United States
Environmental Protection
Agency

Office of Research and
Development
Washington, D.C. 20460

EPA-600/R-99/065
July 1999

# Using MM5 Version 2 with Models-3/CMAQ

# A User's Guide and Tutorial

PROPERTY OF
DIVISION
OF
METEOROLOGY

EPA-600/R-99/065
July 1999

# Using MM5 Version 2 with Models-3/CMAQ

# A User's Guide and Tutorial

By

Tanya L. Otte*

Atmospheric Modeling Division
National Exposure Research Laboratory
U. S. Environmental Protection Agency
Research Triangle Park, NC 27711

* On assignment from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

# DISCLAIMER

The information in this document has been funded wholly or in part by the United States Environmental Protection Agency. It has been subjected to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement nor recommendation for use.

# CONTENTS


Disclaimer ..... ii
Notice ..... v
Figures ..... vi
Tables ..... vii
Acknowledgments ..... viii
Acronyms ..... ix

1. USING MM5 WITH MODELS-3/CMAQ ..... 1
1.1 Introduction ..... 1
1.2 Why Use a Meteorological Model with Models-3/CMAQ ..... 1
1.3 What is MM5? ..... 2
1.4 Requirements to Run MM5 ..... 2
1.5 Meteorology Model Pre-Processing ..... 3
1.6 Defining the Simulation Domain (*TERRAIN*) ..... 4
1.7 Processing the Meteorological Background Fields (*DATAGRID*) ..... 6
1.8 Objective Analysis (*RAWINS*) ..... 9
1.9 Setting the Initial and Boundary Conditions (*INTERP*) ..... 10
1.10 The Meteorology Model (MM5) ..... 12
1.11 Meteorology Model Visualization ..... 14
1.12 Meteorology Model Post-Processing ..... 15
1.13 Changes to the MM5 System's Software for Models-3/CMAQ ..... 15

2. MODELS-3/CMAQ VERSION 3 TUTORIAL — MM5 ..... 17
2.1 Background ..... 17
2.2 Preparing to Run MM5 ..... 18
2.3 Case Overview ..... 19
2.4 Working with the Scripts ..... 19
2.5 Running *TERRAIN* ..... 21
2.6 Running *DATAGRID* ..... 22
2.7 Running *RAWINS* ..... 23
2.8 Running Front-End *INTERP* ..... 25
2.9 Running MM5 ..... 26
2.10 Running Back-End One-Way Nest *INTERP* ..... 28
2.11 Comparison on Various Platforms ..... 29

## CONTENTS (continued)

Appendix A: References ..... 31
Appendix B: File Size ..... 33
Appendix C: MM5 Version 2 Output Variables ..... 35
Appendix D: Sample Program to Read MM5 Version 2 Output ..... 37
Appendix E: EPA-Initiated Changes to MM5 Software ..... 41

# NOTICE

MM5 was developed in cooperation with Penn State and the University Corporation for Atmospheric Research (UCAR). Penn State and UCAR make no proprietary claims, either statutory or otherwise, to any version and release of MM5, and they consider MM5 to be in the public domain for use by any person or entity without any fee or charge. Penn State and UCAR shall be credited in any publications that result from the use of MM5. The names "Penn State" and "UCAR" shall not be used or referenced in any advertising or publicity that endorses or promotes any products or commercial entity associated with or using MM5, or any derivative works thereof, without the written authorization of UCAR and Penn State.

MM5 is provided to the public by Penn State and UCAR on an "as is" basis, and any warranties, either express or implied, including but not limited to implied warranties of non-infringement, originality, merchantability, and fitness for a particular purpose, are disclaimed. Neither UCAR nor Penn State will be obligated to provide the user with any support, consulting, training, or assistance of any kind regarding the use, operation, and performance of MM5, nor will they be obligated to provide the user with any updates, revisions, new versions, error corrections, or "bug" fixes. In no event will UCAR and Penn State be liable for any damages, whatsoever, whether direct, indirect, consequential, or special, which may result from an action in contract, negligence, or other claim that arises out of or about the access, use, or performance of MM5, including infringement actions.

Data that are obtained from NCAR (e.g., global analyses and observations) are prepared and maintained by the Data Support Section, Scientific Computing Division, NCAR. Some of the data that are released by the SCD cannot be freely distributed. Please contact the SCD before disseminating data that either originated from NCAR or was obtained directly from NCAR.

NCAR is operated by UCAR and is sponsored by the National Science Foundation. Any published research that uses data from the NCAR archive shall credit the maintainers of the archive. In addition, the original source(s) of data shall be acknowledged.

# FIGURES


1 Overview of the MM5 system and its connection to Models-3/CMAQ.....3
2 *TERRAIN* inputs and outputs.....6
3 *DATAGRID* inputs and outputs.....8
4 *RAWINS* inputs and outputs..... 10
5 Front-end *INTERP* inputs and outputs..... 12
6 Back-end one-way nest *INTERP* inputs and outputs..... 12
7 MM5 inputs and outputs..... 14
8 MM5 domains for the Models-3/CMAQ Version 3 Tutorial..... 23

# TABLES


1 Path variables in MM5 scripts ..... 20
2 Variables common to "go" scripts ..... 20
3 Variables specific to "terrain.go" ..... 21
4 Variables specific to "datagrid.go" ..... 22
5 Variables specific to "rawins.go" ..... 24
6 Variables specific to "interp.go" ..... 25
7 Variables specific to "mm5.go" ..... 26

# ACKNOWLEDGMENTS

This work was supported under NOAA Interagency Agreement DW13937252 with the EPA. Dr. Sharon LeDuc (NOAA), Dr. Jonathan Pleim (NOAA), Dr. Daewon Byun (NOAA), Thomas Szymkiewicz (EPA), Brian Orndorff (NOAA), and Dr. Carey Jang (EPA) provided comments and suggestions that served to improve this document. In addition, Lara Joyce (DynTel Corp.) performed the initial testing and evaluation of the MM5 workstation system on the EPA Sun ULTRA network, and provided the comparisons between Cray and Sun output.

# ACRONYMS

| | |
|---|---|
| CCTM | CMAQ Chemical Transport Model |
| CMAQ | Community Multi-Scale Air Quality |
| ECMWF | European Centre for Medium-Range Weather Forecasting |
| EPA | Environmental Protection Agency |
| FDDA | four-dimensional data assimilation |
| GMT | Greenwich Mean Time |
| I/O API | Input/Output Application Programming Interface |
| MCIP | Meteorology-Chemistry Interface Processor |
| MM5 | Fifth-Generation PSU/NCAR Mesoscale Model |
| NAAQS | National Ambient Air Quality Standard |
| NARSTO | North American Research Strategy for Tropospheric Ozone |
| NCAR | National Center for Atmospheric Research |
| NCEP | National Centers for Environmental Prediction |
| PBL | planetary boundary layer |
| PSU | The Pennsylvania State University |
| SST | sea-surface temperature |
| UCAR | University Corporation for Atmospheric Research |
| UTC | Universal Time Constant |

# Chapter 1
# USING MM5 WITH MODELS-3/CMAQ

## 1.1 Introduction

Meteorological data are important in many of the processes simulated in the Community Multi-Scale Air Quality (CMAQ) model and the Models-3 framework (Models-3/CMAQ). The first meteorology model that has been selected and evaluated with Models-3/CMAQ is the Fifth-Generation Pennsylvania State University/National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5). All software in the Penn State/NCAR mesoscale modeling system has been dedicated to the public domain and is presently used for both research and operational purposes at many organizations worldwide. Improvements to MM5 within the meteorological community are ongoing, and these enhancements may be evaluated for their applicability to air quality modeling. Other meteorological models are also planned for use with Models-3/CMAQ, but that capability has not been publicly released at this time.

This document is a basic overview of MM5 and its relationship with Models-3/CMAQ. Some general information is provided to help the user get started with MM5 in preparation for Models-3/CMAQ. However, this document is not a substitute for NCAR's MM5 Tutorial Course (offered semi-annually in Boulder, Colorado, for a nominal charge), and it is meant to supplement NCAR's official documentation on MM5. Users should expect to seek additional information from NCAR's official documentation in conjunction with this document in order to successfully run MM5. This document is not endorsed by and has not been reviewed by NCAR. (Please read the Notice at the beginning of this document for credits and disclaimers for user of MM5.) NCAR officially maintains MM5 and its supporting software, and NCAR should be the primary contact for general questions about using MM5. Questions specifically regarding the use of MM5 with Models-3/CMAQ should be directed to the EPA. The application of MM5 with Models-3/CMAQ and a test case ("Models-3/CMAQ Version 3 Tutorial") are described below. Users should thoroughly read the background information on MM5 before proceeding to the tutorial.

## 1.2 Why Use a Meteorological Model with Models-3/CMAQ

In order to properly run an air-quality modeling simulation, the chemistry model needs the best possible characterization of the atmosphere. With some first- and second-generation air-quality models, the meteorology was provided through a series of diagnostic analyses that typically included 12-hourly upper-air analyses and hourly surface analyses. However, more local (higher-resolution) effects must be captured by air quality modeling efforts. The observation network (which provides the basis for the diagnostic analyses) cannot resolve mesobeta-scale (on the order of 20 kilometers and several hours) features, particularly in the planetary boundary layer. One way to "fill the gap" is to use a meteorological model as a "dynamic analysis" tool. A meteorological model can provide a dynamically consistent evolution of the atmosphere at more frequent time intervals and higher spatial resolutions than analyses of observations.

Meteorological models can add scientific credibility to the air quality simulation through the use of advanced theory and techniques. It must be acknowledged that better science does not always lead to a better solution; there will be times when diagnostic analyses are more realistic representations

of the atmosphere than the meteorological model. However, using a meteorological model to provide input for the chemistry model has been shown to add skill, particularly for regional and urban scales.

## 1.3 What is MM5?

MM5 is the Fifth-Generation Penn State/NCAR mesoscale model. It is a complex, state-of-the-science meteorological model. MM5 evolved from the model used by Anthes in the early 1970s, described by Anthes and Warner (1978). MM5 is used both nationally and internationally for operational (forecasting) applications and for research. In addition, MM5 is widely used for air quality applications.

The improvements in MM5 over the previous version (MM4) include the option for non-hydrostatic physics, as well as more sophisticated explicit moisture, boundary layer processes, radiation, and convective parameterization, among other improvements. Dudhia (1993) documented the major changes from MM4 to MM5. Starting with MM5 Version 2 (MM5v2), the model and its supporting software have been restructured to run on various hardware platforms in addition to the Cray supercomputer, with emphasis placed on workstation-based MM5 simulations. Starting with MM5v2 Release 8 (MM5v2.8), the simulation model is also formally supported on parallel hardware architectures.

MM5 can be obtained electronically via anonymous-ftp from NCAR's site, *ftp://ftp.ucar.edu*, in the subdirectory */mesouser*. In addition, the EPA allows free access to versions of MM5 that have been tested with Models-3/CMAQ from *ftp://www.epa.gov/pub/asmdnerl/stand_alone_models3*. Note that the EPA has made some modifications to the NCAR releases of MM5 to tailor the modeling system for air quality applications. (See Section 1.13 and Appendix E for summaries of the changes made to the NCAR release.) The version of MM5 provided on the June 1998 Models-3/CMAQ installation tape was tested on the EPA's Cray C90. The scripts and source codes that accompany the June 1999 release of Models-3/CMAQ Version 3 have been tested on Cray C90, Cray T90, and Sun ULTRA 30.

## 1.4 Requirements to Run MM5

According to the NCAR MM5 home page on the Internet (*http://www.mmm.ucar.edu/mm5*), the following user expertise is recommended before attempting to run MM5:

- Experience with numerical modeling of the atmosphere
- At least a Master of Science level of understanding of atmospheric science
- Basic FORTRAN 77 and 90 knowledge
- Basic UNIX knowledge

NCAR also recommends the following hardware requirements for MM5:

- Runs on Cray, SGI, IBM, Sun, DEC, HP, PCs with LINUX, and NT
  (The pre-processors are not available for all of these platforms.)
- Available for both serial and parallel (distributed or shared) architectures
- 128+ MB of memory and 2+ GB of disk space
- NCAR Graphics license (no longer required)
- Source(s) of observations and gridded background data sets

Note that the EPA versions of the MM5 programs have only been tested on Cray and Sun, but should be easily expanded to run on other platforms with minor modifications. Also, the recommended memory and disk requirements are very conservative for MM5 applications with

Models-3/CMAQ using nested domains. It is likely that a simulation with four domains (e.g., 108-, 36-, 12-, and 4-km horizontal resolution) to cover a five-day episode will require 10-20 GB of disk space and would run more comfortably with 256-512 MB of memory (depending on domain sizes and model physics options).

It is imperative that new MM5 users exercise patience while setting up and attempting to run MM5. MM5 is a well-documented community model, but there are no "cookbook" instructions. A novice user is unlikely to obtain success on the first try with the software. It is not uncommon (even for experienced users) to run multiple experiments with MM5 before achieving the desired results. Users should expect to invest a significant amount of time in learning to run and interpret MM5, even if they meet NCAR's recommended level of expertise.

## 1.5 Meteorology Model Pre-Processing

Before the meteorology model can be run, several smaller ("pre-processing") programs must be run to set up the domain for the simulation and to generate a set of initial and boundary conditions for the meteorology model. Each of these programs plays a critical role in the definition of the MM5 simulation. Many of the decisions that are made for running MM5 will have a direct impact on the Models-3/CMAQ simulation, so careful planning is required prior to running MM5. A general flow diagram for the MM5 programs is shown in Figure 1. The output from MM5 enters the Models-3/CMAQ system via the Meteorology-Chemistry Interface Processor (MCIP).

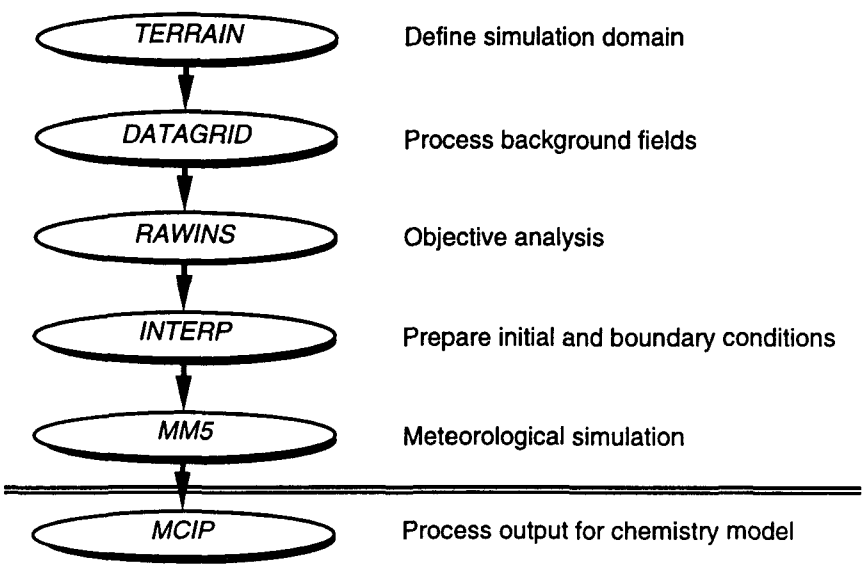


Figure 1. Overview of the MM5 system and its connection to Models-3/CMAQ.

Each of the pre-processors to MM5 is described below. Note that the descriptions specifically pertain to an enhanced version of MM5 Version 2 Release 10 (MM5v2.10), which has been used at the EPA to test and evaluate the June 1999 releases of Models-3/CMAQ Version 3 and the Stand-Alone CMAQ.

## 1.6 Defining the Simulation Domain (*TERRAIN*)

The program, *TERRAIN*, is used to define the MM5 simulation domains. The domains for the meteorology simulations are defined by several primary parameters: number of grid points in each horizontal dimension, grid spacing, center latitude, center longitude, map projection (Mercator, Lambert conformal, or polar stereographic), and number of "nested" domains and their horizontal dimensions. (Some other user-specific parameters are also defined based on the primary parameters, as required.) *TERRAIN* is executed only when a new domain location is required; the same domains can be used for many studies. *TERRAIN* makes use of high-resolution global terrain and land use data sets to create values for each grid point for terrain height and land use specification (e.g., deciduous forest, desert, water). The *TERRAIN* program is thoroughly described by Guo and Chen (1994).

Before running *TERRAIN*, the user must know the region(s) of interest for the Models-3/CMAQ simulation. The settings that are selected for *TERRAIN* (e.g., grid spacing, center latitude, center longitude, map projection) will be applied throughout Models-3/CMAQ. In addition, the MM5 domain(s) defined in *TERRAIN* are generally at least six grid points larger *on each side* than the corresponding Models-3/CMAQ simulation for the same domain. Thus, the MM5 domains should be set for larger areas than the user is planning to model with Models-3/CMAQ.

In order to define the domains, the user must also consider several other factors.

- Major meteorological influences and air pollution precursors in the region of interest. The user must consider emission source distributions and typical flow patterns in the domains, particularly if the user is nesting to higher horizontal resolutions.
- Available meteorological and air quality observations in the region(s) of interest to verify the model solutions. The meteorological observations (particularly special field study data) may also be used for four-dimensional data assimilation (FDDA) in MM5.
- Consistency and completeness of the emissions data for the region(s) and time period(s) of interest, including impacts across international borders.
- Model constraints, especially at very high horizontal resolutions. Although the software may allow the user to establish domains with horizontal resolutions less than one kilometer, these model physics (both MM5 and the CMAQ Chemical Transport Model; CCTM) may not be appropriate for solutions on these scales without adjustments from the user.
- Hardware resources. The user must consider the amount of memory and disk space that would be required for the domain sizes, number of nested domains, and simulation lengths.
- Land use consistency across processors. Each of the models in the Models-3/CMAQ system has a different land-use processor, so care must be taken to set the land use (particularly the land-water mask) consistently between processors. This sometimes means the automated land use categories that are generated in *TERRAIN* must be manually modified to achieve consistency. (For example, users can modify the land use categories in *TERRAIN* with the FUDGE option.) The user must pay particular attention to coastline definitions in urban areas. One goal is to ensure that high-emission grid points are not represented by "water" points in MM5, which would result in insufficient vertical mixing in MM5 in that grid point. This type of problem is most common at lower horizontal resolutions (e.g., 36 km). Also, MM5 uses a dominant land-use category for each grid point, which is different from the diagnostic approach used in MCIP which defines fractional land-use representation in the Models-3/CMAQ grid cells (Byun *et al.* 1999).

There are several user-definable options in *TERRAIN*. Unless the user has sufficient reason to make additional modifications, many of these options should be left on the "default" values specified by NCAR. To run *TERRAIN*, the following items must be defined by the user:

- Horizontal dimensions for each domain (in grid points)
- Horizontal grid spacing for each domain
- Location of each nested domain (with respect to its immediate parent domain)
- Center latitude, center longitude, and map projection
- Input terrain and land-use databases
- Type of nesting (one-way or two-way)

Here is some additional information to keep in mind when using *TERRAIN*:

- MM5 domain dimensions are set by the number of grid points in vector (or "dot") space. Models-3/CMAQ grid dimensions are defined by grid cells, or scalar ("cross") points. The number of scalar points is always one fewer in each dimension than the number of vector points. For example, an MM5 domain with dimensions of 53 x 75 would be represented in Models-3/CMAQ as 52 x 74.
- MM5 grid dimensions are defined as $(y, x)$, where $i$ always refers to the $y$-dimension (relative north-south), and $j$ always refers to the $x$-dimension (relative east-west).
- The MM5 *coarse* domain must have an odd number of grid points in each horizontal dimension such that the center of the domain is located at a [dot] grid point. In the Models-3 framework, the number of rows (NROWS) and columns (NCOLS) must be set to an even integer to satisfy this constraint. (Recall that the number rows and columns in Models-3/CMAQ is one fewer than the number of MM5 grid points.)
- Horizontal grid spacing (delta-$x$ and delta-$y$) for the MM5 nested domains should have a 3:1 ratio. (For one-way nesting, a 3:1 ratio is not strictly required, but is generally recommended.) For example, a valid nesting configuration is 108:36:12:4 km. This nesting configuration is recommended with Models-3/CMAQ.
- The MM5 nest dimensions must be equal to a multiple of 3, plus 1, such that the corner points of the nest are collocated on [dot] grid points of the parent domain. This constraint is from the 3:1 ratio in horizontal grid spacing. For example, a nested domain can have 85 grid points (85 = 28*3 + 1). The grid point dimensions represent the number of "dot points". The number of "cross points" in MM5 is one fewer than the number of dot points in each dimension.
- Nested domains are located by identifying the $(i, j)$ coordinates of the nest lower-left corner on the immediate parent domain (not necessarily the coarse domain).
- In MM5, delta-$x$ is inherently equal to delta-$y$.
- Generally, no more than four domains are run for a simulation because of the amount of data that is created. Memory requirements may limit two-way nesting options, as well.
- Horizontal grid spacing of 6-9 km is often avoided because of the scientific uncertainty in the between-scale physics of cloud processes.
- The smallest horizontal grid spacing that has been tested with Models-3/CMAQ is 4 km. Simulations on higher resolutions may give uncertain results.
- Models-3/CMAQ has only been tested with the Lambert conformal map projection.
- The static land-use databases that are obtainable from NCAR do not account for sea ice. For winter-time cases, ice points will need to be "manually" inserted into the domain. (The FUDGE option in *TERRAIN* may help accomplish this task.)
- The nests for MM5 should not be set within or near the lateral boundaries (external five grid points) of the parent domain. A rule of thumb is for the nest boundaries to be at least 10 grid points from the boundaries of the parent domain.

- Domain boundaries should not be placed along complex topography (e.g., mountain ranges) or coasts. Domain boundaries should be placed at least 10 grid points beyond the region of interest.
- MM5 output files can get fairly large, often 1-2 GB or more for each domain. If disk space is an issue, consider that a domain of 51 x 51 x 30 grid points will generate more than 6.7 MB of data per output interval (called "time step" in much of the Models-3 documentation), or 737 MB for a 110-h simulation with hourly output using the "Models-3/CMAQ Version 3 Tutorial" physics options in MM5v2.10 on a 32-bit workstation. (This is *per domain*, regardless of horizontal grid spacing.) A domain with 85 x 85 x 30 grid points will generate more than 12.6 MB per output interval, or 1.4 GB for the same simulation. Refer to Appendix B for the formula to calculate disk usage. Note that this is only for the model, MM5, not the entire MM5 system; it does not include disk space required for output from the MM5 pre-processors or for temporary files.

The *TERRAIN* program is run one time to process all domains in a nest configuration in the same program execution. The *TERRAIN* program generates a single output file for each domain, with the naming convention "TERRAIN_DOMAIN*n*", where *n* is a positive integer. Lower values of *n* correspond to the lower resolution domains. A schematic of the *TERRAIN* input and output files is shown in Figure 2.

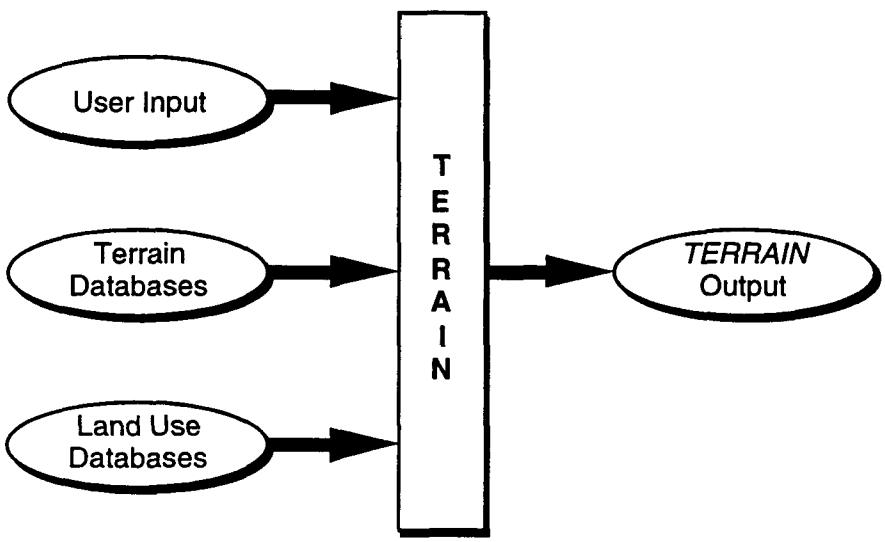


Figure 2. *TERRAIN* inputs and outputs.

## 1.7 Processing the Meteorological Background Fields (*DATAGRID*)

After the simulation domain has been established, the program *DATAGRID* is run to process the meteorological background ("first-guess") fields. *DATAGRID* generates first-guess fields for the model simulation by horizontally interpolating a larger-scale data set (global or regional coverage)

to the simulation domain. *DATAGRID* generates background fields for the model on the pressure levels for which there are data from the larger-scale data set. *DATAGRID* interpolates the background fields to the simulation domain for times throughout the simulation period; these fields are used ultimately to generate lateral boundary conditions for the coarse-domain simulations, and they form the basis for the objective analysis in *RAWINS* (described below). (Note that nested domains obtain lateral boundary conditions from their parent domains.) *DATAGRID* also processes the sea-surface temperature and snow files by interpolating the analyses to the simulation domain. Lastly, *DATAGRID* calculates latitude, longitude, map-scale factors, and Coriolis parameters at each grid point to be used by MM5. The *DATAGRID* program is thoroughly described by Manning and Haagenson (1992).

There are several user-definable options in *DATAGRID*. Unless the user has sufficient reason to make additional modifications, many of these options should be left on the "default" values specified by NCAR. To run *DATAGRID*, the following items must be defined by the user:

- Source of input larger-scale gridded data
- Source of sea-surface temperature (SST) and snow fields
- Simulation start and end times
- Model "top"

Here is some additional information to keep in mind when using *DATAGRID*:

- *DATAGRID* is specifically written to accommodate sources of data that can be obtained from the NCAR archive. These data include 2.5°-resolution global analyses from the National Centers for Environmental Prediction (NCEP) and from the European Centre for Medium Range Forecasting (ECMWF). Other sources of data (e.g., NCEP's Eta model output) can be and have been used with MM5, but the standard version of *DATAGRID* does not accommodate these sources. The user may query the MM5 community for access to software to read other data sources.
- Starting with MM5 Version 3 (MM5v3; released by NCAR in July 1999), *DATAGRID* is renamed and split into two subprograms to more freely accommodate different sources of larger-scale gridded background data.
- When establishing model start times in *DATAGRID*, a "spin-up" time (e.g., 12 hours) prior to the start of the Models-3/CMAQ simulation is generally allowed. That is, the first portion of the MM5 simulation is generally not applied in the air quality simulation to enable the spin up of clouds and other physical processes. (Note: This does *not* imply that the first 12 hours of MM5 simulations cannot be used for other applications.)
- MM5 can be run for longer simulation lengths than in typical forecasting applications, when FDDA is used throughout retrospective MM5 simulations for air quality. Typical MM5 simulation lengths for air quality with FDDA have been on the order of 4.5 to 6.5 days. There is no known research that suggests an optimal or a maximum simulation length, as long as FDDA is used throughout the simulation period to moderate the error growth.
- MM5 simulation times are given in terms of Greenwich Mean Time (GMT), which is also known as Universal Time Constant (UTC) and Zulu.
- If the simulation period is longer than a typical MM5 run for air quality work, the MM5 simulation may be broken into multiple parts (e.g., run in 4.5-day segments). Since each part of the simulation is re-initialized (assuming the "restart" option is not being used), there will be a physical discontinuity between each segment of the simulation. The user should set the start times (accounting for the spin-up periods) to minimize discontinuities during CCTM simulations (typically a 24-hour period from 00 UTC to 00 UTC). If a

spin-up period is allowed, there will be temporally overlapping segments in the MM5 simulation.

- The calculation of the temporal tendency term in Models-3/CMAQ currently requires MM5 output to be available for two hours beyond the period of interest. For example, if the Models-3/CMAQ period of interest ends at 00 UTC, the MM5 simulation must be run through at least 02 UTC on that day. To adequately capture the lateral boundary conditions and analyses for FDDA, *DATAGRID* must [typically] include the background fields of the next synoptic time (00 UTC or 12 UTC) beyond the Models-3/CMAQ period of interest.
- A typical model top is 100 mb.
- The user should examine the availability of the background data to ensure that each synoptic time (i.e., 00 UTC and 12 UTC) has valid analyses and there are no missing data during the time period of interest. NCAR data availability is listed in the "catalog", available electronically from the NCAR anonymous-ftp site. If data are missing during the time period of interest, it is strongly recommended that the user find an alternate source of background data or opt to simulate a different period.

The *DATAGRID* program is run one time to process all domains in a nest configuration in the same program execution. The *DATAGRID* program generates a single output file for each domain, with the naming convention "DATAGRID_DOMAIN*n*", where *n* is a positive integer. Lower values of *n* correspond to the lower resolution domains. A schematic of the *DATAGRID* input and output files is shown in Figure 3.

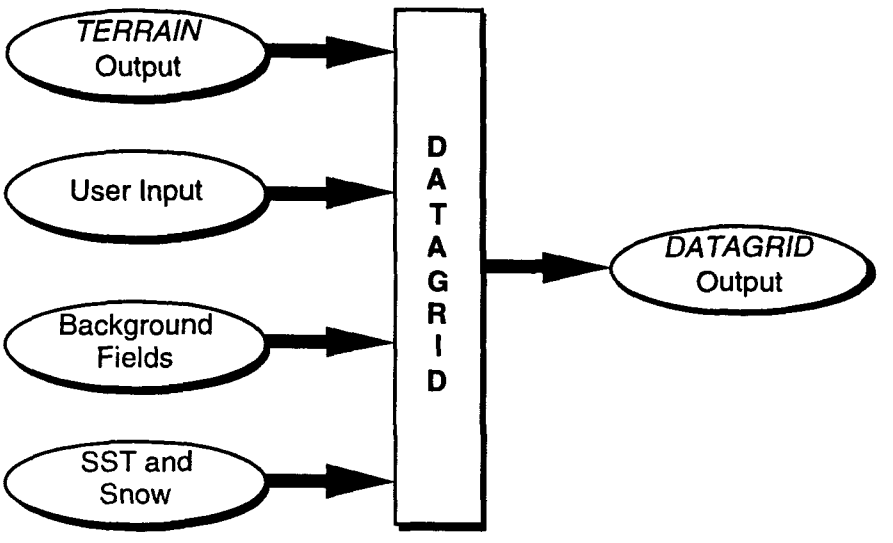


Figure 3. *DATAGRID* inputs and outputs.

## 1.8 Objective Analysis (*RAWINS*)

The program *RAWINS* performs an objective analysis by blending the first-guess fields generated by *DATAGRID* with [generally] twelve-hourly upper-air and three-hourly surface observations. *RAWINS* performs some simple quality control and buddy checking on observations with user-defined thresholds. However, observation quality control is generally the responsibility of the user. The objective analyses in *RAWINS* are performed on pressure levels: both "mandatory" levels and user-defined "supplemental" levels. *RAWINS* generates analyses on this collection of levels, and it also generates a file of surface analyses that can be used in analysis nudging FDDA. The *RAWINS* program is thoroughly described by Manning and Haagenson (1992).

There are several user-definable options in *RAWINS*. Unless the user has sufficient reason to make additional modifications, many of these options should be left on the "default" values specified by NCAR. To run *RAWINS*, the following items must be defined by the user:

- "Supplemental" pressure levels
- Analysis method (4 options)
- Source(s) of observations

Here is some additional information to keep in mind when using *RAWINS*:

- *RAWINS* is specifically written to accommodate observations that can be obtained from the NCAR archive. These data include twelve-hourly rawinsonde soundings and three-hourly surface data. Other sources of observations may be used with MM5, but the user must put the new observations in the format *RAWINS* expects (a non-trivial task).
- When specifying the supplemental pressure levels, note whether or not 925 mb has been included in the background data set. Even though 925 mb has been a mandatory pressure level since 1992, some of NCAR's background data sets since then do not include it. In this case, the user should include 925 mb in the supplemental levels.
- There should typically be 20-25 total pressure levels for analysis.
- A typical analysis method is the "banana" scheme.

The *RAWINS* program can either be run once to include all domains for a simulation, or each domain can be run separately. (To support Models-3/CMAQ, each domain is generally run separately.) *RAWINS* generates several output files for each domain, including "RAWINS_DOMAIN*n*" and "RW4DDA_DOMAIN*n*", where *n* is a positive integer. Lower values of *n* correspond to the lower resolution domains. The RW4DDA file contains surface analyses used for analysis nudging. A schematic of the *RAWINS* input and output files is shown in Figure 4.

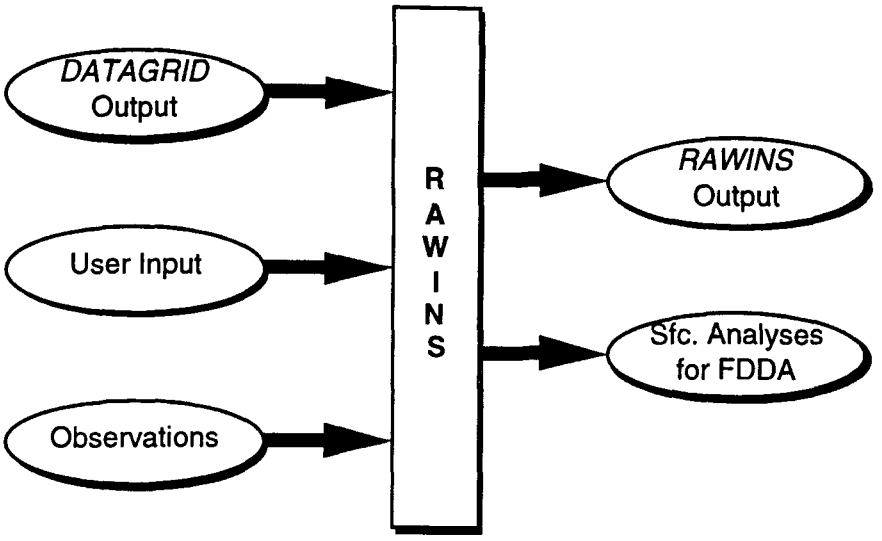


Figure 4. *RAWINS* inputs and outputs.

## 1.9 Setting the Initial and Boundary Conditions (*INTERP*)

The *INTERP* program is designed to fill multiple purposes in the MM5 system. To support Models-3/CMAQ, *INTERP* is used in two different ways to set initial and lateral boundary conditions: "front-end" and "back-end one-way nest". (*INTERP* also serves as an MM5 post-processor in the standard MM5 system, but that functionality is accomplished by MCIP in Models-3/CMAQ.) There is no formal NCAR documentation that is specifically written for *INTERP*. See Dudhia *et al.* (1998) for additional information on *INTERP*.

In the "front-end" mode of *INTERP*, the analyses from *RAWINS* are interpolated to MM5's staggered grid configuration and from their native vertical coordinate (pressure) to MM5's vertical coordinate (a terrain-following, pressure-based "sigma" coordinate). In addition, the state variables are converted as necessary, e.g., relative humidity to specific humidity. The analyses from one time (generally the first time analyzed by *RAWINS*) are interpolated by *INTERP* to provide MM5's initial conditions, while analyses from all times are interpolated to generate MM5's lateral boundary conditions and 3-D analyses for analysis nudging FDDA.

In the "back-end one-way nest" mode of *INTERP*, the MM5 output from the parent domain is interpolated to the nest to provide initial and lateral boundary conditions for the nested run. The one-way nested initial conditions file only contains information for the requested initial time period (unlike the standard front-end initial conditions file which includes subsequent times, as well). The original standard front-end initial conditions file is still used for analysis nudging FDDA for one-way nest applications, since the analyses at subsequent times are not available in the one-way nested initial conditions file. The lateral boundary conditions for one-way nesting are typically at a much higher temporal frequency than the standard front-end processing (i.e., hourly vs. twelve-hourly), since model output is often written more frequently than synoptic intervals.

There are several user-definable options in *INTERP*. Unless the user has sufficient reason to make additional modifications, many of these options should be left on the "default" values specified by NCAR. To run *INTERP*, the following items must be defined by the user:

- Mode of operation (front-end or back-end one-way nest)
- Simulation length
- Vertical (sigma) structure for MM5
- Non-hydrostatic reference variables (temperature, pressure, lapse rate)

Here is some additional information to keep in mind when using *INTERP*:

- There are likely to be more vertical layers in MM5 than in the CCTM. In MM5, the horizontal resolution should be consistent with the vertical resolution for appropriate representation of the physical processes in the atmosphere. Keep in mind that the same vertical resolution is used for all MM5 domains in the simulation. It is recommended that no fewer than 25 layers be used for simulations that include a 4-km domain. The vertical layers should not be compromised in the planetary boundary layer, and there should be more concentrated layering in the planetary boundary layer for air-quality modeling support. Consult with an experienced mesoscale modeler before deviating from the vertical structure provided.
- Starting with MM5v3, *INTERP* is split into three programs that represent the three functionally independent modes of operation.
- The number of flux (full) sigma levels specified in the *INTERP* input is one more than the number of scalar (half-sigma) levels in MM5.
- The simulation end time can be shorter than specified in *DATAGRID*, since the data have already been processed for the long period.
- The simulation start time should remain as it was set in *DATAGRID*. A change to the start time will require modifications to the FDDA files (assuming FDDA is used).
- The non-hydrostatic reference temperature may need to be adjusted seasonally, but there is no specific research that supports this hypothesis.
- The non-hydrostatic reference lapse rate does not reverse at the tropopause. Perturbations from the reference lapse rate above the tropopause will tend to become large, so upper stratospheric results may be less reliable.
- The input information for front-end and back-end one-way nest is very different, but the same scripts and source code can be used for both modes of *INTERP*.

The *INTERP* program can either be run once to include all domains for a simulation, or each domain can be run separately. (To support Models-3/CMAQ, each domain is generally run separately.) The front-end *INTERP* program generates two output files for each domain, with the naming conventions "MMINPUT_DOMAIN*n*" and "BDYOUT_DOMAIN*n*", where *n* is a positive integer. Lower values of *n* correspond to the lower resolution domains. The MMINPUT files are used for both the initial conditions and the 3-D analyses for analysis nudging. One-way nested output files are called "MMINPUT_1WAY_DOMAIN*n*" and "BDYOUT_1WAY_DOMAIN*n*". A schematic of the *INTERP* input and output files for the front-end mode is shown in Figure 5, and the back-end one-way nest mode is shown in Figure 6.

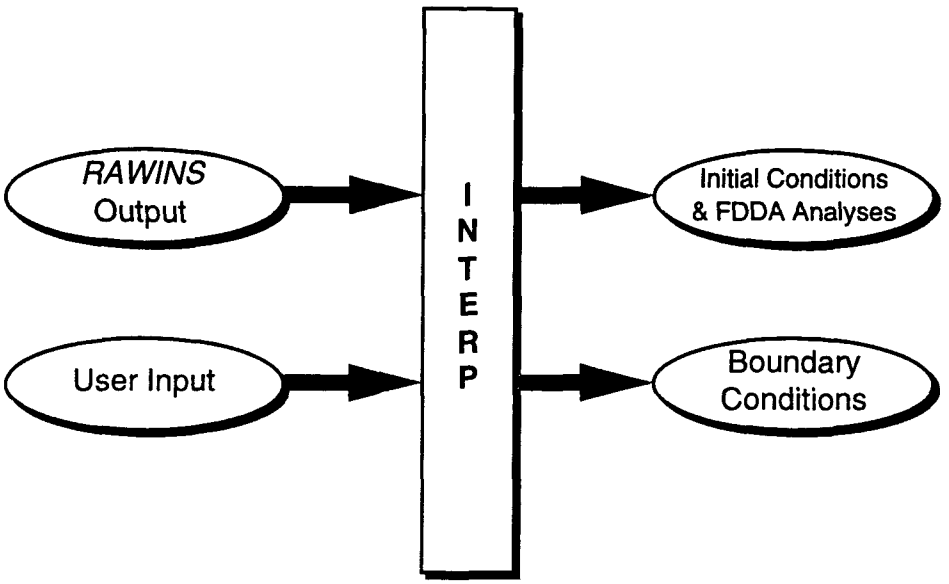


Figure 5. Front-end *INTERP* inputs and outputs.

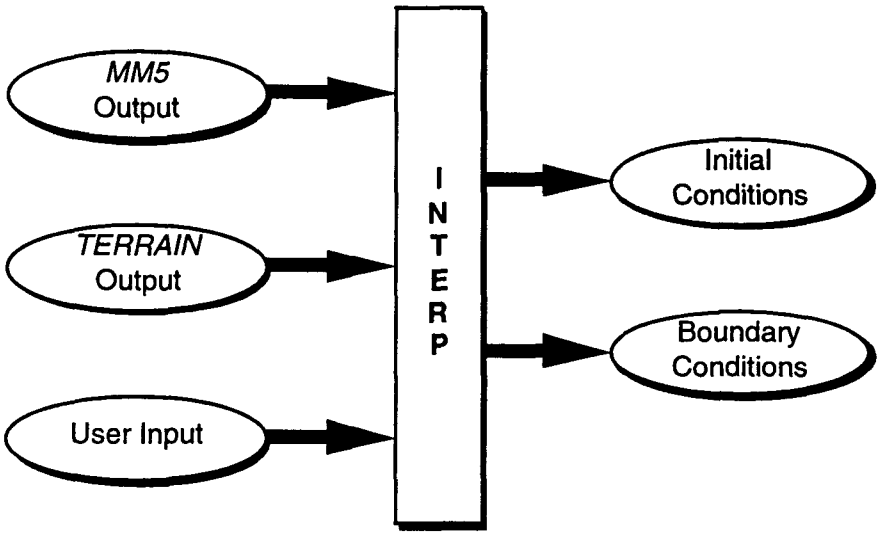


Figure 6. Back-end one-way nest *INTERP* inputs and outputs.

## 1.10 The Meteorology Model (MM5)

MM5 is a complex, state-of-the-science, mesoscale meteorological model. MM5 uses full-physics equations of momentum, thermodynamics, and moisture. It can be run using either hydrostatic or

non-hydrostatic equations. (Non-hydrostatic equations are generally used with Models-3/CMAQ. Hydrostatic equations are not available in MM5v3.) The state variables include temperature, specific humidity, grid-relative wind components, and pressure. The state variables are mass-weighted with a modified surface pressure throughout the model. In the non-hydrostatic equations, pressure, temperature, and density are defined by a reference state and perturbations from the reference state. The vertical (sigma) coordinate is defined as a function of reference pressure in the non-hydrostatic model, and it is time-independent.

Thorough descriptions of the standard MM5 options are found in Grell *et al.* (1994) and Dudhia *et al.* (1998). See Otte (1999) for information on how these options relate to Models-3/CMAQ. The source code for an early release of MM5 is documented in Haagenson *et al.* (1994). The Models-3/CMAQ package does not include NCAR's most current release of MM5. MM5 user options have been tailored and expanded in the Models-3 release to support air quality modeling. Some of these options have not yet been accepted into the official version that is maintained by NCAR. Caution should be exercised when deviating from the version of MM5 that is included with Models-3/CMAQ. Refer to Section 1.13 and Appendix E for summaries of the EPA-initiated changes to MM5 for Models-3/CMAQ.

There are several user-definable options in MM5. Unless the user has sufficient reason to make additional modifications, many of these options should be left on the "default" values specified by NCAR. To run MM5v2.10 with the EPA scripts, the following items must be defined by the user:

- Domain start and end times
- Time step for model integration (frequency of calculations)
- Output frequency (e.g., 1 hour, 3 hours, 15 minutes)
- Planetary boundary layer scheme (4 options)
- Cumulus parameterization scheme (7 options)
- Explicit moisture scheme (7 options)
- Radiation scheme (4 options)
- Type(s) of FDDA and nudging coefficients (2 types of FDDA)

Here is some additional information to keep in mind when using MM5:

- The output from MM5 is in a binary "MM5" format. It is not in I/O API like the other Models-3/CMAQ programs. A sample FORTRAN 90 program to read the MM5 output is provided in Appendix D.
- There are several fields that are output by MM5 at each output interval. These fields are described in Appendix C.
- If MM5 is run on the Cray supercomputer and the other Models-3/CMAQ programs are to be run on a UNIX workstation or Windows NT, the MM5 output may need to be converted to IEEE format. A utility script and a program (*ieee.csh* and *ieee.f*) are available via anonymous-ftp from the NCAR site to perform the IEEE conversion.
- MM5 cannot be run through the Models-3 Framework. MM5 is run by a series of C-shell scripts.
- There are many physics options in MM5 with various degrees of sophistication. The physics options are not interchangeable, and several options have been developed for specific purposes. The current release of Models-3/CMAQ was evaluated with only one set of physics options on each domain. There is no "standard" set of physics options for MM5. Consult with an experienced MM5 user before selecting physics options. Also, using certain physics options will require modifications to MCIP to properly accommodate those options in Models-3/CMAQ.

- FDDA should be used throughout the simulation for air quality studies. Use the "Models-3/CMAQ Version 3 Tutorial" settings for FDDA in Models-3/CMAQ runs. (Note that the nudging coefficients vary with horizontal resolution and state variable.) Consult with an experienced MM5 user before altering the settings.
- The 4-km domains have not been set up with FDDA for the tutorial simulations. Analysis nudging is not recommended at that resolution (*cf.* Stauffer and Seaman 1994). Observation nudging will be used in the MM5 12-km and 4-km domains for a subsequent release of Models-3/CMAQ. Until then, the results on the MM5 4-km domain should be used with caution for lengthy simulations (i.e., beyond 36 hours, depending on many factors).

For one-way nested applications of MM5, each domain is run separately. MM5 generates a single output file for each domain, with the naming convention "MMOUT_DOMAIN*n*", where *n* is a positive integer. Lower values of *n* correspond to the lower resolution domains. Additional output files may be generated, as well, depending on the user options selected. However, this is the primary output file, and it is the only MM5 file read by MCIP. A schematic of the MM5 input and output files is shown in Figure 7.

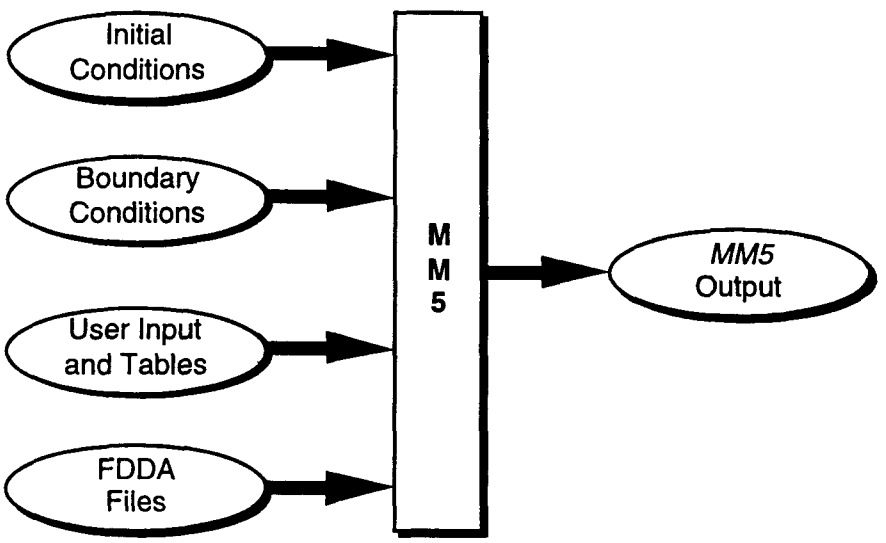


Figure 7. MM5 inputs and outputs.

## 1.11 Meteorology Model Visualization

There are several visualization programs that are compatible with MM5, including include Vis5D, NCAR Graphics, GrADS, and GEMPAK. Some programs are accessible as "free-ware" or "share-ware"; others require a software license. These programs often require a change of format (using a converter) before the visualization program can read the data.

Vis5D is available as "free-ware" from *http://www.ssec.wisc.edu/~billh/vis5d.html*. It is a well-documented and versatile visualization program, and it is compatible with MM5 and Models-3/CMAQ. A converter program, "to_v5d", is available from the EPA anonymous-ftp site in the *util* directory. This converter program can run on Cray, Sun, and SGI platforms.

## 1.12 Meteorology Model Post-Processing

The output variables that are generated by an MM5 simulation are not always useful in their raw form. Those variables must be converted into fields that are required by the chemistry and emission models. The conversion of MM5 output for the other Models-3/CMAQ programs is accomplished in MCIP, which is discussed in Byun *et al.* (1999). MCIP computes variables that are useful to the subsequent models in Models-3, and it creates an output file written with the I/O API libraries in netCDF format that is standard in the Models-3/CMAQ modeling system.

## 1.13 Changes to the MM5 System's Software for Models-3/CMAQ

An enhanced version of MM5 Version 2, Release 10, has been evaluated with the June 1999 releases of Models-3/CMAQ and the Stand-Alone CMAQ. The system is complete with upgrades and "bug fixes" included through September 1998. Only minor changes have been made since September 1998. (Including all of the changes to MM5 would have jeopardized testing and evaluation of Models-3/CMAQ, which uses the MM5 output.) This version of MM5 includes most of the science options in NCAR's current version. The omission of upgrades should not result in substandard or corrupted meteorological simulations.

Below is a summary of EPA-initiated changes that were made to the NCAR release of MM5 and its supporting software to accommodate Models-3/CMAQ. A complete listing of line changes to MM5v2.10 is in Appendix E.

All Programs:

- Standardized the radius of the earth as 6370.997 km to be consistent with chemistry and emission models.
- Changed "configure.make" to enable some machine-specific variables to be set in the C-shell script.

*TERRAIN*:

- Set terrain height over ocean to zero when using the 30-second terrain database.
- Improved representation of urban areas along coasts.

*INTERP*:

- Increased loop indices and parameter sizes to accommodate one-way nested simulation lengths greater than 100 hours.

MM5:

- Modified source code to enable analysis nudging on a one-way nested simulation.
- Modified KZ profile in Blackadar and Bulk PBL schemes.
- Modified cloud constraint in Kain-Fritsch cumulus parameterization scheme.

# Chapter 2
# MODELS-3/CMAQ VERSION 3 TUTORIAL — MM5

## 2.1 Background

This tutorial for MM5 accompanies the Models-3/CMAQ Version 3 Tutorial. (Models-3/CMAQ Version 3 and the Stand-Alone CMAQ were released in June 1999.) This tutorial is not related to and is not a substitute for NCAR's MM5 Tutorial, offered semi-annually in Boulder, Colorado, for a nominal fee. Although the MM5 output is provided with the Models-3/CMAQ Version 3 Tutorial so users have the opportunity to begin with Models-3/CMAQ, *all users are strongly encouraged to run MM5 and create their own output.* This tutorial is based on MM5v2.10 enhanced with EPA modifications, and it is designed to be run on either a Sun workstation or a Cray supercomputer. The user may acquire the latest version of MM5v2 from NCAR via anonymous-ftp from *ftp://ftp.ucar.edu* in the directory */mesouser* and [carefully] retrofit the EPA modifications in Appendix E to the code, or the user may acquire the EPA code from *ftp://www.epa.gov/pub/asmdnerl/stand_alone_models3*. Note that this tutorial will not work with MM5v3, which has significantly different data formats than MM5v2. The tutorial is designed to use the C-shell scripts that are available from the EPA. The NCAR scripts can also be used, with some minor changes. The novice user should read the previous sections of this document to become familiar with MM5. The user should also acquire the NCAR documentation (Grell *et al.* 1994; Guo and Chen 1994; Haagenson *et al.* 1994; and Manning and Haagenson 1992) and tutorial notes (Dudhia *et al.* 1998 or later), which are available electronically in PostScript format via anonymous-ftp or in hard copy from NCAR.

MM5v2.10 (which was evaluated with Models-3/CMAQ and is provided by the EPA for the user's convenience) is not Year 2000 Compliant. Users can continue to run MM5 after 31 December 1999 on retrospective cases up to that date. MM5v2.12 (released by NCAR on 3 April 1999) is Year 2000 Compliant for all code *except* FDDA. MM5v3 (released by NCAR in July 1999) is fully Year 2000 Compliant.

The MM5 output data sets provided with the Models-3/CMAQ Version 3 Tutorial were run on a Cray C90 using a FORTRAN 90 compiler. Users who attempt to run this tutorial on a workstation will not exactly (bit-for-bit) replicate the output in the Cray-generated files. For the Sun workstation, *DATAGRID* and MM5 should be compiled in FORTRAN 77; the other programs use FORTRAN 90. See Section 2.11 for a summary of differences between Cray-generated and Sun-generated MM5 output files for the Models-3/CMAQ Version 3 Tutorial case.

It is imperative that new MM5 users exercise patience while setting up and attempting to run MM5. MM5 is a well-documented community model, but there are no "cookbook" instructions. A novice user is unlikely to obtain success on the first try with the software. It is not uncommon (even for experienced users) to run multiple experiments with MM5 before achieving the desired results. Users should expect to invest a significant amount of time in learning to run and interpret MM5, even if they meet NCAR's recommended level of expertise.

This tutorial assumes a basic knowledge of FORTRAN, UNIX, and C-shell scripting. In addition, the user is required to think through some of the concepts presented in Chapter 1. This

tutorial is not a "cookbook", per se, and step-by-step instructions have not been provided. Users should consult the appropriate documentation before requesting assistance from either NCAR or the EPA.

This tutorial only addresses the mechanics of running MM5. There is no attempt made here to interpret or evaluate results, although this is a very important aspect of running any model. It is strongly recommended that users consider the appropriateness of MM5 runs for their air quality simulations, and they should make adjustments if necessary (e.g., modify physics or settings).

The *mm5* directory on the EPA ftp site includes the following files:

| | |
|---|---|
| Datagrid_EPA.tar.gz | README |
| Datagrid_ws_EPA.tar.gz | Rawins_EPA.tar.gz |
| Interp_EPA.tar.gz | Rawins_ws_EPA.tar.gz |
| MM5_scripts_EPA.tar.gz | Terrain_EPA.tar.gz |
| MM5v2-10_EPA.tar.gz | Terrain_ws_EPA.tar.gz |

Users should access and read the "README" file before downloading any software. Note that these files are based on NCAR releases of the code, but these files have been modified by the EPA to tailor the code for air quality modeling applications. Note that the "README" files *inside the software "tar" files* were prepared by NCAR, and the information in those files does not necessarily pertain to this tutorial or the EPA scripts.

## 2.2 Preparing to Run MM5

Before running MM5, the user must do several things to set up.

- Acknowledge NCAR's recommendations for running MM5.
- Ensure access to hardware that meets the memory and disk space requirements.
- Ensure access to the following compilers and utilities and note the "paths":
  - FORTRAN compiler (either *f77* or *f90*)
  - C compiler (*cc*)
  - C pre-processor (*cpp*)
  - Make utility (*make*)
  - GNU file compressor (*gzip* and *gunzip*)
- Determine whether or not NCAR Graphics libraries and license exist.
- Determine period and region of interest for air quality modeling.
- Acquire MM5 code for pre-processors and model. Make EPA-initiated modifications, if obtaining code directly from NCAR.
- Acquire scripts to run MM5.
- Acquire terrain and land use databases from NCAR anonymous-ftp site.
- Acquire background ("first-guess") data from NCAR or other source.
- Acquire observations from NCAR and/or field study.

If the first-guess and observation data are acquired from a source other than NCAR, the user must make the appropriate modifications to *DATAGRID* and *RAWINS* to accommodate the new data sets.

MM5 and the pre-processors each generate log files and standard output for the user. Much of the information in these files is quite useful, as long as the user learns to interpret the information. All of the programs mark the completion of the code with "`STOP 99999`" in the standard output. (No additional information on the log and standard output files is provided in this tutorial.) *Note that the appearance of this signal in the output does not imply that the program worked properly; it merely completed.* However, all programs in the MM5 system that run to a successful completion will contain this signal in the output. Users should get in the habit of carefully examining the log and output files to ensure a certain level of confidence in their output, rather than just "turning the crank". Although MM5 is a mature modeling system, "operator error" and problems in the model sometimes arise. It is to the user's advantage to address potential MM5 errors (as best as can be done) at the MM5 stage rather than rely on problems in Models-3/CMAQ to discover problems in MM5.

## 2.3 Case Overview

The Models-3/CMAQ Version 3 Tutorial focuses on a high-ozone event in the northeastern United States on 14 July 1995. The one-hour National Ambient Air Quality Standard (NAAQS) for ozone was exceeded at 37 ozone reporting stations in the Northeast on that day, including parts of Maryland, Delaware, Pennsylvania, New Jersey, Long Island, and southern New England (*cf.* Seaman and Michelson 1999). On that day, a weak stationary front was draped from east-to-west north of the U.S.-Canada border in the northeast. High pressure dominated the northeastern U.S., and high temperatures were predominantly in the 90-100°F range. A trough developed in the lee of the Appalachians, and it became oriented along the northeast urban corridor. This trough contributed to the transport and trapping of ozone through the corridor. A large field study sponsored by the North American Research Strategy for Tropospheric Ozone (NARSTO) captured many of the supplemental meteorological and air quality observations in this case.

This case provides the opportunity to simulate important mesoscale meteorological features in a case with weak synoptic forcing. In addition, the effects of the horizontal scale can be evaluated.

## 2.4 Working with the Scripts

The MM5 scripts on the EPA anonymous ftp site are set up to run on either a Cray supercomputer or a Sun ULTRA 30 workstation. The scripts automatically check the type of machine and adjust various machine-specific syntax for either type of hardware. The default compiler is FORTRAN 90, but FORTRAN 77 will be used if the other compiler is not available. (On the Sun workstation, *DATAGRID* and MM5 must be compiled in FORTRAN 77.) The scripts can be adapted to other vendors' machines with limited modifications, e.g., FORTRAN compiler options, location of libraries and compilers, and file unit syntax. NCAR provides template scripts for other platforms in addition to Cray and Sun, so users will have some modest suggestions for adapting the scripts to other hardware configurations.

The scripts for each processor consist of two parts: a "go" and a "deck". The "go" script is used to modify most variables that the average MM5 user will want to change. The "go" script starts the "deck" script. Other than changing library and compiler paths and changing syntax for another type of hardware, the "deck" scripts should not need to be modified for individual case studies.

Near the top of each "go" script, there is a series of lines that begins with "#QSUB". These lines are required for processing on the Cray, and they control memory and time limits for the "batch" jobs. Workstations interpret these lines as comments and otherwise ignore them.

Some variables are common to many of the "go" scripts. Table 1 defines the path variables that appear in more than one "go" script. The path variables are provided as a convenient way to organize the system on the user's machine. The user is free to modify these variables and adjust

the "deck" scripts in the appropriate locations. Note that the locations of the various XxxxDir in Table 1 should not be nested within each other. A sample directory structure that is compatible with this tutorial can be found in the scripts. Users should address the disk space requirements for each XxxxDir before establishing a final directory structure.

Table 1. Path variables in MM5 scripts.

| Variable | Description |
|---|---|
| Prog | Name of the program, e.g., Terrain, Datagrid. |
| ExpName | Name of the experiment. |
| BaseName | Leading path that is common to most other paths. |
| ProjName | Path to this experiment. |
| DataBase | Path to first-guess fields and observations. |
| DataDir | Path to input and output data for this experiment. |
| DeckDir | Path to the "deck" scripts. |
| ProgDir | Path to the source code for this program. |
| StatDir | Path to static files, such as terrain and land use databases. |
| TempDir | Path to temporary directory where this script will run. |
| UtilDir | Path to utility files that are common to many scripts. |

Some additional variables appear often in scripts, and these variables control some of the functions within the scripts. Table 2 defines control variables that are common to several scripts.

Table 2. Variables common to "go" scripts.

| Variable | Description |
|---|---|
| DomId | Domain Identifier. Coarsest domain is "1", first nest is "2", etc. |
| OutIEEE | For Cray only. Yes, if additional IEEE files will be made. |
| Host | Location of source code. Can be either a local directory or a remote machine. |
| Repository | Yes, if code is in a "*cvs*" repository. |

## 2.5 Running *TERRAIN*

Table 3 lists the primary user-definable variables in the *TERRAIN* script. See the "terrain.go" script for details on how these variables should be set.

Table 3. Variables specific to "terrain.go".

| Variable | Description |
|---|---|
| compile | Indicate whether or not to compile *TERRAIN*. Otherwise assume the executable exists. |
| NCAR | Indicates whether NCAR Graphics is available. |
| TerPlt | Indicates whether *TERRAIN* will be run, or if only maps will be generated in NCAR Graphics. |
| CentLat | Center latitude. |
| CentLon | Center longitude. |
| Projection | Map projection. |
| NumDoms | Number of domains. |
| IDim | I ($y$) dimension of each domain. |
| JDim | J ($x$) dimension of each domain. |
| DeltaX | Horizontal grid spacing for each domain. |
| NestLLI | Lower-left corner $i$ coordinate for each nest with respect to immediate parent domain |
| NestLLJ | Lower-left corner $j$ coordinate for each nest with respect to immediate parent domain |
| Expand | Indicates whether or not an "expanded grid" is used for coarse domain objective analysis |
| ExpandSize | Amount of grid expansion on each side of coarse domain. |
| TerLURes | Terrain and land-use databases that will be used. |
| NestType | Type of nesting for each domain. |

For the Models-3/CMAQ Version 3 Tutorial, four nested domains will be run. The nests have increasing focus over the northeastern United States and ultimately cover the New York City metropolitan area. Figure 8 shows the MM5 simulation domains for the Models-3/CMAQ Version 3 Tutorial. The following settings should be used for the tutorial:

- The simulation domains are centered at 40°N, 90°W, on a Lambert conformal projection.
- Domain 1 is a 108-km domain with 53 x 75 grid points, using 30-minute terrain and land use databases and one-way nesting.
- Domain 2 is a 36-km domain with 43 x 52 grid points and a lower-left corner of (21, 37), using 10-minute terrain and land use databases and one-way nesting.

- Domain 3 is a 12-km domain with 64 x 58 grid points and a lower-left corner of (13, 28), using 5-minute terrain and 10-minute land use databases and one-way nesting.
- Domain 4 is a 4-km domain with 49 x 49 grid points and a lower-left corner of (32, 32), using 30-second terrain and 10-minute land use databases and one-way nesting.
- An expanded grid of 500 km is used for the coarse domain.

## 2.6 Running *DATAGRID*

Table 4 lists the primary user-definable variables in the *DATAGRID* script. See the "datagrid.go" script for details on how these variables should be set.

Table 4. Variables specific to "datagrid.go".

| Variable | Description |
|---|---|
| NumDoms | Number of domains. |
| YYYY | Four-digit year for simulation start date. |
| MM | Two-digit month for simulation start date. |
| DD | Two-digit day for simulation start date. |
| HH | Two-digit hour (UTC) for simulation start time. |
| TimIntvl | Time interval of available first-guess fields. |
| NumFiles | Number of first-guess times requested for processing. Generally set as simulation length (hours) divided by 12, plus 1. (Round up to next highest integer.) |
| ModelTop | Pressure at the model top (millibars). |
| MaxI | Maximum grid points in north-south direction on any domain, including expanded dimension in coarse domain. |
| MaxJ | Maximum grid points in east-west direction on any domain, including expanded dimension in coarse domain. |
| Analysis | Source of first-guess files, e.g., NMC, ECMWF, TOGA. See the script for details on these data sources. |
| InAnalysis1 | NCAR catalog number for first first-guess file. |
| InAnalysis2 | NCAR catalog number for second first-guess file. |
| SST | Source of sea-surface temperature analysis (e.g., NMC, Navy, climatology). |
| InSST1 | NCAR catalog number for first SST file. |
| InSST2 | NCAR catalog number for second SST file. |
| InSnow1 | NCAR catalog number for first snow analysis file. |
| InSnow2 | NCAR catalog number for second snow analysis file. |
| ANDLEV2 | Special variable for EPA research. |

The Models-3/CMAQ Version 3 Tutorial focuses on 14 July 1995. Recall that a "spin-up period" must be considered in MM5 prior to the Models-3/CMAQ time period of interest. In addition, the calculation of the temporal tendency term in Models-3/CMAQ currently requires MM5 output to be available for two hours beyond the period of interest. Thus, for the 24-h period from 00 UTC 14 July 1995 to 00 UTC 15 July 1995, MM5 must be run for a 38-h simulation (including a 12-h spin-up period and an extra 2 h at the end). The following settings should be used for the MM5 tutorial:

- All of the domains that were processed in *TERRAIN* will be run.
- The MM5 simulation start time is 12 UTC 13 July 1995.
- The expanded grid added 10 grid points to each dimension in the coarse domain. Consider this when determining the maximum *i* and *j* grid points.
- The model top is 100 millibars.
- The background and SST files are from NMC (NCEP), and they are available at 12-h intervals. The catalog numbers for the period can be found in the NCAR catalog.
- Since this is a summer case, no snow files will be used.
- The EPA research variable should be set to "no" for each domain.

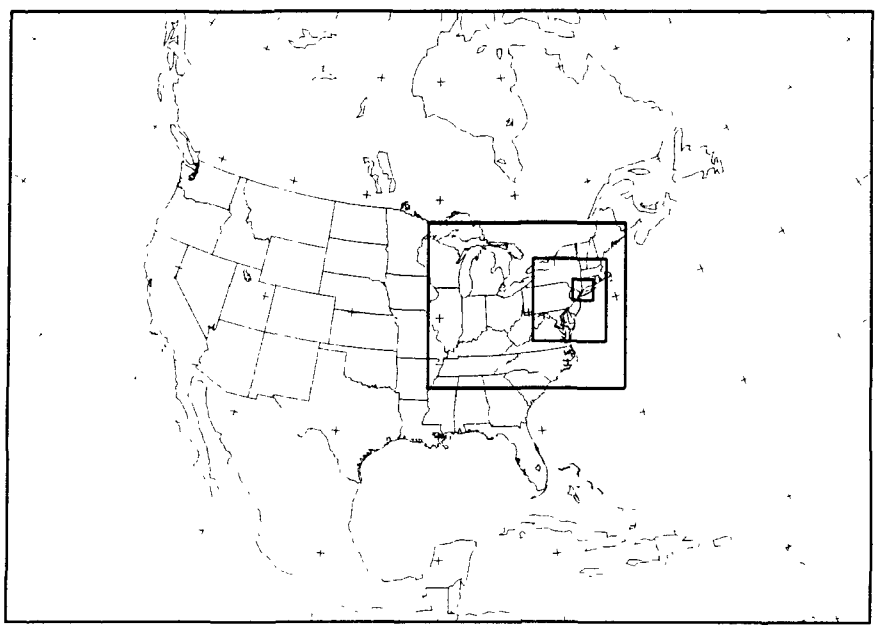

Figure 8. MM5 domains for Models-3/CMAQ Version 3 Tutorial.

## 2.7 Running *RAWINS*

Table 5 lists the primary user-definable variables in the *RAWINS* script. See the "rawins.go" script for details on how these variables should be set. Since *RAWINS* is generally run once for each domain to support Models-3/CMAQ, it would be helpful to develop separate "go" scripts for each domain.

The following settings should be used in *RAWINS* for the tutorial:

- Each domain should be processed separately using the same options.
- There are 23 total pressure levels, including the surface.
- The supplemental pressure levels are: 975, 950, 925, 900, 875, 800, 750, 650, 600, 550, 450, and 350.
- The observation files are from the NCAR archive. Upper-air, 6-hourly, and 3-hourly observations are used. The catalog numbers can be found in the NCAR catalog.
- The analysis method is "banana".
- FDDA surface files should be generated.
- There is no autobogusing and no bogusing.
- The buddy-weighting factor is 1.25. The tolerances for temperature, wind, and pressure are 10, 13, and 6, respectively.
- The 4-km domain does not need to be run through *RAWINS*, since the initial and boundary conditions are derived from the 12-km domain, and there is no analysis nudging in the 4-km domain.

Table 5. Variables specific to "rawins.go".

| Variable | Description |
|---|---|
| Imax | Grid points in *i* dimension, including expanded grid. |
| Jmax | Grid points in *j* dimension, including expanded grid. |
| Nprs | Number of pressure levels for analysis. |
| NewPrs | User-defined supplemental pressure levels. |
| INOBS | Source of observation files. |
| SFCsw | Indicates whether or not surface observations are used. |
| Submit | Indicates level of autobogus. |
| Method | Analysis method. |
| BudWgt | Weighting factor for buddy-checking of observations. |
| ErrMxT | Tolerance for temperature differences. |
| ErrMxW | Tolerance for wind differences. |
| ErrMxP | Tolerance for pressure differences. |
| IfFDDA | Indicates whether or not to generate surface FDDA files. |
| BOGUSsw | Type of bogusing. |
| Raobs1 | NCAR catalog number for first upper-air obs file. |
| Raobs2 | NCAR catalog number for second upper-air obs file. |
| Sfc6h1 | NCAR catalog number for first 6-h surface obs file. |
| Sfc6h2 | NCAR catalog number for second 6-h surface obs file. |
| Sfc3h1 | NCAR catalog number for first 3-h surface obs file. |

| | |
|---|---|
| Sfc3h2 | NCAR catalog number for second 3-h surface obs file. |
| NCAR | Indicates whether NCAR Graphics is available. |

## 2.8 Running Front-End *INTERP*

Table 6 lists the primary user-definable variables in the *INTERP* script. See the "interp.go" script for details on how these variables should be set. Since *INTERP* is generally run once for each domain to support Models-3/CMAQ, it would be helpful to develop separate "go" scripts for each domain and for each mode of *INTERP*. Note that the same basic scripts are used for both front-end and back-end one-way nest *INTERP* runs.

Table 6. Variables specific to "interp.go".

| Variable | Description |
|---|---|
| ForBsw | Sets mode for *INTERP*. |
| NumDoms | Number of domains in this run. |
| LevIDN | Level of nest for each domain with respect to the configuration set in *TERRAIN*. Represents the number of layers of nesting for this domain from the coarse domain, usually the domain ID minus 1. (0, 1, 2, 3, ...) |
| NumNC | Mother domain ID for each domain with respect to the configuration set in *TERRAIN*. Represents the domain ID of the parent domain to this domain. For Domain 1, set mother domain ID to "1". (1, 1, 2, 3, ...) |
| IDim | Grid points in *i* dimension for each domain in this run. |
| JDim | Grid points in *j* dimension for each domain in this run. |
| NestLLI | Lower-left corner *i*-coordinate for nest. |
| NestLLJ | Lower-left corner *j*-coordinate for nest. |
| Nprs | Number of pressure levels in analysis (excluding surface). |
| FullSig | Full sigma levels for MM5. |
| BegTime | First time processed by *INTERP* (simulation start). |
| EndTime | Final time processed by *INTERP*. |
| TimInt | Time interval between input files. |
| HYDROsw | Sets up for hydrostatic or non-hydrostatic simulation. |
| RefTS0 | Reference surface temperature for non-hydrostatic run. |
| RefTLP | Reference lapse rate for non-hydrostatic run. |
| RefP0 | Reference surface pressure for non-hydrostatic run. |
| ANDLEV2 | Special variable for EPA research. |

The following settings should be used in front-end *INTERP* for the tutorial:

- Each domain should be processed separately using the same options.
- There are 22 total pressure levels in the analysis (excluding the surface).
- The full-sigma levels are: 1.000, 0.995, 0.990, 0.985, 0.980, 0.970, 0.960, 0.945, 0.930, 0.910, 0.890, 0.865, 0.840, 0.810, 0.780, 0.740, 0.700, 0.650, 0.600, 0.550, 0.500, 0.450, 0.400, 0.350, 0.300, 0.250, 0.200, 0.150, 0.100, 0.050, and 0.000.
- The *INTERP* start time is 12 UTC 13 July 1995, and the end time is 12 UTC 15 July 1995. Recall that the MM5 simulation ends on a non-standard time (not 00 UTC or 12 UTC). Thus, the lateral boundary conditions and the 3-D analyses for FDDA must be processed for the next synoptic time (twelve-hourly increment) beyond the simulation end for temporal interpolation.
- The input files are available on 12-hourly intervals.
- MM5 will be run with non-hydrostatic equations. The reference surface temperature is 290 K, the reference lapse rate is 50 K, and the reference surface pressure is 1000 mb.
- The EPA research variable should be set to "no" for each domain.
- The 4-km domain does not need to be run through front-end *INTERP*, since the initial and boundary conditions are derived from the 12-km domain, and there is no analysis nudging in the 4-km domain.

## 2.9 Running MM5

Table 7 lists the primary user-definable variables in the MM5 script. See the "mm5.go" script for details on how these variables should be set. Since MM5 is generally run once for each domain to support Models-3/CMAQ, it would be helpful to develop separate "go" scripts for each domain.

Table 7. Variables specific to "mm5.go".

| Variable | Description |
|---|---|
| NCPUS | Number of CPUs for this job...generally set to 1. |
| ONEWAYsw | Indicates whether domains are started as one-way nests. |
| TIMAX | Length of simulation, in minutes. |
| OutFrq | Output frequency. |
| TiStep | Model integration time step. |
| StDom | Domain start time, in minutes into simulation. |
| EndDom | Domain end time, in minutes into simulation. |
| Restart | Indicates if this run is restarted from (physically linked to) another run. |
| ResTime | Restart time, in minutes into simulation. |
| SavRes | Indicates whether to save data for a restart. |
| SavTim | Frequency to save output for restart. |
| IDim | Number of grid points in *i* dimension (north-south). |

| | |
|---|---|
| JDim | Number of grid points in *j* dimension (east-west). |
| KDim | Number of vertical half-sigma layers (full-sigma minus 1) |
| NestLLI | Lower-left corner *i*-coordinate for nest. |
| NestLLJ | Lower-left corner *j*-coordinate for nest. |
| IMPhys | Explicit moist physics scheme. |
| MPhysTbl | Indicate whether a look-up table is used for moist physics. |
| ICuPa | Cumulus parameterization scheme. |
| IBLTyp | Planetary boundary layer scheme. |
| FRad | Atmospheric radiation scheme. |
| ISoil | Indicate whether the multi-layer soil model is used. |
| IShallow | Indicate whether the shallow convection scheme is used. |
| AnlyNud | Indicate whether analysis nudging FDDA is used. |
| AVCoeff | Analysis nudging coefficient for wind. |
| ATCoeff | Analysis nudging coefficient for temperature. |
| AQCoeff | Analysis nudging coefficient for specific humidity. |
| ObsNud | Indicate whether observation nudging is used. |
| OVCoeff | Observation nudging coefficient for wind. |
| OTCoeff | Observation nudging coefficient for temperature. |
| OQCoeff | Observation nudging coefficient for specific humidity. |
| RINXY | Surface radius of influence for observation nudging. |
| TWINDO | Half-window over which observations will be nudged. |

The following settings should be used in MM5 for the tutorial. Note that these settings only represent one configuration for the air quality simulation. They are not the only options that are valid on the various horizontal resolutions. However, the user should not randomly modify the physics and FDDA options without insight into the impact of using those options. Some physics options have been designed for specific scales and phenomena. The physics options are generally *not* interchangeable.

- Each domain should be processed separately.
- The coarse domain has the ONEWAYsw set to "NO1WAY", since it is not initialized as a one-way nest from another MM5 run. All other domains are set to "1WAY".
- The domain dimensions are as set in *TERRAIN*.
- The lower-left corner of the domain should be set to "1" for each domain, since only one domain is run at a time (not two-way nesting).
- All domains use 30 half-sigma levels. (See front-end *INTERP* instructions.)
- Each domain should be run for a 38-h simulation with hourly output. Each domain starts at "0" and ends at 38 h.
- The model integration time steps for 108-, 36-, 12-, and 4-km are 240 s, 90 s, 30 s, and 10 s, respectively.

- There is no restarting and no need to save restart files.
- Each domain uses Reisner mixed-phase explicit moist physics with the look-up table.
- The cumulus parameterization schemes are as follows: Anthes-Kuo for 108-km; Kain-Fritsch for 36-km; Grell for 12-km; and none for 4-km.
- Each domain uses the Blackadar planetary boundary layer scheme.
- The 4-km domain uses the Dudhia cloud radiation scheme. The other domains use the "none" option for surface radiation only.
- The multi-layer soil model and shallow convection schemes are not used.
- Analysis nudging is used for the 108-, 36-, and 12-km domains.
- The wind and temperature nudging coefficients are $3 \times 10^{-4}$ $s^{-1}$ for the 108- and 36-km domains. The wind and temperature nudging coefficients are reduced to $1 \times 10^{-4}$ $s^{-1}$ for the 12-km domain. Surface wind analyses are used for nudging, but no temperature nudging occurs within the PBL.
- The moisture nudging coefficient is $1 \times 10^{-5}$ $s^{-1}$ for all three domains. No moisture nudging occurs within the PBL.
- No observation nudging is used. (It will be recommended in the future for the 12-km and 4-km domains.)

## 2.10 Running Back-End One-Way Nest *INTERP*

See Table 6 for descriptions of the primary user-definable variables in the *INTERP* script. The following settings should be used in back-end one-way nest *INTERP* for the tutorial:

- Each domain should be processed separately using the same options.
- There are 22 total pressure levels in the analysis (excluding the surface).
- The full-sigma levels are the same as in the front-end runs.
- For purposes of each run, there are two domains in each one-way nest run: the parent domain which provides the MM5 output, and the nest for which the initial and lateral boundary conditions are being generated.
- The domain ID represents the domain for which the initial and lateral boundary conditions are being generated.
- The level of nest and mother domain ID are each set for the parent and nest domains in this one-way run. The level of nest and mother domain ID are fixed for each domain respective to the nesting established in *TERRAIN*. (They are not relative to this one-way run.) The level of nest is the number of nest layers from the coarse domain (usually the domain ID minus 1). The mother domain ID is the domain ID of the immediate parent domain of the domain of interest.
- The domain dimensions are each set for the parent and nest. The lower-left corner coordinates are for the nest (with respect to the parent), as they are defined in *TERRAIN*.
- The *INTERP* start time is 12 UTC 13 July 1995, and the end time is 02 UTC 15 July 1995. (Recall that MM5 must be run for two hours beyond the time period of interest for Models-3/CMAQ.) Since the MM5 output for the parent domain is available hourly for the full 38-h simulation, *INTERP* does not need to be (and cannot be) run beyond the simulation length.
- The input files are available on hourly intervals.

- MM5 will be run with non-hydrostatic equations using the same reference variables as in the front-end runs.
- The EPA research variable should be set to "no" for each domain.
- The 108-km domain does not need to be run through back-end one-way nest *INTERP*, since it is not initialized as a one-way nest from another MM5 run.

## 2.11 Comparison on Various Platforms

MM5v2.10 enhanced with EPA modifications has been run on the Cray C90 and Sun ULTRA 30 (300 MHz and 384 MB RAM). The MM5 Cray simulations on the tutorial domains are 6.45 to 7.75 times faster than the Sun simulations. Note that the Sun simulations were run with FORTRAN 77 and the Cray simulations used FORTRAN 90. FORTRAN 90 has been shown to accelerate MM5 simulations by 10-25%.

Because of the differences in the numerics on the Cray and the Sun (e.g., 64-bit vs. 32-bit), there are differences in the simulation results between the two machines. Changes to FORTRAN compiler options also occasionally result in differences in model simulations. The terrain databases on the different hardware architectures induced differences on the order of $10^{-5}$ m in the tutorial simulation domains, which acted in part to perpetuate differences between the simulations. In addition, the latitude and longitude differences for identical grid points are on the order of $10^{-5}$ degrees. The land use arrays in both simulations were identical. On the 12-km domain (which includes the compounded effects of differences on the 108-km and 36-km domains), maximum differences after a 36-h simulation were on the order of 2 K in temperature, and 2 m $s^{-1}$ in wind. There are also noticeable differences in the heat flux terms and PBL heights at isolated grid points, which is likely related to if-then branching of PBL regimes in the Blackadar PBL scheme. However, most of the largest differences between the Cray and Sun simulations are confined to the lateral boundaries of the domains. Qualitatively, the simulations are very similar.

This information is provided for reference only. It suggests that simulation results may be slightly different when run on different platforms. This is an important consideration for attempting to get bit-for-bit replication of the MM5 portion of the Models-3/CMAQ Version 3 Tutorial. It is also important to note these differences when comparing to other MM5 data sets for the same case that have been generated on different platforms.

# Appendix A. References


Anthes, R. A., and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev.*, **106**, 1045–1078.

Byun, D. W., J. Pleim, R. Tang, and A. Bourgeois, 1999: Meteorology-chemistry interface processor (MCIP) for Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System.* D. W. Byun and J. K. S. Ching, Eds. EPA-600/R-99/030, March 1999. [Available from the United States Environmental Protection Agency, Office of Research and Development, Washington, D.C. 20460.]

Dudhia, J., 1993: A nonhydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493–1513.

Dudhia, J., D. Gill, Y.-R. Guo, D. Hansen, K. Manning, and W. Wang, 1998: PSU/NCAR mesoscale modeling system tutorial class notes (MM5 modeling system version 2). [Available from the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307.]

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 117 pp. [Available from the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307.]

Guo, Y.-R., and S. Chen, 1994: Terrain and land use for the fifth-generation Penn State/NCAR mesoscale modeling system (MM5): Program TERRAIN. NCAR Tech. Note, NCAR/TN-397+IA, 119 pp. [Available from the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307.]

Haagenson, P., J. Dudhia, D. Stauffer, and G. Grell, 1994: The Penn State/NCAR mesoscale model (MM5) source code documentation. NCAR Tech. Note, NCAR/TN-392+STR, 200 pp. [Available from the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307.]

Manning, K., and P. Haagenson, 1992: Data ingest and objective analysis for the PSU/NCAR modeling system: Programs DATAGRID and RAWINS. NCAR Tech. Note, NCAR/TN-396+IA, 209 pp. [Available from the National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307.]

Otte, T. L., 1999: Developing meteorological fields. *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System.* D. W. Byun and J. K. S. Ching, Eds. EPA-600/R-99/030, March 1999. [Available from the United States Environmental Protection Agency, Office of Research and Development, Washington, D.C. 20460.]

Seaman, N. L., and S. A. Michelson, 1999: Mesoscale meteorological structure of a high-ozone episode during the 1995 NARSTO-Northeast study. *J. Appl. Meteor.* In press.

Stauffer, D. R., and N. L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.*, **33**, 416–434.

# Appendix B. File Size

Below is a formula that can be used to predict the size of the output file that will be generated from MM5. The inputs to the formula are: number of grid points in north-south dimension (*ix*), number of gridpoints in east-west dimension (*jx*), number of vertical layers (*kx*), number of output times (*times*), number of 3-D variables (*n3d*), number of 2-D variables (*n2d*), the number of bytes per word (*word*), the size of the header record, which is controlled by the word size (*hdr*), and an adjustment factor (*factor*; 1.002 for Cray and 1.000 for workstation). The output of the formula, *size*, will be the approximate size of the file in bytes. Number of control characters on a record is *ctrl*. Note that *hdr* is fixed at 3.52 MB for a 64-bit machine and 3.36 MB for a 32-bit machine. (*Hdr* may be affected by control characters on some architectures.) The size of *ctrl* will vary depending on the hardware architecture.

$$size = \{times * [hdr + ((n2d * ix * jx + ctrl) * bit) + ((n3d * ix * jx + ctrl) * kx * word)]\} * factor$$

If the user selects the same physics options and model version as in the Models-3/CMAQ Version 3 Tutorial, there will be 10 3-D variables and 21 2-D variables in the MM5 output file.

Example 1: User is running a domain with dimensions 94 x 100 x 30, and output is generated hourly for a 110-hour simulation on a Cray C90. Thus, *ix* is 94, *jx* is 100, *kx* is 30, and *times* is 111 (initial time, plus 110 simulation hours). *Hdr* is 3,520,000, *ctrl* is 0, and *word* is 8.

*size* = { 111 * [ 3,520,000 + (21 * 94 * 100 * 8) + (10 * 94 * 100 * 30 * 8) ] } * 1.002

*size* = { 111 * [ 3,520,000 + 1,579,200 + 22,560,000 ] } * 1.002

*size* = { 111 * 27,659,200 } * 1.002

*size* = 3,070,171,200 * 1.002

*size* = 3,076,311,542 bytes (or 3.1 GB). Actual file size is 3,076,210,688 bytes.

Example 2: User is running a 43 x 52 x 30 domain on a Cray C90, with 63 output times.

*size* = { 63 * [ 3,520,000 + (21 * 43 * 52 * 8) + (10 * 43 * 52 * 30 * 8) ] } * 1.002

*size* = { 63 * [ 3,520,000 + 375,648 + 5,366,400 ] } * 1.002

*size* = { 63 * 9,262,048 } * 1.002

*size* = 583,509,024 * 1.002

*size* = 584,676,042 bytes (or 585 MB). Actual file size is 584,667,136 bytes.

Example 3: User is running a 43 x 52 x 30 domain on a Sun ULTRA 30, with 63 output times.

*size* = { 63 * [ 3,360,000 + (21 * (43*52+2) * 4) + (10 * (43*52*30+2) * 4) ] } * 1

*size* = { 63 * [ 3,360,000 + 187,992 + 2,683,280 ] }

*size* = { 63 * 6,231,272 } = 392,570,136 bytes (or 393 MB) ... actual file size!

# Appendix C. MM5 Version 2 Output Variables

| Field | Variable Name | Data Type | Size of Array | Units | Cross or Dot? | Output When... |
|---|---|---|---|---|---|---|
| Header | MIF,<br>MRF,<br>MIFC,<br>MRFC | Integer<br>Real<br>Char*80<br>Char*80 | 1000 x 20<br>1000 x 20<br>1000 x 20<br>1000 x 20 | N/A | N/A | |
| U-Component Wind coupled with P-Star | U | Real | IX x JX x KX | kPa m $s^{-1}$ | Dot | |
| V-Component Wind coupled with P-Star | V | Real | IX x JX x KX | kPa m $s^{-1}$ | Dot | |
| Temperature coupled with P-Star | T | Real | IX x JX x KX | kPa K | Cross | |
| Specific Humidity coupled with P-Star | Q | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IDRY = 0 |
| Cloud Water Mixing Ratio coupled w/ P-Star | CLW | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IMPHYS ≥ 3 |
| Rain Water Mixing Ratio coupled w/ P-Star | RNW | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IMPHYS ≥ 3 |
| Ice Water Mixing Ratio coupled with P-Star | ICE | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IMPHYS ≥ 5 |
| Snow Water Mixing Ratio coupled w/ P-Star | SNOW | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IMPHYS ≥ 5 |
| Graupel Mixing Ratio coupled with P-Star | GRAUPEL | Real | IX x JX x KX | kPa kg $kg^{-1}$ | Cross | IMPHYS ≥ 6 |
| Number Concentration of Graupel coupled with P-Star | NCI | Real | IX x JX x KX | kPa number $m^{-3}$ J $kg^{-1}$ | Cross | IMPHYS ≥ 6 |
| Turb. Kinetic Energy | TKE | Real | IX x JX x KX | J $kg^{-1}$ | Cross | IBLTYP = 3 |
| Atmospheric Radiative Tendency | RAD TEND | Real | IX x JX x KX | K $day^{-1}$ | Cross | IFRAD ≥ 2 |
| Vertical Velocity coupled with P-Star | W | Real | IX x JX x KX+1 | kPa m $s^{-1}$ | Cross | INHYD = 1 |
| Pressure Perturbation coupled with P-Star | PP | Real | IX x JX x KX | kPa Pa | Cross | INHYD = 1 |

| Field | Variable Name | Data Type | Size of Array | Units | Cross or Dot? | Output When... |
|---|---|---|---|---|---|---|
| P-Star | PSTAR | Real | IX x JX | kPa | Cross | |
| Ground Temperature | GROUND T | Real | IX x JX | K | Cross | ITGFLG = 1 |
| Accumulated Convective Precipitation | RAIN CON | Real | IX x JX | cm | Cross | IDRY = 0 |
| Accumulated Non-Convective Precipitation | RAIN NON | Real | IX x JX | cm | Cross | IDRY = 0 |
| Terrain Elevation | TERRAIN | Real | IX x JX | m | Cross | |
| Map-Scale Factor | MAPFACCR | Real | IX x JX | dimensionless | Cross | |
| Map-Scale Factor | MAPFACDT | Real | IX x JX | dimensionless | Dot | |
| Coriolis Parameter | CORIOLIS | Real | IX x JX | $s^{-1}$ | Dot | |
| Substrate Temperature | RES TEMP | Real | IX x JX | K | Cross | ITGFLG $\neq$ 3 |
| Latitude | LATITCRS | Real | IX x JX | degree | Cross | |
| Longitude | LONGICRS | Real | IX x JX | degree | Cross | |
| Land Use | LAND USE | Real | IX x JX | category | Cross | |
| Snow Cover | SNOWCOVR | Real | IX x JX | dimensionless | Cross | |
| PBL Height | PBL HGT | Real | IX x JX | m | Cross | PBL =1,2,5 |
| PBL Regime | REGIME | Real | IX x JX | category (1-4) | Cross | PBL =1,2,5 |
| Sensible Heat Flux | SHFLUX | Real | IX x JX | W $m^{-2}$ | Cross | PBL =1,2,5 |
| Latent Heat Flux | LHFLUX | Real | IX x JX | W $m^{-2}$ | Cross | PBL =1,2,5 |
| Frictional Velocity | UST | Real | IX x JX | m $s^{-1}$ | Cross | PBL =1,2,5 |
| Surface Downward Shortwave Radiation | SWDOWN | Real | IX x JX | W $m^{-2}$ | Cross | |
| Surface Downward Longwave Radiation | LWDOWN | Real | IX x JX | W $m^{-2}$ | Cross | |
| Soil Temp. in Layers | SOIL T *n* | Real | IX x JX | K | Cross | IFSOIL = 1 |

# Appendix D. Sample Program to Read MM5 Version 2 Output

```
      PROGRAM RDMM5V2

C-----------------------------------------------------------------------
C     THIS IS A FORTRAN 90 PROGRAM TO READ MM5V2 OUTPUT.
C-----------------------------------------------------------------------

      IMPLICIT NONE

      REAL,          ALLOCATABLE   :: DUM2D      ( : , : )
      REAL,          ALLOCATABLE   :: DUM3D      ( : , : , : )
      REAL,          ALLOCATABLE   :: DUM3D2     ( : , : , : )
      INTEGER                      :: I
      INTEGER                      :: IMAX
      INTEGER                      :: INCR
      INTEGER                      :: INDEX
      INTEGER                      :: ISTAT
      INTEGER,       PARAMETER     :: IMM5IN     = 10
      INTEGER                      :: JMAX
      INTEGER                      :: KSIG
      INTEGER                      :: LOOP
      INTEGER                      :: MDOMEND
      INTEGER                      :: MDOMST
      INTEGER                      :: MIF        ( 1000, 20 )
      CHARACTER*80                 :: MIFC       ( 1000, 20 )
      INTEGER                      :: MRUNLEN
      REAL                         :: MRF        ( 1000, 20 )
      CHARACTER*80                 :: MRFC       ( 1000, 20 )
      INTEGER                      :: NUM2D
      INTEGER                      :: NUM3D
      REAL                         :: OUTITVL
      REAL                         :: SCALE
      CHARACTER*8                  :: VAR

C-----------------------------------------------------------------------
C     READ MM5 HEADER FOR THIS DOMAIN.  REWIND UNIT SO WE CAN START
C     FRESH ON THE LOOP.
C-----------------------------------------------------------------------

      VAR = 'HEADER'
      READ (IMM5IN, IOSTAT=ISTAT, ERR=8000, END=8000) MIF, MRF,
     &                                                MIFC, MRFC
      REWIND (IMM5IN)

C-----------------------------------------------------------------------
C     VERIFY THAT THIS IS MM5 OUTPUT.  IF NOT, STOP.
C-----------------------------------------------------------------------

      INDEX = MIF(1,1)
      IF ( INDEX /= 6 ) STOP 10

C-----------------------------------------------------------------------
C     EXTRACT DOMAIN ATTRIBUTES.
C-----------------------------------------------------------------------

      IMAX    = MIF(104,1)
```

```
      JMAX    = MIF(105,1)
      KSIG    = NINT(MRF(101,INDEX))
      OUTITVL = MRF(302,INDEX)
      IF ( NINT(OUTITVL) >= 60 ) THEN
        SCALE = 60.0
      ELSE
        SCALE = OUTITVL
      ENDIF
      INCR    = NINT( OUTITVL        / SCALE )
      MDOMST  = NINT( MRF(306,INDEX) / SCALE )
      MDOMEND = NINT( MRF(307,INDEX) / SCALE )
      MRUNLEN = NINT( MRF(308,INDEX) / SCALE )
      MDOMEND = AMIN0 ( MDOMEND, MRUNLEN )
      NUM2D   = MIF(202,INDEX)
      NUM3D   = MIF(201,INDEX)

C-----------------------------------------------------------------------
C     ALLOCATE NECESSARY VARIABLES.
C-----------------------------------------------------------------------

      ALLOCATE ( DUM2D  (IMAX, JMAX) )
      ALLOCATE ( DUM3D  (IMAX, JMAX, KSIG) )
      ALLOCATE ( DUM3D2 (IMAX, JMAX, KSIG+1) )

C-----------------------------------------------------------------------
C     PROCESS EACH TIME OUTPUT OF MODEL DATA
C-----------------------------------------------------------------------

      DO LOOP = MDOMST, MDOMEND, INCR

C-----------------------------------------------------------------------
C       READ HEADER.
C-----------------------------------------------------------------------

        VAR = 'HEADER'
        READ  (IMM5IN, IOSTAT=ISTAT, ERR=8000, END=8000) MIF, MRF,
     &                                                   MIFC, MRFC

C-----------------------------------------------------------------------
C       READ 3-D VARIABLES.
C-----------------------------------------------------------------------

        DO I = 1, NUM3D
          VAR = MIFC(204+I,INDEX)(1:8)
          IF ( VAR == 'W       ' ) THEN
            READ  (IMM5IN, IOSTAT=ISTAT, ERR=8000, END=8000) DUM3D2
          ELSE
            READ  (IMM5IN, IOSTAT=ISTAT, ERR=8000, END=8000) DUM3D
          ENDIF
        ENDDO

C-----------------------------------------------------------------------
C       READ 2-D VARIABLES.
C-----------------------------------------------------------------------

        DO I = 1, NUM2D
          VAR = MIFC(204+NUM3D+I,INDEX)(1:8)
          READ  (IMM5IN, IOSTAT=ISTAT, ERR=8000, END=8000) DUM2D
```

```
        ENDDO

      ENDDO

C-----------------------------------------------------------------------
C     DEALLOCATE VARIABLES.
C-----------------------------------------------------------------------

      DEALLOCATE ( DUM2D )
      DEALLOCATE ( DUM3D )
      DEALLOCATE ( DUM3D2 )

      STOP

C-----------------------------------------------------------------------
C     ERROR-HANDLING SECTION.
C-----------------------------------------------------------------------

 8000 WRITE(6,9000) IMM5IN, TRIM(VAR), ISTAT
      CALL ABORT
      STOP

 9000 FORMAT (/, 1X, 70('*'),
     &        /, 1X, '*** PROGRAM RWMM5',
     &        /, 1X, '***   ERROR READING MM OUT FILE, UNIT = ', I3,
     &        /, 1X, '***   VARIABLE = ', A,
     &        /, 1X, '***   IOSTAT = ', I4,
     &        /, 1X, 70('*'))

      END PROGRAM RDMM5V2
```

# Appendix E. EPA-Initiated Changes to MM5 Software

All Programs:

*configure.make*: Modified to allow machine-specific information to be passed from script, rather than set explicitly by user.

TERRAIN and TERRAIN_ws:

*crlnd.F*: Improve representation of urban areas along coastlines.

- Add after CRLND.75:

```
DIMENSION URBAN(IX,JX)
```

- Add after CRLND.127:

```
IF ( ILEV.EQ.1 ) THEN
  URBAN(II,JJ) = LNDOUT(II,JJ)
ENDIF
```

- Add after CRLND.133:

```
 IF ( ILEV.EQ.7 .AND. LNDOUT(II,JJ).LT.80.0 .AND.
&     URBAN(II,JJ).GE.10.0 ) GOTO 30
```

- Comment out CRLND.280 and replace with:

```
 IF ((NWATER.GE.NDOMINANT.AND.URBAN(I,J).LT.10.0)
&       .OR.(URBAN(I,J).GE.10.0.AND.NWATER.GE.40)) THEN
```

*interp.F*: Set terrain height over ocean to zero for 30-second terrain.

- Comment out INTERP.63-70
- Add after INTERP.70:

```
HTIN(KK,LL)=max(HTIN(KK,LL),0.)
```

*plots.F*: Standardize earth radius throughout CMAQ.

- In subroutine mpdrml, change parameter "re" to 6370.997.
- In subroutine xytoll, change parameter "re" to 6370.997.

*setup.F*: Standardize earth radius throughout CMAQ.

- In SETUP.56, change variable "A" to 6370.997.

DATAGRID:

*lbclat.F*: Standardize earth radius throughout CMAQ.

- In LBCLAT.23, set variable "AA" to 6370.997.

DATAGRID_ws:

All DATAGRID modifications, plus...

*rnavy.F*: Correct compile error.

- Change integer declaration of "NUMR" to data, as:

```
DATA NUMR /0/
```

*toga2dg.F*: Correct REWIND error.

- In 20 loop, comment out "`rewind iunit`" and replace with:

```
close (iunit)
```

RAWINS:

*lltoxy.F*: Standardize earth radius throughout CMAQ.

- In LLTOXY.9, change parameter "RE" to 6370.997.

*plots.F*: Standardize earth radius throughout CMAQ.

- In subroutine mpdrml, change parameter "re" to 6370.997.
- In subroutine xytoll, change parameter "re" to 6370.997.

*savstn.F*: Standardize earth radius throughout CMAQ.

- In SAVSTN.31, change variable "A" to 6370.997.

*sigdat.F*: Correct precision errors in "real" compare.

- Replace SIGDAT.143 with:

```
IF((NINT(CORRH(5,L,N))).LE.(NINT(GLVL(K)))) GO TO 65
```

RAWINS_ws:

All RAWINS modifications, plus...

*comadp.INCL:* Correct compile error.

- Remove `ifdef` surrounding definition of MWDSZ
- Set MWDSZ to 32.

*dadlib.F*: Correct compile error.

- In subroutine opendad, change integer declaration of "ICALL" to data, as:
  ```
  data icall /0/
  ```
- In subroutine opendad, remove ifdef surrounding definition of LENS, and set to 4.

*inacct.F*: Correct station types.

- Replace line INACCT.91 with:
  ```
  IF (SSTA .EQ. 'RIGG   1') THEN
  ```
- Replace line INACCT.97 with:
  ```
  IF (SSTA .EQ. 'PLAT   1') THEN
  ```
- Replace line INACCT.103 with:
  ```
  IF (SSTA .EQ. 'BUOY   1') THEN
  ```
- Replace line INACCT.109 with:
  ```
  IF (SSTA .EQ. 'SHIP   1') THEN
  ```

*swap.F*: Correct compile error.

- In subroutine filt, change declaration of m, n, l, and u from logical*1 to integer.

INTERP:

*interp.F*: Allow one-way nested processing on more than 100 output times.

- In INTERP.197, change loop maximum to 200.

*daddyo.incl*: Processing with large number of variables and output times.

- In DADDYO.3, change "LENTAB" to 10000 and "LTPERM" to 2000.

MM5:

*domain/initial/param.F*: Change KZ profile.

- Add after PARAM.1476:

```
C... AQM mods based on Seaman and Stauffer's
C FOR REGIME 1, ADJUST KZO: NLS/DRS SEPT 1991
      DO 5678 K=1,KL
      IF(K.EQ.1)THEN
      KZO(K)=(KZOC*DQCRIT)/DSIGMA(K)
      ELSE
      KZO(K)=(KZOC*DQCRIT)/(0.5*(DSIGMA(K)+DSIGMA(K-1)))
      ENDIF
      IF(KZO(K).GT.0.5)THEN
      KZO(K)=0.5
      ENDIF
5678  CONTINUE
      PRINT*,'******************** KZO *************************'
      DO 5680 K=1,MKX
5680  WRITE(6,5679)K,KZO(K)
      FORMAT(10X,'KZO(',I2,') = ',F5.2)
      PRINT*,'************************************************** '
```

*fdda/grid/in4dgd.F*: Allow analysis nudging with one-way nest.

- Replace IN4DGD.63 with:

```
      IUN1 = 200 + IN
#ifdef NO1WAY
      IUN1 = 10 + IN
#endif
```

- Add after IN4DGD.326:

```
#ifdef NO1WAY
```

- Add after IN4DGD.340:

```
#endif
```

*physics/cumulus/kf/kfpara.F*: Cumulus parameterization mods

- Uncomment 21APR97.328 and comment out 21APR97.329.

*physics/pbl_sfc/bulk/blkpbl.F*: KZO mods

- In BLKPBL.14, remove KZO from declaration list.
- Change BLKPBL.243 to:
  ```
  HFX(I,J)=AMAX1(HFX(I,J),-100.)
  ```
- Comment out BLKPBL.392
- In BLKPBL.394, change KZO to KZO(K)
- In BLKPBL.396, change KZO to KZO(K)

*physics/pbl_sfc/hirpbl/hirpbl.F*: KZO mods

- In HIRPBL.44, remove KZO from declaration list.
- Comment out HIRPBL.432
- In HIRPBL.434, change KZO to KZO(KL)
- In HIRPBL.436, change KZO to KZO(KL)
- Replace HIRPBL.450 with:
  ```
  PSIM(I)=-0.5*GZ1OZ0(I)
  ```
- Comment out HIRPBL.573
- Add after line HIRPBL.574
  ```
  HFX(I,J)=AMAX1(HFX(I,J),-100.)
  ```
- Comment out HIRPBL.575-578
- Comment out HIRPBL.580
- Comment out HIRPBL.635 and HIRPBL.637
- Comment out HIRPBL.712 and HIRPBL.714
- Comment out HIRPBL.856
- In HIRPBL.858, change KZO to KZO(K)
- In HIRPBL.860, change KZO to KZO(K)

*include/param3.incl*: KZO mods

- In PARAM3.12, add KZO to the end of the PARAM3 common block.
- Add after PARAM3.13:

```
      REAL KZOC, KZO(MKX)
```

- Add after PARAM3.19:

```
C Define critical KZO and DQ (DSIGMA): PBL MODS SEPT 1991
      DATA KZOC/0.1/
      DATA DQCRIT/0.05/
```