AERMOD Helper Files for the CMV run group

Introduction

The SMOKE4AERMOD postprocessing scripts
(https://github.com/CEMPD/SMOKE/tree/master/scripts/aermod) create "helper" files for the port and underway commercial marine vessel (CMV) run group. These helper files are used by the AERMOD modeler[1] to create AERMOD inputs for each source in the run group. This document describes the format of the helper files for the CMV run group, and their content based on the 2014 National Air Toxics Assessment (NATA) application of the SMOKE4AERMOD tool. As with all AERMOD file preparation for NATA, helper files contain only those NEI sources with HAP or DIESEL-PM10 emissions and no sources with FIPS 85 (Federal waters).

The "port and underway CMV" run group is comprised of sources within the NEI NONPOINT data category that have SCC codes of 2280002200 (c1/c2 underway), 2280002100 (c1/c2 ports), 2280003200 (c3 underway), 2280003100 (c3 ports) and that reside in US **State** waters, i.e., any US State FIPS including DC, Puerto Rico (72) and the Virgin Islands (78). Therefore, FIPS 85 and non-US emissions ARE NOT included in this run group.

In the NEI, CMV emissions are inventoried as GIS-based shapes that reflect locations of CMV hoteling or maneuvering at ports or CMV cruising or reduce speed zones while underway. SMOKE gridding is **not** needed for these CMV sources because locational information is linked to the FF10 field "SHAPE_ID" which references a GIS shape that has been simplified into one or more a polygon sources for input into AERMOD. For ports the NEI "SHAPE_ID" represents a single simplified polygon. For underway, a set of simplified shapes consisting of multiple polygons per NEI shape is used; emissions from the NEI SHAPE_ID is apportioned to these polygons based on area of the polygon.

Each shape in the inventory is associated with a single county (shapes do not cross counties). For underway shapes, previous NATA modeling created simplified shapes (based on sub-dividing NEI shapes and simplifying them into polygons) and for port shapes, the 2014 NEI contains the already simplified shapes that are described by polygon vertices.

The "BEFORE-AERMOD" helper files consist of the following 3 types: (1) Location, (2) Area_Parameter, and (3) Temporal. The helper files contain all sources for that run group. The "AFTER AERMOD" helper file contains pollutant-specific emissions by source. The "before AERMOD" files do not provide the unit emissions rate or divide by the source area. The AERMOD modeler[1] will apply the 1000 g/s and divide by area where appropriate in his post processing step.

---

[1] James Thurman, Thurman.james@epa.gov, is the AERMOD modeler for the 2014 National Air Toxics Assessment

The POLYGON file – A key input to the CMV Emissions Processing

Unlike the other run groups, CMV emissions processing uses a key non-FF10 ancillary file containing the port and underway polygon/shape information: number of vertices, latitude & longitude of each vertex, relevant id to relate the polygons to the NEI shapes/polygons and the polygon areas to apportion emissions from NEI shapes that are modeled as multiple polygons. NEI Underway shapes may have multiple polygons to be used in modeling. NEI port shapes are identical to the polygons used in the modeling such that each NEI shape is a single AERMOD polygon. The "POLYGON" file contains both underway and port shape/polygon information.

The POLYGON file for the 2014 NATA v2 is:
CMV_polygon_combined_update20161209_31jan2018_31jan2018_v0

Below is the format of the POLYGON file

Field Name	Description
FACID	Facility id – basically an ID for the polygon that has the FIPS in it and that can be associated with and NEI Port or Underway Shape. For underway: concatenation of 6 strings as follows: "U" , NEI Shape id, "F", 5-digit FIPS, "P", source #, where source number begins at 001 and increases by 1 for each polygon with in the Shape. E.g., U0118F02013P001 For port: concatenation of 4 strings as follows: "P", NEI ShapeID, "F", 5-digit FIPS
LON	Longitude of first polygon vertex for shape
LAT	Latitude of first polygon vertex for shape
NUMVERT	Number of vertices for the polygon
AREA	Area of the polygon in square meters

"BEFORE-AERMOD" HELPER FILES
File naming conventions

The names of the helper files are:

File type	NAME	Description
location	CMV_locations.csv	Provides one shape vertex and cross walk between the polygon and the 12km CMAQ grid cell (needed for sources in the CONUS in order to assign correct gridded MET data)
Area source parameters	CMV_area_params.csv	Release height, σ_z, all vertices of the simplified shapes
Temporal factors	CMV_temporal.csv	Temporal factors – for CMV these are just monthly factors and they depend on whether the source is c1/c2 versus c3 marine vessels.

Source and naming conventions

A source is a unique combination of polygon-shape/vessel type. The polygon could be a sub-shape within an NEI shape or its own shape. Each unique NEI shape is a separate facility. The number of sources at the facility depends on the number of polygons for the shape and whether there are both c1/c2 and c3 emissions for the NEI shape.

Source Naming Convention

The Facid in the POLYGON file is used for constructing the source id and is comprised of the NEI shape id, the polygon, the type of CMV, the FIPS.

For port: "P" concatenated with the NEI shape id, concatenated with "F" concatenated with the 5-digit FIPS

for underway: "U" concatenated with the NEI shape id (put in leading zeroes such that the shape id is always 4 characters in length), concatenated with "F" concatenated with the 5-digit FIPS concatenated with an "P" concatenated with a number

The polygon source id is constructed from the Facid as follows:

For ports: use all characters of the Facid up until (and excluding) the "F", appended with "c1" or "c3" for sources for which c1/c2 or c3 emissions occur at the port, respectively.

for underway: All characters of the Facid up until (and excluding) the "F", the end of the Facid beginning with (and including) the "P", appended with "c1" or "c3" for sources for which c1/c2 or c3 emissions occur in the underway shape, respectively.

 Note that the first part will be 5 or 6 characters at the most, and the 2nd part will be 6 characters at the most. So the AERMOD source id will be 12 characters at the most.

Location files

 ONLY SOURCES with emissions (HAP , diesel PM) should be included in this FILE. Each polygon that has emissions associated with it needs to be in the ancillary file that has the vertices. The Location files also contain the col row of the grid cell that the (polygon) fits into. If the polygon is in multiple grid cells then use the grid cell that contains most of the area of the shape. If the polygon is so large that it encompases multiple full grid cells, then choose the grid cell it most represents.

IMPORTANT QA STEP: There should be no FF10 shape ids (with nonzero HAP and/or DIESEL-PM10 emissions) that are not covered in the polygon vertices files. Every source in the Location files need to be in the Area Source parameter files unless it has no emissions. Any source in the location files that has no emissions should be listed on the QA spreadsheet.

Table 1 Fields for CMV Location Helper Files

Filename **CMV_ locations.csv**, 1 row per source id polygon/ship type (c1c2 vs c3) combination (i.e., one row per source)	
Notes: CONUS and nonCONUS sources can be run together	
fieldname	**Description**

State 2 digit FIPS	Meta data, not used, based on the FIPS which is part of the FacId (5 digits after "F")
FIPS	Meta data, not used
Facid	This is basically an id for the shape that includes the FIPS in it. Because it is too long for the AERMOD source id (except for port facid and some of the newest underway polygons), it isn't used other than as extra meta data.
Source id	This is the ID for the polygon shape and is unique for each polygon For ports: use all characters of the Facid up until (and excluding) the "F" , appended with "c1" or "c3" for sources for which c1/c2 or c3 emissions occur at the port, respectively. for underway: All characters of the Facid up until (but excluding) the "F", concatenated with the end of the Facid beginning with (and including) the "P" , appended with "c1" or "c3" for sources for which c1/c2 or c3 emissions occur at the port, respectively. Note that the first part will be 5 or 6 characters at the most, and the 2nd part will be 4 characters at the most
Source Type	AREAPOLY
COL ROW	CMAQ grid cell which contains most of the area of the polgyon shape. If the polygon shape >>> bigger than grid cell then choose most representative grid cell. LEADING ZERO is fine. For nonCONUS shapes, a "pseudo" grid cell based on the preponderance of vertices of the polygon is used.
x-coord- (UTM)	UTM x for one vertex of the polygon shape based on the lat lon in the polygon vertices file and the UTM zone
y-coord- (UTM)	UTM y for one vertex of the polygon shape
UTM zone	Zone of UTM – use the centroid of the 12km CMAQ grid cell for the source. (note that the rest of SMOKE-AERMOD uses the SW corner of the grid cell). For non-CONUS polygons, use the centroid of the 12km "pseudo" grid cell based on the preponderance of vertices of the polygon
lon	Decimal degrees longitude of 1st vertex of polygon shape
lat	Decimal degrees latitude of 1st vertex of polygon shape

Suggested QA procedures for the CMV_locations.csv file
    1.   Ensure each source id is unique.
    2.   Count the number of unique source ids and provide in QA output file
    3.   Make sure each polygon is in a single UTM zone
    4. List sources that have no emissions.	

AREA Source Parameter Files (RELEASE PARAMETER FILES AND VERTICES)

Table 2 Fields for Area Source Parameter Helper Files for CMV Sources (Port or Underway)

Filename **CMV_area_parameters.csv**	
1 row per source id and vertex.	
field	**description**
State 2 digit FIPS	

Facid	value in the polygon vertices and emission allocation files, this is basically an id for the polygon that includes the FIPS in it. Because it is too long for the AERMOD source id(except for port facid and some of the newest underway polygons), it isn't used other than as extra meta data.
Source id	This is the ID for the polygon shape and is unique for each polygon See location helper file for how it is constructed.
SOURCE TYPE	AREAPOLY
area	area of the polygon in square meters from the POLYGON file.
RELEASE HEIGHT (M)	C1/c2 Sources – i.e., those that have c1 appended to them have a release height of 8.4 meters. C3 have the same release height we are using in CMAQ which is 65.62 ft (20 m) for all C3 sources.
Number of vertices	Same as in v1
SZ (M)	C1/c2 Sources – SZ is RH/2.15 C3 – The sigmaz is developed to account for the plume rise that occurs in CMAQ . For every county, we found that 50% of the emissions were above layer 3 and 50% were below. layer 3 is 60.7 meters. Therefore sigmaz is 60.7-20=40.7 m for all c3 sources.
UTM_X coord-i	VERTICES are in any order, but the first one must be the same one listed in the Location file.
UTM_y coord-i	VERTICES are in any order, but the first one must be the same one listed in the Location file.
Lon coord-i	Longitude associated with the UTM coordinate i, decimal degrees
Lat coord-i	Latitude associated with the UTM coordinate i, decimal degrees
Suggested QA procedures for the CMV_area_parameters.csv file 1. Check that this file and the location file has exactly the same source ids in them.	

TEMPORAL Factor files by source id

Table 3 Fields for CMV Temporal (monthly) Helper Files

Filename **CMV_temporal.csv** One row per polygon For CMV we are using uniform day of week and diurnal profiles. Monthly temporal profiles can be county-specific, and for the 2014 Platform used for the NATA modeling, they are county-specific profiles for c1/c2 sources on the Great Lakes. Generally, C1/c2 sources use a different profile than c3 (which is profile 19531).	
Field name	**description**
State 2 digit FIPS	
Facid	See location helper file.
Source id	Same as in location helper file
qflag	MONTH. For CMV we are using uniform day of week and diurnal profiles

Monthly scalar	For qflag=MONTH, N=12. The scalars equal the monthly temporal profile factors (after normalization) in the PTPRO SMOKE temporal profile file AND are based on the SCC AND the county. Scalar1=January, Scalar2=February, Scalar3=March…Scalar12=December **Note that the AERMOD modeler will compute hourly emissions as follows:** Compute hourly emissions for each hour of month, i as: $$Hourly\ emissions_i = Annual \times \frac{Scalar_i}{\sum_{i=1}^{12} Scalar_i \times \#days\_in\_month_i} \times \frac{1}{24}$$

Suggested QA procedures for the CMV_temporal.csv file
1. Sum of scalars is 1
2. Check that this file and the location file has exactly the same polygons (source id)

AFTER-AERMOD Emissions files

For the "After AERMOD" (part 2), SMOKE should produce pollutant specific emissions by polygon source id (separately for underway and port polygons) and SMOKE name.

For CMV, we are creating annual chi/q so the emissions are annual.

Table 4 Fields for CMV EMISSIONS Helper Files

Filename **CMV_emis.csv**	
One row per SMOKE name pollutant and source	
Field name	**description**
State abbrev	
Run group	Port or underway
Facid	value in the polygon vertices and emission allocation files, this is basically an id for the polygon that includes the FIPS in it. Because it is too long for the AERMOD source id (except for port facid and some of the newest underway polygons), it isn't used other than as extra meta data.
Source id	Same as in location helper file
Source group	For underway, the source group is "underway" irrespective of the **Source id** For port the source group is either "c1c2port" or "c3port" depending on the **Source id**
pollutant name	Use SMOKEshortname in INVTABLE (should include only HAPs and diesel PM10)
Emissions (tons)	• Port for **Source id**'s that end in c1: Emissions from the FF10 summed by SCC 2280002100 and shape • Port for **Source id**'s that end in c3: Emissions from the FF10 summed by SCC 2280003100 and shape • Underway for sources that end in c1: Emissions from the FF10 summed by SCC 2280002200 and shape, multiplied by area of polygon/area of shape. Use area of polygon/area of shape data from the emission allocation files for underway.

	• Underway for **Source id**'s that end in c3: Emissions from the FF10 summed by SCC 2280003200 and shape, multiplied by area of polygon/area of shape. Use area of polygon/area of shape data from the mission allocation files for underway.

Suggested QA procedures for the CMV_emis.csv file

1. The Metal/CN speciation factor should be 1 for both port and underway. This should be checked.
2. Check all facids and source ids that are in the locations file are in this file
3. Sum emissions by county and they should match FF10 for NEI v1 where 2280002100 and 2280003100 are aggregated to "port"; 2280002200 and 2280003200 are aggregated to "underway".