventory refers to a file or a database containing emission data for a specific set of pollutants for a specific time period (typically for the entirety of a specific year) for an area (country, states, counties).

Irix. Operating system for SGI computers.

Keyword. A keyword is a word or phrase, up to 40 characters long, that can be used to locate an entity in help text or a CMAQ object (e.g., dataset, program, study, or model) using a Find screen.

Layer collapsing. A procedure in which the layer structure prepared by a meteorological model is modified by reducing the number of layers.

"This is not recommended."

Linux. An open-source, UNIX-like operating system. It is available and redistributable under the terms of the GNU Public License.

Makefile. A Makefile is a list of UNIX commands that perform activities such as mounting files and compiling source code.

Massively parallel processing. Computer processing employing multiple CPUs to execute multiple computational tasks simultaneously. Massively parallel systems employ a large number of CPUs simultaneously on the same task. In contrast, conventional computer design uses a single CPU to perform computational tasks in a strictly linear, sequential order.

Mesoscale. Medium scale. In meteorology, mesoscale denotes medium horizontal and vertical spatial scale. The horizontal scale extends to several hundred kilometers. The vertical scale extends from tens of meters to the depth of the troposphere.

Metadata. Information about data and about the processes that produced them. In particular, information about the data's structure, internal characteristics and history, and location within a data storage environment, as well as information derived from the data.

Meteorological model. This type of model provides descriptions of atmospheric motions, momentum, moisture, heat fluxes, turbulence

characteristics, clouds and precipitation, and atmospheric radiative characteristics. Most meteorological models currently in use for air quality modeling were originally developed for the prediction of weather. CMAQ models require information from a meteorological model that is designed to address specific issues relevant to air quality modeling, such as planetary boundary layer information, cloud distribution and mixing characteristics, precipitation, and surface fluxes.

'''Mie scattering. '''A generalized particulate light-scattering mechanism that follows from the laws of electromagnetism applied to particulate matter.

Mixed-media. Simultaneously involving more than one environmental pollutant medium, such as air and water.

Model developer. Those scientists and software developers who construct and study theories to explain physical and chemical processes, and use computer models to study their theories.

Model users. Research, production, and quality assurance personnel (i.e., applied scientists and engineers) who generate valid input for the modeling systems, run the modeling systems, maintain audit trails, analyze input and output for correctness, and produce analyses of the modeling results.

Model. A representation of a planned or existing object or condition.

Modeling structure. A design specification that provides the paradigm of operation and the interface specifications for the modules used to construct a particular family of models. In a CMAQ model, for example, the paradigm is that modules act as operators upon a shared concentration field, and four types of interfaces (call interfaces, INCLUDE-file interfaces, I/O interfaces, and UNIX-environment

interfaces) must be specified.

Modeling system. A set of computational models and associated data processors that together provide for the simulation of processes of interest.

CMAQ components. The various subsystems within the CMAQ framework. Each component is represented by its own icon. The available components are Dataset Manager, Model Builder, Program Manager, Science Manager, Strategy Manager, Study Planner, and Tools Manager.

CMAQ. The third-generation air quality modeling system. It is a flexible system that addresses multiple air quality issues, such as regional- and urban-scale oxidant and acid deposition.

Module. A subset that is part of a larger program (such as a meteorological model, an emissions model, or CMAQ). In a modular program, all modules of a particular type (e.g., those that compute dry deposition) are interchangeable, allowing you to replace one module with another to determine, for example, how each module affects modeling results. Examples of modules include science modules and analysis and visualization modules.

Monotonic. A quality of strictly increasing or decreasing within some interval.

Multilevel nesting. Multilevel nesting refers to employing nested grids within nested grids, possibly several levels deep.

National Emissions Inventory. A database at EPA containing the information about sources that emit criteria air pollutants and their precursors, and hazardous air pollutants.

Nested grids. Nesting refers to fitting a finer-resolution grid over part of a coarser-resolution grid. The finer-resolution grid receives information (such as boundary conditions) from the coarser-grid simulation.

Nonconforming datasets. Nonconforming datasets are ones that are not in I/O API format. They can be used in the CMAQ framework by programs that are specifically designed to read those datasets. When nonconforming datasets and programs are used, however, you must know how to match programs and datasets, and which data formats and programs are transportable to different machine architectures. Those considerations are automatically managed by the Models-3 framework for those who use conforming datasets and conforming programs.

Nonhydrostatic. Used to indicate that the model does not assume that the atmosphere is in hydrostatic equilibrium. Air is not assumed to have only horizontal motion relative to the earth.

Open-source software. Open-source software began as a marketing campaign for free software. OSS can be defined as computer software for which the human-readable source code is made available under a copyright license (or arrangement such as the public domain) that meets the “open source” definition. This permits users to use, change, and improve the software, and to redistribute it in modified or unmodified form. It is very often developed in a public, collaborative manner. Open-source software is the most prominent example of open-source development and often compared to user-generated content.

Operational models. These models offer fully functional modeling of relevant science processes, such as atmospheric, emissions, and chemical transport processes. They represent the current state-of-the-art that has undergone significant quality assurance review, peer review, and evaluation. The turnaround time for these models is generally much

longer than for screening models but short enough to allow episodic studies.

Parameterize. To create an algorithm that describes the average large-scale behavior of a physical phenomenon, rather than describing the subgrid-scale behavior in terms of the underlying physics and chemistry. For example, a parameterized cloud algorithm might describe average cloud behavior over 80-km-square cells, although the individual clouds are much smaller than 80 km across.

Planetary boundary layer. The portion of the atmosphere adjacent to the earth's surface. This layer generally ranges from 100 m to 2 km in height, and is significantly influenced by surface effects (e.g., heating, friction). These effects can cause the layer to be well-mixed, which affects the distribution of pollutants in the air. (See also "troposphere.")

Popup window. A popup window is a special window for displaying an on-line help entry. The window opens when you select a specially designated hypertext link. Pop-up windows are useful for quickly displaying short, concise units of information. You cannot jump or step to other entries from a pop-up window.

Prepare. Read and process a file or a set of data.

Process analysis. Traces the source(s) of a chemical species within a simulation. One example of process analysis is determining whether a species concentration originated within the cell in which it is found or instead within some other cell(s). Another example is determining what chemical constituents were involved in creating a species produced by reaction (rather than transported from somewhere else).

Process. Read a file or a set of data, perform the desired functionality (quality control, reformatting, algebraic operations, etc.) and submit the

processed data to the next set of actions.

Quality control (QC). The act of reading data (inventories, files) and checking for correctness, completeness, and consistency. QC may involve automatic correction, substitution, or the filling of missing data. All QC processes are followed by QC reports.

Register data. When you register data, you are making something that already exists (e.g., a file) known to the system.

Rule effectiveness percent. An adjustment to projected estimated emissions data to account for emissions underestimates due to compliance failures and the inability of most inventory techniques to include these failures in an emission estimate. The adjustment accounts for known underestimates due to noncompliance with existing rules, control equipment downtime, or operational problems and process upsets. Valid values: 0 to 100.

Rule penetration percent. An adjustment to projected estimated emissions data to account for emissions underestimates due to failure to implement rules throughout the area of intent. Valid values: 0 to 100.

Scalable. In the context of parallel computer architectures and algorithms, a parallel architecture or algorithm is termed scalable if its performance behaves linearly in terms of the number of processors employed. For example, doubling the number of processors does not cause new communications bottlenecks to appear but doubles the performance achieved.

Scale flexibility. The modeling system's ability to accurately simulate physical and chemical processes that describe atmospheric pollutants at different spatial and temporal scales. A modeling system with scalable algorithms for physical and chemical processes and with a generic grid

system has this quality.

Science module. A component that is part of a modeling program (such as a meteorological model, an emissions model, or CMAQ) and that computes data for a discrete category of environmental phenomena (e.g., dry deposition, photochemistry, vertical advection).

Screening models. These models have simplified science processes designed for quick assessment studies. These models are employed when users are willing to sacrifice some accuracy for faster turnaround time. A typical screening study involves making a large number of runs in order to identify episodes or situations that warrant more detailed modeling.

Source Classification Code (SCC). The SCC system is used for classifying emissions sources at the level of individual processes (e.g., automobiles, boilers, solvent sources) within an emissions data inventory.

Source. With respect to air pollution, a point, area, or mobile source that produces and/or emits air pollutants.

Speciation. In CMAQ, speciation refers to using one of the chemical mechanisms available with CMAQ to disaggregate a chemical substance (pollutant) into simpler compounds.

Species. Typically, a chemical substance or group of related substances whose behavior is modeled during environmental simulation modeling.

Sub-grid-scale process. Physical process that occurs on a scale smaller than the grid resolution of the modeling system, such as point-source plumes and convective clouds. Since the scale is smaller than the grid resolution, these processes must be estimated or parameterized.

Summary report. Generally refers to an automatic, short report generated after the execution of a process.

Surface fluxes. The exchange of material, energy, or momentum between the surface of the earth and the atmosphere.

Time step. A time step is a fixed unit of time. A model may have one or more internal time steps for various processors. In the CMAQ framework, a time step is used to indicate the length of time between output of variables from the model or a process within the model. Another term might be “output time interval.”

Troposphere. The troposphere is the lowest portion of Earth's atmosphere. It contains approximately 75% of the atmosphere's mass and almost all of its water vapor and aerosols. The average depth of the troposphere is about 11 km (7 miles) in the middle latitudes. It is deeper in the tropical regions (up to 20 km [12 miles]) and shallower near the poles (about 7 km [4 miles] in summer, indistinct in winter). Within the troposphere, temperature decreases with altitude. The lowest part of the troposphere, where friction with the Earth's surface influences air flow, is the planetary boundary layer (PBL). This layer is typically a few hundred meters to 2 km (1.2 miles) deep, depending on the landform and time of day. The border between the troposphere and stratosphere is called the tropopause. Above this layer is a temperature inversion—that is, in the stratosphere temperature increases with altitude.

Visualization. An important aspect of scientific computing that provides a method for presenting easily understandable data quickly and compactly in the form of charts and graphs.

[1] Future efforts toward fourth-generation systems will extend linkages and process feedback to include air, water, land, and biota to provide a more holistic approach to

simulating transport and fate of chemicals and nutrients throughout an ecosystem

[2] The CVS "modules" file has no intrinsic relationship with the CMAQ classes/module design implementation.

[<< Previous Chapter](#) - [Home](#)