

# 10. Developing New CMAQ Simulations

---

For application users of CMAQ, the CMAQ model builder Bldmake is used only at the beginning of a simulation to compile executables for a specific science configuration. Since the horizontal grid and vertical layer structure are defined dynamically at execution of the model, there is typically no need to recompile the programs when changing these parameters. Compilation is required only when either developing CMAQ simulations for the first time or when the chemistry/science configuration within the model is changed.

A model developer would use Bldmake to check out a working version of the CMAQ source code and create a Makefile to facilitate the interchange of science components within a model, to modify reaction details within an existing chemistry mechanism, or to experiment with source code modifications.

The purpose of this chapter is to demonstrate how to build the executables for the CMAQ programs beyond running the installation test case. Before proceeding with this chapter, review [Chapter 5](#) for an overview of the system requirements for CMAQ. In general, there are three major steps in compiling CMAQ executables:

1. Install third-party libraries (netCDF, I/O API, MPI)
2. Compile Bldmake
3. Configure and build the executables for the various CMAQ

programs.

All compilations of libraries and CMAQ must be done with the *same compilers and settings*. The details of these three major steps with respect to creating new CMAQ simulations are covered in this chapter.

## General Introduction to Model Building

---

Before using CMAQ for operational modeling in a new computing environment, it is recommended that the model be benchmarked using the test dataset that is distributed with the model. [Chapter 5](#) describes how to install and benchmark CMAQ on a Linux system. After benchmarking CMAQ, it can be configured for other simulations. The same steps that are required to build the model for the test case apply to building it for new simulations. However, not all of the steps need to be repeated for a new model configuration unless new code becomes available or bug fixes are identified. In particular, it is not necessary to rebuild any of the libraries that CMAQ uses once working versions are built on a user's system. A single installation of the libraries (netCDF, I/O API, MPI) can be linked to for multiple configurations and applications of the model. Likewise, the CMAQ model builder, Bldmake, only needs to be compiled once and used for all applications of the model.

Except for MCIP, all of the CMAQ programs need to be recompiled when the chemistry mechanism or science configuration of the model change. If the science configuration does not change between applications, the CMAQ executables can be reused for different applications on the same Linux system. MCIP needs to be compiled only once on a user's Linux system and the executables may be reused for all applications of the model, unless new source code, including libraries, become available.

# Configuring New Simulations

---

The reason for modeling a particular time period evolves from a research question, such as determining why a severe air pollution episode happened, or studying the dominant sources of visibility degradation in a specific geographic region. Once a modeling episode is selected, several steps must be taken before CMAQ can be run.

1. Run and evaluate the WRF meteorology model for the episode of interest
2. Convert the WRF data for input to SMOKE and CMAQ using MCIP
3. Run and evaluate SMOKE to process emissions input data for CMAQ
4. Prepare biogenic emissions input data, either for inline or offline processing
5. Prepare initial and boundary conditions for CMAQ
6. Run and evaluate the CCTM.

The horizontal model grid, vertical layer structure, and model time periods must be consistent across the meteorology data, emissions data, and CMAQ.

Configuring CMAQ for new simulations involves defining the model grid, vertical layers, time periods, initial and boundary conditions, input/output file locations, and science options of the model. The following sections cover these topics in the context of setting up a new CMAQ simulation.

## Defining a new horizontal grid

The grid-dependent CMAQ programs (CCTM, ICON, BCON) use a GRIDDESC grid description file to define the map projection and horizontal grid for a simulation. The GRIDDESC file is either output by

MCIP or it can be created manually using a text editor. The CMAQ run scripts for the grid-dependent CMAQ programs must refer to a GRIDDESC file that contains the definition of the model grid to be simulated. The name of the grid for the current application must be specified in the CMAQ run script because a single GRIDDESC file can contain multiple grid definitions. Additional information about setting up horizontal grids in CMAQ is contained in [Chapter 9](#).

The following error from any of the CMAQ programs can occur if the name and/or location of the GRIDDESC file and/or the name of the horizontal grid are incorrect in the run script:

```
Failure defining horizontal domain
```

For CCTM, which uses both a GRIDDESC file and gridded input data (emissions, meteorology, ICs/BCs), the grid defined in the GRIDDESC file must be consistent across all of the gridded input files, or the simulation will fail.

To configure a new CMAQ simulation, the following steps must be taken to set up the horizontal model grid:

- Produce emissions and meteorology data on a consistent horizontal model grid to be modeled with CCTM
- Create a GRIDDESC file that defines the horizontal model grid
- Generate ICs and BCs on the horizontal model grid; for nested simulations, generate BCs from a parent grid
- Configure the CCTM script to use the GRIDDESC file and input data on the desired horizontal model grid

## Defining a new vertical layer structure

The CMAQ programs that produce 3-D output (CCTM, ICON, BCON) use

a MET\_CRO\_3D file to define a vertical layer structure. The MET\_CRO\_3D file is output from MCIP and takes the vertical layer structure from the input meteorology. The vertical layer structure must be consistent across all of the CMAQ programs used to complete a CCTM simulation. New vertical layer structures for CMAQ simulations are configured with MCIP when processing raw meteorological model data.

## **Setting a new episode time period**

The temporal extent of a CMAQ simulation is limited by the availability of input meteorology and emission data. Similar to the horizontal grid and vertical layer structures, the time period to model with CMAQ must be considered when designing the meteorological modeling simulation used to produce CMAQ input data.

The model output time step and run length of each CMAQ simulation are flexible and can be configured in the run script for the applicable CMAQ program. For CCTM, it is possible to have output files with a different number of time steps than the corresponding meteorology or emission input data. For example, it is common to have input meteorology files in 4- or 5-day intervals, input emission files in 1- or 2-day intervals, and CCTM output files in 1-day intervals. The CMAQ scripts allow the user to configure the model to output data with any number of time steps. The number of CCTM output time steps does not need to correspond with the input meteorology and emissions output time steps. To keep CMAQ output file sizes manageable, CMAQ output is typically stored in 1-day (24-hour) blocks.

To configure a new CMAQ simulation, the follow steps must be taken to set up the modeling time period:

- Produce emissions and meteorology data for the time period to be

modeled with CMAQ. When deciding upon a modeling period, it is necessary to have a “spin-up” interval prior to the beginning of the initial time of interest. The spin-up period is a sequence of days at the beginning of an air quality simulation that are not used in the analysis of the modeling results. These days are simulated to minimize the impacts of the initial conditions on the CCTM modeling results. Spin-up periods vary in length depending on the size of the modeling domain, the magnitude of the emissions sources within the modeling domain, and the pollutants being studied. As a general rule of thumb for regional modeling, a period of at least 10 days should be considered for a spin-up period.

- Generate ICs for the first time step of the CCTM simulation; if running multiple days in sequence, configure the CCTM run script to use the ICs from ICON for the first model day and to initialize the subsequent days with the previous day’s CCTM output
- Generate either time-independent BCs for the grid to be modeled, or for a nested simulation generate temporally resolved BCs for the model time period from CCTM outputs for the parent grid
- Configure the CCTM run script to loop through the days to be modeled, using the correct input files for each model day and writing output files with the desired number of time steps

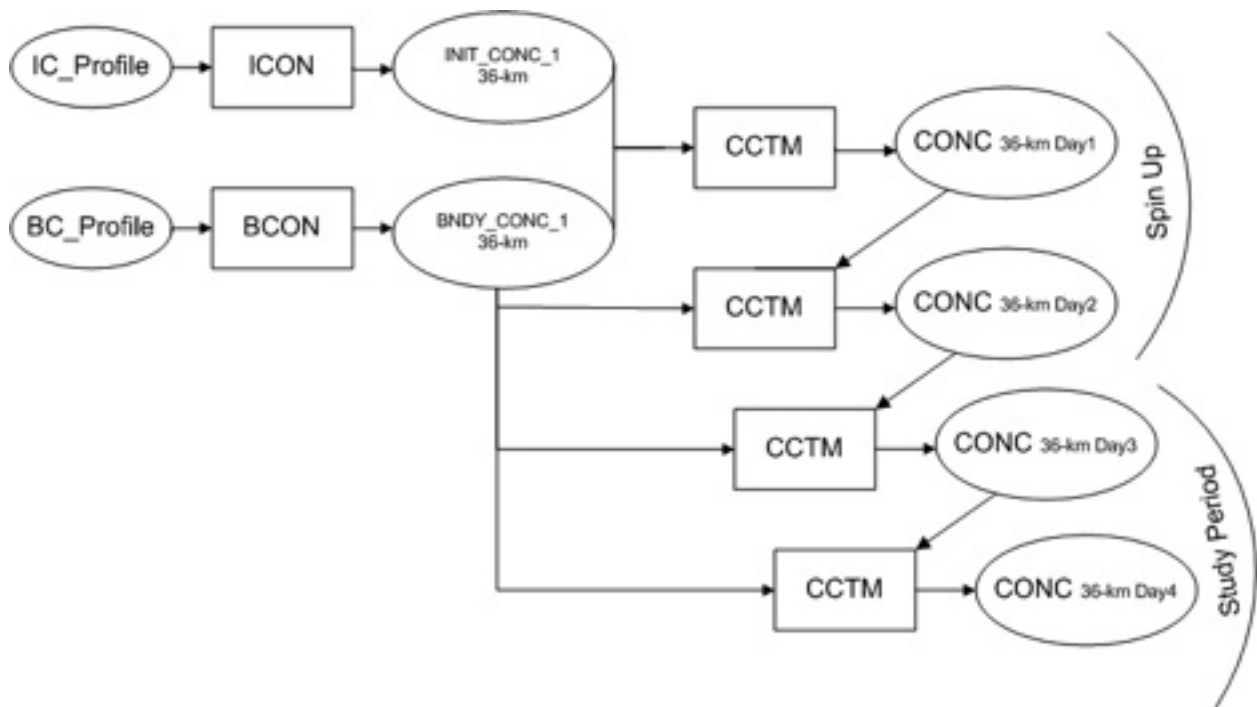
## **Initial and boundary conditions**

After preparing the meteorology and emissions input data files and determining the model grid and model time periods, the next step for setting up a new CMAQ simulation is creating initial and boundary conditions (ICs and BCs) for CCTM. The ICON processor provides initial chemical fields of individual species concentrations for a specific modeling domain. BCON provides concentrations of individual chemical species for the grid cells surrounding the modeling domain. ICON and

BCON both require two inputs: concentration values for the chemical species needed in the simulation, and a predefined chemical mechanism.

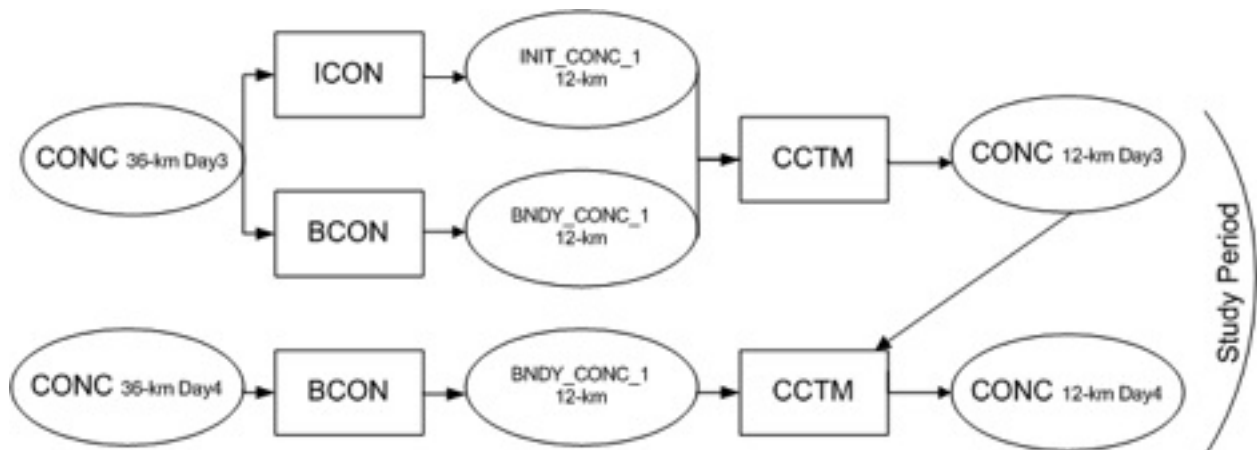
As described in Chapter 5, there are two types of input concentrations for ICON and BCON: either (1) tabulated tropospheric vertical profiles or (2) three-dimensional fields from a previous CMAQ or larger-scale CTM simulation, such as GEOS-Chem(2009) or MOZART (2009). The input file type to use for ICON and BCON depends on whether the simulation is a restart of a previous run and/or whether the simulation is a nest of a larger parent grid. The same chemical mechanism must be selected for ICON, BCON, and CCTM. Both ICON and BCON assume that the input species concentrations are for the selected mechanism. Refer to Chapter 5 for information on how to configure ICON and BCON for the two input file types.

Standard operation of CMAQ uses boundary conditions that are extracted from a global model, such as GEOS-Chem or MOZART for the outermost modeling domain. Nested grids use temporally resolved boundary conditions extracted from the results of a parent-grid CCTM simulation. The initial conditions produced by ICON are time independent for the first spin-up day. If the initial conditions were generated from vertical profile data then they will also be spatially uniform, meaning that for a given layer they will contain the same concentration in every grid cell. The remaining days of a simulation should use spatially heterogeneous output from the previous day's CCTM simulation to initialize each day. See [Figure 10-1](#) for a schematic of the initial and boundary conditions used for a CCTM simulation with a two-day spin-up followed by a two-day period of interest. In this example, each of the days was run separately.



**Figure 10-1. 36-km four-day modeling period IC/BC schematic**

When using a nested-grid configuration, the fine-grid-resolution grids can use time-varying boundary conditions generated by BCON with input from time-varying output concentrations from the coarse-grid CCTM simulation. The initial conditions for start-up of the fine grid are also generated using concentrations from the coarse grid. Subsequent runs in the period of interest use the last hour of the concentration file generated from the previous day's run. See [Figure 10-2](#) for an example of a one-way, nested simulation. This example uses initial and boundary conditions from the coarser 36-km grid simulation to perform the nested, finer-grid simulation for the same two-day period of interest.



**Figure 10-2. 12-km nested two-day modeling period IC/BC schematic**



# Input/output file names and locations

Configuring the CMAQ run scripts for a new simulation involves creating an application identifier to use in the name of the CMAQ outputs and to specify the correct input and output file names and locations on the system where CMAQ will be run. Application identifiers are selected to uniquely identify the output files with respect to the simulation. For example if you are running a base simulation for the year 2007, a simulation identifier, or APPL setting in CMAQ terms, could be Base07. For time-dependent output files, a date identifier is commonly added to the file name to identify the time period covered by the file. Following the previous example, if you configured a simulation to run on January 5, 2007 (Julian date 2007005), an example file name for a CMAQ version 5.2 CCTM concentration file would be CCTMv52.CONC.Base07.2007005.ncf. Additional identifying information, such as the chemistry mechanism name, versioning identifiers for the input meteorology and emissions, and the operating system version (i.e. Linux2\_x86\_64) are also commonly added to CMAQ output file names.

Before starting a new CMAQ simulation, the user needs to confirm the existence of the input files for all of the CMAQ programs, and to check the names of the output files and directory locations. A common error that occurs with all of the CMAQ programs is an execution failure caused by missing or misnamed input files.

## Science option configuration

The CMAQ scripts as they are distributed use a default science option configuration. While this default configuration is acceptable for all applications of the model, it is not the only configuration that can be used with CMAQ. The ICON and BCON science options depend on the

nature of the input data, the chemical mechanism chosen, and whether one-way nests are being used. The CCTM science configuration presents the largest number of combinations of different transport algorithms, chemistry options, and special features, such as process analysis. To see a choice of all the different options available for configuring CCTM, refer to [Chapter 7](#).

Not all of the combinations of the various CCTM configuration options have been tested. It is possible that some combinations of different science modules will not build a working executable. Generally, here are few rules to follow when configuring CCTM:

- For configurations that use the Yamartino model driver, the Yamartino options must be used for the initialization module, the advection routines, and the concentration adjustment scheme.
- Ensure that there is consistency in the selection of the chemistry mechanism and the aerosol module; for example, the *aero6* aerosol module cannot be used with a chemistry mechanism that is tagged “ae5”.
- The EBI chemistry solver is mechanism-dependent and must be consistent with the chemistry mechanism; for example, the EBI solver for CB05 cannot be used with a SAPRC-07 chemistry mechanism.

The availability of different science options in CMAQ creates the flexibility to select a configuration that optimizes the model performance for different applications. Through testing the various science configurations with respect to model accuracy and speed, the CMAQ user can find the combination of science options that gives the best performance for a particular modeling study.

## References for Chapter 10: New

# Simulations

---

GEOS-CHEM (2009) [http://wiki.seas.harvard.edu/geos-chem/index.php/Main\\_Page](http://wiki.seas.harvard.edu/geos-chem/index.php/Main_Page)

MOZART (2009),  
<http://www.mpimet.mpg.de/en/wissenschaft/modelle/mozart.html>

---

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

CMAQ Operational Guidance Document (c) 2016