Appendix B - Emission Inventory Development for the California Air Toxics Assessment (CATA)

CEPAM 2019 v1.01

Triennial Modeling

Second Iteration – Modeling Year 2017

CALIFORNIA
AIR RESOURCES BOARD

Table of Contents

Table of Figures

List of Tables

I. Development of Toxics Emission Inventories

The California Air Toxics Program establishes the process for the identification and control of toxic air contaminants (TACs). Air quality models (AQMs) can inform policy makers of the progress in the mitigation or potential public exposure to TACs. The air quality models require model-specific formatted emission inventory inputs often referred to as Toxics Emission Inventories.

There are two AQMs of interest at the California Air Resources Board, the dispersion model CALPUFF and the Community Multiscale Air Quality (CMAQ) photochemical model. The following sections of chapter 1 provide a brief description of how toxics emission inventories are developed, while chapters 2 and 3 provide more detail on explicit spatial, temporal, and chemical extent of toxic inventory development for CALPUFF and CMAQ, respectively.

II. Background

The Office of Community Air Protection Program (OCAPP) was created according to the regulations specified in Assembly Bill 617 (Garcia, 2017) with a goal of providing environmental justice to impacted communities in the State of California that suffer from toxic air pollutants. Community level air dispersion modeling coupled with regional photochemical modeling have been identified as tools that can help with the selection of communities, health risk assessments and/or the progress expected in those communities (Office of Air Quality Planning and Standards, 2002). These tools require model-specific emission inventory inputs that incorporate both anthropogenic and biogenic sources.

III. Description of CALPUFF dispersion modeling

CALPUFF is a Lagrangian, multi-layer, multi-species non-steady-state puff dispersion model that simulates the effects of time- and space-varying meteorological conditions on the transport, transformation, and removal of ambient pollutants (U.S. E.P.A., Air Quality Dispersion Modeling - Preferred and Recommended Models, 2017). It includes algorithms for sub grid scale effects, such as terrain impingement, and longer-range effects such as pollutant removal due to wet scavenging and dry deposition, simple chemical transformation, and visibility effects of particulate matter concentrations. CALPUFF is particularly useful for inert pollutants such as diesel particulate matter (DPM). It can process stationary, line, and area sources; as a result multiple 1km regional subdomains that cover specific counties are used to simulate the community level at a higher resolution than can be achieved with regional modeling.

IV. Description of CMAQ photochemical modeling

CMAQ is a numerical air quality model that simulates emissions, chemistry, and physical processes in the atmosphere. CMAQ uses mathematical representations of coupled atmospheric dynamics, chemical, and physical processes to describe the fate of air pollutants. The model is formulated to conserve mass in the 3-D atmosphere within a specific modeling domain (U.S. E.P.A., CMAQ User's Guide, 2020). The resultant partial differential governing equations are numerically solved over a 3-D grid representing the geographic area of interest (U.S. E.P.A., CMAQ User's Guide, 2020). CMAQ belongs to the Eulerian class of mathematical models that calculate a mass balance over each discrete grid volume by accounting for transport across the boundaries of the grid volume and relevant source and sink terms within the grid volume over a given period of time. CMAQ requires two primary types of inputs: meteorological information and emission rates from both anthropogenic and biogenic sources. To obtain emission inputs CMAQ relies heavily on an emissions processor to estimate the magnitude, location, speciation, and temporal variability of pollution sources. The Sparse Matrix Operator Kernel Emissions (SMOKE) processor is used by the California Air Resources Board (CARB) for computing emissions inputs for CMAQ from planning emissions inventories. These emission inputs must be specified on CMAQ's horizontal and vertical grid structure and cover the time period of the air quality model simulation. The emission inputs must also represent chemical species that conform with the gas and aerosol chemical mechanism employed in the CMAQ configuration. CARB utilizes the Statewide Air Pollution Research Center (SAPRC) chemical mechanism and therefore total organic gas (TOG) for each emission source is speciated to represent a specific SAPRC model species.

V. Spatial Extent of Emission Inventories

To simulate toxic air contaminants within communities, model ready emission inventories are prepared for different regions throughout California. Table 1 shows a summary of the parameters for different modeling domains, while Figure 1 shows their corresponding spatial extents. All modeling domains use the Lambert projection system, which assumes a spherical Earth. This projection system uses a Lambert Conical Projection with two parallels i.e. at 30° and 60° N latitude, and with a central meridian at 120.5° W longitude. The coordinate system origin is offset to 37° N latitude. There are two different grid resolutions used for toxics modeling and they differ based on the AQM. A single 2-km resolution California Statewide domain is used for CMAQ, which has a high computational burden because of its detailed chemistry. Six region specific 1-km resolution domains are used for CALPUFF, which focuses on inert pollutants such as diesel particulate matter (DPM). The multiple 1-km regional subdomains cover highly populated areas within California and allow us to simulate community level emissions at a higher resolution than CMAQ.

Table 1: Toxic Modeling Domain Parameters

2-K Grid Domain Definition

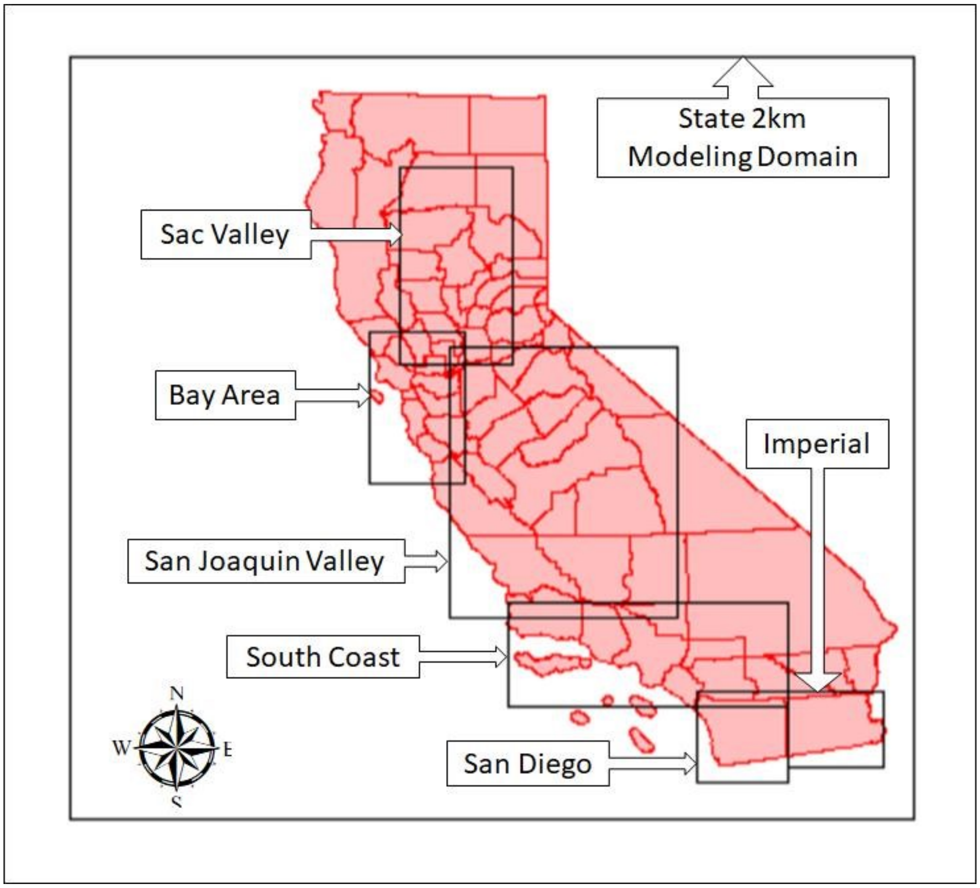
Name	**X_origin**	**Y_origin**	**x_size**	**y_size**	**x_extent**	**y_extent**
'CA_State2k'	-684000	-564000	2000	2000	642	582

1-K Grid Subdomain Definition

Name	X_origin	Y_origin	x_size	y_size	x_extent	y_extent
'SJV_348x412'	-108000	-256000	1000	1000	348	412
'SAC_172x300'	-184000	132000	1000	1000	172	300
'IMP_148x116'	408000	-484000	1000	1000	148	116
'SDC_136x149'	272000	-508000	1000	1000	136	149

Name	X_origin	Y_origin	x_size	y_size	x_extent	y_extent
'SC_424x160'	-16000	-392000	1000	1000	424	160
'BA_144x232'	-228000	-52000	1000	1000	144	232

Figure 1: Different Modeling Domains used for Toxics Modeling



VI. Inventory Year and Terminology

Current toxics modeling is based on version 1.01 of the California Emission Projection Analysis Model (CEPAMv1.01) also tracked as rf3043-snap20190508. Previous inventory versions have been archived and can be referenced in past protocols (CARB A. , Toxics Emission Inventory Protocol, 2019). As part of the triennial modeling for California toxics assessments, the current inventory is the second iteration and established on a 2017 base year with emission estimates for each county, air basin and district (Coabdis). The Coabdis is interchangeably referred to as a geographic area of interest (GAI) which is a numerical code that corresponds to unique Coabdis regions.

The emission inventory is processed in multiple sectors, which are merged depending on type of AQM used. For CMAQ, all gridded sources are combined into a 2-km grid cell, while stationary sources are processed in line. For CALPUFF, sources are split by sectors that contribute to toxic air contaminants and their spatial extent (line, polygon, point and gridded). Table 2 describes each of the individual sectors and their spatial extent that are processed for both CMAQ and CALPUFF modeling.

Table 2. Types of emission sources and spatial processing

Inline or Point	Gridded	Line / Polygon
OGV at Berth Stationary Criteria Stationary Toxics	Agburn Agriculture Military Aircraft Charbroil Commercial Harbor Craft Construction Fertilizer Livestock Locomotive Managed Burns Ocean Going Vessels (Military) Others Onroad (CMAQ only) Paved Road Dust Unpaved Road Dust Residential Wood Combustion Structural Burns TRU Wildfires Prescribed Burns Mexico (Area, Nonroad, Point and Onroad)	Onroad (CALPUFF only) OGV Maneuvering OGV Transit Aircraft 3D

VII. Toxic Air Contaminants Inventory

In addition to criteria pollutant-species, three types of toxic air contaminants (TACs) are considered: 1) DPM, 2) toxic heavy metals, and 3) toxic VOCs. These TACs can be classified as non-reactive species (DPM and heavy metals) and reactive species (VOCs). The toxic VOCs are determined from total organic gas (TOG) through speciation of CEPAMv1.01 estimates. Table 3 illustrates a list of specific TACs or precursors of secondary pollutants, such as O_3, that are simulated in CMAQ and/or CALPUFF. Table 4 summarizes the type of TAC, emission category and model selection.

Table 3. List of Modeled Species

Pollutant	Pollutant Description	Type
AACD	Acetic Acid	VOC
ACET	Acetone	VOC
ACRO	Acrolein	VOC
ACRY	Acrylonitrile	VOC
ACYE	Acetylene	VOC
ALK1	Alkanes that react only with OH, and have kOH between 2 and 5 x 10^2 ppm^{-1} min^{-1}	VOC
ALK2	Alkanes that react only with OH, and have kOH between 5 x 10^2 and 2.5 x 10^3 ppm^{-1} min^{-1}	VOC

Pollutant	Pollutant Description	Type
ALK3	Alkanes that react only with OH, and have kOH between 2.5 x 10^3 and 5 x 10^3 ppm^{-1} min^{-1}	VOC
ALK4	Alkanes that react only with OH, and have kOH between 5 x 10^3 and 1 x 10^4 ppm^{-1} min^{-1}	VOC
ALK5	Alkanes that react only with OH, and have kOH greater than 1 x 10^4 ppm^{-1} min^{-1}	VOC
APIN	a-Pinene	VOC
ARO1	Aromatics with kOH less than 2x10^4 ppm^{-1} min^{-1}	
ARO2	Aromatics with kOH greater than 2x10^4 ppm^{-1} min^{-1}	VOC
B124	1,2,4-Trimethylbenzene	VOC
BACL	Biacetyl	
BALD	Aromatic aldehydes (e.g., benzaldehyde)	VOC
BDE13	1,3-Butadiene	
BENZ	Benzene	VOC
CCHO	Acetaldehyde	VOC
CCL4	Carbon tetrachloride	VOC
CH4	Methane	
CO	Carbon Monoxide	
CRES	Phenols and Cresols	VOC
CRVI	Chromium, hexavalent (& compounds)	Heavy Metal
DPM	Diesel PM (total)	PM
DPM10	Diesel PM_{10}	PM
DPM25	Diesel $PM_{2.5}$	PM
ETHE	Ethene	VOC
ETOH	Ethanol	VOC
FACD	Formic Acid	VOC
GLY	Glyoxal	VOC
HCHO	Formaldehyde	VOC
IPRD	Lumped isoprene product species	VOC
ISOP	Isoprene	VOC
MACR	Methacrolein	VOC
MECL	Methylene chloride	VOC
MEK	Ketones which react with OH radicals faster than 5 x 10^{-13} but slower than 5 x 10^{-12} cm^3 $molec^{-2}$ sec^{-1}	VOC
MEOH	Methanol	VOC
MGLY	Methyl Glyoxal	VOC
MVK	Methyl Vinyl Ketone	VOC
MXYL	m-Xylene	VOC
NH3	Ammonia	
NOX	Nitrogen Oxides	
OLE1	Alkenes (other than ethene) with kOH less than 7x10^4 ppm^{-1} min^{-1}	VOC
OLE2	Alkenes with kOH greater than 7x10^4 ppm^{-1} min^{-1}	VOC
OXYL	o-Xylene	

Pollutant	Pollutant Description	Type
PACD	Higher organic acids and peroxy acids	
PAHO*	PAHs, total, with individual components also reported	
PAHW*	PAHs, total, w/o individual components reported	
PAS	$PM_{2.5}$ Arsenic	Heavy Metal
PCD	$PM_{2.5}$ Cadmium	Heavy Metal
PDCB	p-Dichlorobenzene	
PEC	$PM_{2.5}$ EC	PM
PERC	Perchloroethylene	VOC
PM	Total Particulate Matter	PM
PM10	Particulate matter up to 10 micrometers	PM
PM25	Particulate matter up to 2.5 micrometers	PM
PMC	PM coarse	PM
PMOTHR	$PM_{2.5}$ Others	PM
PNCOM	PM2.5 Non-Carbon organic matter	PM
PNI	$PM_{2.5}$ Nickel	Heavy Metal
PNO3	$PM_{2.5}$ Nitrate	Heavy Metal
POC	$PM_{2.5}$ OC	PM
PPB	$PM_{2.5}$ Lead	Heavy Metal
PRD2	Ketones and other non-aldehyde oxygenated products which react with OH radicals faster than 5×10^{-12} cm^3 $molec^{-2}$ sec^{-1}.	VOC
PRPE	Propene	VOC
PSO4	$PM_{2.5}$ Sulfate	PM
PXYL	p-Xylene	VOC
RCHO	Lumped $C3^+$ Aldehydes	VOC
RNO3 +	Lumped Organic Nitrates	
ROG	Reactive organic gas	
SESQ	Sesquiterpenes	VOC
SOX	Sulfur Oxides	
TERP	Terpenes	
TOLU	Toluene	VOC
ETCL3	Trichloroethylene	VOC
VICL	Vinyl chloride	VOC
AACD	Acetic Acid	VOC
ACET	Acetone	VOC
ACRO	Acrolein	VOC
ACRY	Acrylonitrile	VOC
ACYE	Acetylene	VOC
ALK1	Alkanes that react only with OH, and have kOH between 2 and 5×10^2 ppm^{-1} min^{-1}	VOC
ALK2	Alkanes that react only with OH, and have kOH between 5×10^2 and 2.5×10^3 ppm^{-1} min^{-1}	VOC
ALK3	Alkanes that react only with OH, and have kOH between 2.5×10^3 and 5×10^3 ppm^{-1} min^{-1}	VOC

Pollutant	Pollutant Description	Type
ALK4	Alkanes that react only with OH, and have kOH between 5×10^3 and 1×10^4 ppm^{-1} min^{-1}	VOC
ALK5	Alkanes that react only with OH, and have kOH greater than 1×10^4 ppm^{-1} min^{-1}	VOC
APIN	a-Pinene	VOC
ARO1	Aromatics with kOH less than $2x10^4$ ppm^{-1} min^{-1}	
ARO2	Aromatics with kOH greater than $2x10^4$ ppm^{-1} min^{-1}	VOC
B124	1,2,4-Trimethylbenzene	VOC
BACL	Biacetyl	
BALD	Aromatic aldehydes (e.g., benzaldehyde)	VOC
BDE13	1,3-Butadiene	
BENZ	Benzene	VOC
CCHO	Acetaldehyde	VOC
CCL4	Carbon tetrachloride	VOC
CH4	Methane	
CO	Carbon Monoxide	
CRES	Phenols and Cresols	VOC
CRVI	Chromium, hexavalent (& compounds)	Heavy Metal
DPM	Diesel PM (total)	PM
DPM10	Diesel PM_{10}	PM
DPM25	Diesel $PM_{2.5}$	PM
ETHE	Ethene	VOC
ETOH	Ethanol	VOC
FACD	Formic Acid	VOC
GLY	Glyoxal	VOC
HCHO	Formaldehyde	VOC
IPRD	Lumped isoprene product species	VOC
ISOP	Isoprene	VOC
MACR	Methacrolein	VOC
MECL	Methylene chloride	VOC
MEK	Ketones which react with OH radicals faster than 5×10^{-13} but slower than 5×10^{-12} cm^3 $molec^{-2}$ sec^{-1}	VOC
MEOH	Methanol	VOC
MGLY	Methyl Glyoxal	VOC
MVK	Methyl Vinyl Ketone	VOC
MXYL	m-Xylene	VOC
NH3	Ammonia	
NOX	Nitrogen Oxides	
OLE1	Alkenes (other than ethene) with kOH less than $7x10^4$ ppm^{-1} min^{-1}	VOC
OLE2	Alkenes with kOH greater than $7x10^4$ ppm^{-1} min^{-1}	VOC
OXYL	o-Xylene	
PACD	Higher organic acids and peroxy acids	
PAHO*	PAHs, total, with individual components also reported	

Pollutant	Pollutant Description	Type
PAHW*	PAHs, total, w/o individual components reported	
PAS	$PM_{2.5}$ Arsenic	Heavy Metal
PCD	$PM_{2.5}$ Cadmium	Heavy Metal
PDCB	p-Dichlorobenzene	
PEC	$PM_{2.5}$ EC	PM
PERC	Perchloroethylene	VOC
PM	Total Particulate Matter	PM
PM10	Particulate matter up to 10 micrometers	PM
PM25	Particulate matter up to 2.5 micrometers	PM
PMC	PM coarse	PM
PMOTHR	$PM_{2.5}$ Others	PM
PNCOM	PM2.5 Non-Carbon organic matter	PM
PNI	$PM_{2.5}$ Nickel	Heavy Metal
PNO3	$PM_{2.5}$ Nitrate	Heavy Metal
POC	$PM_{2.5}$ OC	PM
PPB	$PM_{2.5}$ Lead	Heavy Metal
PRD2	Ketones and other non-aldehyde oxygenated products which react with OH radicals faster than 5×10^{-12} cm^3 $molec^{-2}$ sec^{-1}.	VOC
PRPE	Propene	VOC
PSO4	$PM_{2.5}$ Sulfate	PM
PXYL	p-Xylene	VOC
RCHO	Lumped $C3^+$ Aldehydes	VOC
RNO3 +	Lumped Organic Nitrates	
ROG	Reactive organic gas	
SESQ	Sesquiterpenes	VOC
SOX	Sulfur Oxides	
TERP	Terpenes	
TOLU	Toluene	VOC
ETCL3	Trichloroethylene	VOC
VICL	Vinyl chloride	VOC

Table 4. Toxic Air Contaminants and Model Selection

TACs	TACs/Emission categories	CALPUFF	CMAQ
DPM	On-road major traffic links On-road-centroid connectors Off-road mobile (locomotives, seaports, airports, etc.) Area sources (ag, construction, etc.) Point sources (stationary)	Explicit links (line sources) 1 x 1 km area sources Explicit links (line/area/point) 1 x 1 km area sources Points sources	

TACs	TACs/Emission categories	CALPUFF	CMAQ
HEAVY METALS	**Heavy metals** On-road major traffic links On-road - centroid connectors Off-road mobile Area sources Point sources	Explicit links (line sources) 1 x 1 km area sources Explicit links (line/area/point) 1 x 1 km area sources Points	
PM2.5 and O3	**All sources**		Gridded 2 x 2 km
VOC	**All sources**		Gridded 2 x 2 km

Diesel particulate matter (DPM), which is a subset of PM, is determined by classifying specific EICs from CEPAMv1.01 that emit diesel particulate matter based on source and material used. Table 5 lists all emission inventory categories that emit DPM, while a complete list of all EICs and classification reasons can be found in the master list (Toxics Inventory and Special Projects Section, 2020). DPM is further split into 10-micron and 2.5-micron modes using PM size fractions.

Toxic VOCs are primarily determined through use of CARB's SARPC07 chemical mechanism speciation profiles of total TOG from CEPAMv1.01estimates. However, some toxics VOCs are currently not available in the SARPC chemical mechanism profiles used in CMAQ modeling. The reporting of these stationary source species comes from California's AB2588 Air Toxics "HotSpots" emission inventory criteria and guidelines regulation, or more commonly known as "CTI" (Toxics Inventory and Special Projections Section, 2019). These species include Vinyl Chloride (VICL), Trichloroethylene (ETCL3), Methylene chloride (MECL), Carbon tetrachloride (CCL4) and Acrylonitrile (ACRY).

Stationary source heavy metals from particulate matter, DPM and other toxics pollutants are also reported in the CTI such as $PM_{2.5}$ Arsenic, $PM_{2.5}$ Cadmium, $PM_{2.5}$ Nickel, $PM_{2.5}$ Lead and Chromium, hexavalent (& compounds) (CRVI). Depending on the specific type of model and level of reporting within a district, the stationary source species may be directly from the CTI or estimated with the use of speciation profiles from CEPAMv1.01. Sections 2.1.2 and 3.1.2 further describe this difference.

Table 5. Diesel PM Emission Categories

EICSUM	Category Description
10	ELECTRIC UTILITIES
20	COGENERATION
30	OIL AND GAS PRODUCTION (COMBUSTION)
40	PETROLEUM REFINING (COMBUSTION)
50	MANUFACTURING AND INDUSTRIAL
52	FOOD AND AGRICULTURAL PROCESSING
60	SERVICE AND COMMERCIAL
99	OTHER (FUEL COMBUSTION)
710	LIGHT DUTY PASSENGER (LDA)

EICSUM	Category Description
720	LIGHT AND MEDIUM DUTY TRUCKS
722	LIGHT DUTY TRUCKS - 1 (LDT1)
723	LIGHT DUTY TRUCKS - 2 (LDT2)
724	MEDIUM DUTY TRUCKS (MDV)
740	HEAVY DUTY DIESEL TRUCKS (ALL)
742	LIGHT HEAVY DUTY DIESEL TRUCKS - 1 (LHDDT1)
743	LIGHT HEAVY DUTY DIESEL TRUCKS - 2 (LHDDT2)
744	MEDIUM HEAVY DUTY DIESEL TRUCKS (MHDDT)
746	HEAVY HEAVY DUTY DIESEL TRUCKS (HHDDT)
760	HEAVY DUTY DIESEL URBAN BUSES (UBD)
770	SCHOOL BUSES (SB)
772	SCHOOL BUSES - DIESEL (SBD)
776	OTHER BUSES (OB)
778	OTHER BUSES - MOTOR COACH - DIESEL (OBC)
779	ALL OTHER BUSES - DIESEL (OBD)
780	MOTOR HOMES (MH)
820	TRAINS
830	SHIPS AND COMMERCIAL BOATS
833	OCEAN GOING VESSELS
835	COMMERCIAL HARBOR CRAFT
840	RECREATIONAL BOATS
850	OFF-ROAD RECREATIONAL VEHICLES
860	OFF-ROAD EQUIPMENT
870	FARM EQUIPMENT

Emissions Processing for CALPUFF

CALPUFF emissions data are formatted as comma separated text files that specify the emission rate, spatial distribution, chemical speciation, and temporal distribution. CMAQ requires a different file format to CALPUFF, Network Common Data Form, or netCDF, interface files for array-oriented data. The remainder of this document has been split into two chapters: the first one describing CALPUFF dispersion modeling requirements and the second one for CMAQ photochemical modeling requirements.

I. Model Emissions Inventory Development for CALPUFF

A. Area Sources

Gridded area source emissions are prepared with the aid of the MCNC Environmental Modeling Center (EMC) emissions processor SMOKE (Sparse Matrix Operator Kernel of Emissions) version 4.0. Emission records are formatted into a SMOKE-ready "FF10" inventory file which is a format consisting of comma-delimited emissions that is read by SMOKE (Baek, 2015). Emission records are associated with GAI code and CARB's emission inventory codes

(EICs), which respectively identify in-state regions and detailed emission categories in 14 digits (EIC14 hereafter). The CEPAM 2019 inventory version 1.01 rf3043-snap20190508 is the current starting point for the area source toxics inventory. The highest DPM emitting area sources were processed separately from the remaining area sources to form more granular health risk assessments of impacted communities. Table 6 lists the current area source sectors while a master list (CARB A. , AreaSector_EIC_rf3043.xlsx, 2020) contains all EICs contributing to each isolated emission sector.

Many emission sources are impacted by meteorology fluxes in wind and rain, therefore SMOKE's GenTpro utility was used to apply meteorology-based temporal profiles. The sectors that utilize SMOKE's GenTpro include Residential Wood combustion, Agriculture Pesticides, Fertilizer and Livestock, and Paved and Unpaved road dust. Sections 2.1.1.1 – 2.1.1.4 describe detailed processing of these sectors. Day-specific burn events are detailed in section 2.1.1.5-2.1.1.6.

Table 6. Area Source Sectors

Area Source Sector	Description
AGBURN+	Agricultural Burning
AGRICULTURE	Offroad-Mobile Agriculture
AIRCRAFT	Aircraft
CHARBROIL	Charbroiling/Food Cooking
CHC	Commercial Harbor Craft
CHE	Cargo Handling Equipment
CONSTRUCTION	Offroad-Mobile Construction
FERTILIZER	Agriculture Fertilizer
LIVESTOCK	Agriculture Livestock
LOCO	Locomotives
MNGBURN	Managed Burns
OGV_MIL	Military Ocean Going Vessels
OTHERS	Aggregated Area (everything else)
PAVED RDUST	Paved Road Dust
RWC	Residential Wood Combustion
STRCBURN	Structural Burns
TRU	Transport Refrigeration Units
UNPAVED RDUST	Unpaved Road Dust
MXAREA*	Mexico Area
MXNONROAD*	Mexico Nonroad
MXONROAD*	Mexico Onroad
Rx Burn	Prescribed Burns
Wildfire	Wildfires

+See section on day-specific reporting of Agburns
*If applicable in domain (San Diego and Imperial only)

1. Residential Wood Combustion

Residential wood combustion (RWC), which consists of woodstoves and fireplaces, has a higher demand when the ambient temperature is low. As a result, RWC emissions must be adjusted in response to temperature and are assumed to have zero emissions if temperature stays 50 °F or above. SMOKE's GenTpro program adopts Adelman's equation for RWC (Adelman, 2010) to determine county-specific, daily allocation factors based on county-wide averages of modeled ambient temperature above 2m (Baek, 2015). Descriptions about input requirements, environment settings and expected output files for GenTpro can be found in (Kwok, 2016).

GenTpro generates a temporal cross reference (Gentpro_tref) file, one monthly profile (Gentpro_tpro_mon) file and one daily profile (Gentpro_tpro_day) file. Gentpro_tref lists lines of reference, each beginning with an RWC EIC14 followed by GAI which is appended with a keyword (either "MONTHLY", "DAILY", or "ALLDAY"). SMOKE will subsequently allocate monthly and daily RWC emissions with these GenTpro output files but uses the assigned hourly temporal profiles.

Three air districts (Sacramento Valley, San Joaquin Valley and South Coast) have RWC curtailment programs that were accounted for in this inventory by reducing emissions in specific days where a no burn day was called by each district. To implement the reductions, gridded hourly unspeciated SMOKE reports were processed and converted to pMED (CARB A. , MEDS_Format_20151027, 2015). District-specific reduction factors were applied to the RWC in the respective regions on days where $PM_{2.5}$ was greater than or equal to 35 µg/m^3. Although the Bay Area has a RWC reduction program, reduction factors were not used as advised by District staff. However, the RWC reductions from other sub-regions have impacts on this region.

2. Agriculture Pesticides and Fertilizers Processing

For pesticide and fertilizers application, TOG and NH_3 emissions are subject to GenTpro temporal variation which is based on MCIP (Meteorology-Chemistry Interface Processor) modeled meteorology. GenTpro creates GAI-level hourly fractions over the entire calendar year. The Russell-Cass algorithm that depends on 2-meter temperature and 10-m wind speed was adopted (Baek, 2015).

3. Agriculture Livestock Processing

Hourly fractions for agriculture livestock emissions are based on ground soil temperature and aerodynamic resistance. The GenTpro generated profiles are applied to NH3, PM, and TOG emissions.

Livestock emissions at known facilities are represented as FF10 point sources since these facilities have latitude and longitude coordinates. However, the emissions are not released from stacks and these emissions should be temporally allocated with meteorology-modulated fractions over a given calendar year through SMOKE's preprocessing program GenTpro. Since GenTpro only works with area sources, the point source FF10 is converted into the area format. To preserve location-specific emission portions, area-to-point files are constructed and used instead of spatial surrogates.

A Fortran program was written to transform the point source FF10 file into the area FF10 file. First the point source FF10 is split by pollutant to create individual files for NH3, PM, and TOG. The Fortran program aggregates the point source emissions by unique EIC and facility codes and then writes to an FF10 area format. At the same time, the program also writes out the area-to-point (ARTOPNT) file of the current pollutant. The ARTOPNT file ensures that fractions sum up to unity for all facilities within a GAI. These fractions represent emissions shared by facilities within each common GAI.

4. Paved and Unpaved Road Dust Processing

Rainfall can impact the total amount of road dust in a certain region and therefore an external process was applied through SMOKE's day-specific control matrices (CNTLMAT). AMSS staff utilized the Special Assessment Branch 2017-02-02 unadjusted paved road dust estimates and 2016-11-30 unadjusted unpaved road dust emissions. A python script creates the CNTLMAT files based on the following criteria:

1. Whenever the rain exceeds 0.01 inches in a corresponding GAI:
    - paved road dust emissions are to be reduced by 25%
    - unpaved emissions are completed removed

The resulting CNTLMAT contains the reductions for the road dust which were converted into standard SMOKE reports for CALPUFF processing.

5. Day-Specific Agburn Processing

Agriculture burns (Agburn) are either reported as annual estimates or day-specific burn events depending on the district reporting. Table 7 summarizes the districts that report day-specific burns.

The day specific burns are given a street location as it's burn site. The location would then manually be assigned a latitude/longitude by AMSS staff using Google or Geolocation tools. The specific point location is then assigned to a 1km x 1km grid cell within the modeling sub-domain of interest.

For the districts that do not provide day-specific events, annual totals are disaggregated using typical SMOKE temporal and spatial processing.

Table 7. Special Agburn Processing by District

Districts that Report Day-Specific burning	Districts that report Annual Data only
Bay Area Marsh	Amador
Great Basin	Antelope Valley
Imperial	Colusa
Sacramento	Kern
San Diego	Glenn
Santa Barbara (South Coast)	Mariposa
Siskiyou	Monterey Bay
SJV	Mendocino
SLO	North Coast

Districts that Report Day-Specific burning	Districts that report Annual Data only
Sutter	Sonoma
Yuba	Tuolumne
Tehama	Ventura
Solano	
Yolo	

6. Wildfires and Prescribed Burns

The emission inventory analysis section of AQPSD (Scott, 2017) provide the wildfire emissions and pre-gridded file. Both the emissions and the gridded files have a FIRE ID field to identify the fire. The gridded file has a start date, end date, duration of the fire and the fraction for the grid cell that has fire emissions. The final burn perimeter is used to allocate the emissions constantly for the duration of the fire event. A processing script uses the two files to create an inline file for CMAQ. Then CMAQ is run with this inline file with the 3D point source emissions diagnostic option. The 3D point source emissions diagnostic file is then reformatted to the CALPUFF format. This methodology applies to Rx and WFU emissions as well.

Future improvements look to develop inventories that change daily to better reflect the spatio-temporal aspect of the fires.

7. 3D Aircraft emission processing with GATE

Commercial aircraft emissions are processed with the use of the Gridded Aircraft Trajectory Emissions (GATE) Model. GATE is a command-line tool for processing raw aircraft emissions data into spatially and temporally allocated emission inventories, suitable for photochemical and dispersion modeling (CARB A. , Gridded Aircraft Trajectory Emissions Model (GATE), 2021). GATE is an open-source tool based on Python 3.6 that outputs CMAQ model ready files. The model was designed by CARB's staff to improve the vertical allocation of emissions for takeoff, landing and taxiing of aircraft. For CALPUFF dispersion modeling, AMSS staff modified GATE to output each flight path coordinates with the emissions and then reformat the file to CALPUFF suitable inputs (comma-separated) format.

B. Point Sources

CALPUFF stationary source inventory utilizes CEPMA2019 v1.01 for criteria pollutants, DPM and speciated heavy metals. DPM is determined by classifying specific EICs from CEPAM that emit diesel particulate matter based on material used. The 2017 CTI inventory is used for toxic VOCs reported explicitly by a facility and are not available through model species speciation. However, for the San Diego subdomain, the 2017 CTI was utilized for all toxic air contaminates as it was found to sufficiently report total DPM, heavy metals and toxics VOCs.

C. On-road Sources

The on-road portion of emission inventories are generated from the EMissionFACtor (EMFAC) model 2017 (CARB-MSEI, 2015). The EMFAC emissions model is used by CARB to assess emissions from on-road vehicles including cars, trucks, and buses in California, and to support air quality planning efforts to meet the Federal Highway Administration's transportation planning requirements. EMFAC is designed to produce county-level, average-day estimates (ARB-MSEI, 2015). As part of toxics inventory development, DPM_{10}, $DPM_{2.5}$ and PEC species are explicitly developed in addition to regular criteria pollutants. To generate toxics species from criteria pollutants, SAPRC07T chemical mechanism (gspro.WithInvpols.toxics.S07TT.072618.extra.CTI) fractions were used. These speciation fractions are denoted by vehicle type, process type and technology type. The PEC emissions were separated using speciation fractions applied on $PM_{2.5}$ totals and DPM emissions were separated using vehicle, technology, and process classification (CARB-MSEI, 2015).

For each model year, two weeks - one from summer and one from winter months - are selected as representative weeks in the year. The GAI level emission totals obtained from EMFAC are then distributed using surrogate fractions prepared for link and gridded sources. CALPUFF differs from onroad processing done for CMAQ and the Emissions Spatial and Temporal Allocator (ESTA) in that only link-based vehicle miles traveled (VMT) data are used (CARB A. , Emissions Spatial and Temporal Allocation (ESTA), 2019). Two files are used for processing, one file represents light-duty VMT, and the other file represents heavy-duty vehicles. The link files are created using data from the Integrated Transportation Network (ITN). A formal reference for version 2 of the ITN is available (Geophysics, 2005). Currently ITNv3 is used however formal documentation is not available. Future amendments through AMSS staff development are in progress for an ITNv4 (CARB A. , Alternative Onroad Processing Scripts, 2020). Further information regarding on-road spatial distribution for CALPUFF is referenced in section 2.2.3.

D. Ocean Going Vessels

Ocean going vessel (OGV) emissions are provided along with other area, stationary, and mobile source emissions from the latest available version of CEPAM. The preprocessing of this data into FF10 format is the same as other area and offroad mobile sources that are processed through SMOKE. The emissions data contain criteria pollutants broken down by EIC and COABDIS (county, air basin, district). At the EIC level the ships are categorized by vessel type, activity, source, and fuel. The offshore air basins extend out to 100 nautical miles from the coast. For CALPUFF the emissions data was processed into 4-vertices polygons for transit, maneuvering, and anchorage while the at-berth emissions were placed as stationary point sources. As a final step in this processing, the emission totals were compared against annual totals in CEPAMv1.01to check whether they reconcile with initial estimates. No vertical elevation is applied to OGV emissions as dispersion is calculated by CALPUFF. However, dispersion is only done to At Berth. Stack parameters are supplied to CALPUFF so that the model can calculate plume rise.

II. Spatial Distribution for CALPUFF

For dispersion models, when the actual location of a particular source of TACs is unknown or non-identifiable, the basic principle for spatially estimating emissions from such sources is to use the best available spatial surrogate. The following sections describe in more detail the type of spatial allocation used for each sector of the emissions inventory for CALPUFF modeling which is simulated on multiple 1km-subdomains (refer to section 1.2).

A. Area Sources

Each area source EIC is assigned to a specific spatial surrogate to disaggregate GAI-level total emissions to grid cells on the CA Statewide 1km modeling subdomains. Subdomains are broken down to cover specific regions or counties and these domains nest within the master 1km domain. Therefore, the CA Statewide 1km spatial surrogates are utilized to allocate emissions for each subdomain. The spatial surrogate snapshot 03-23-2020 was used with CEPAM 2019 v1.01. More information, methodologies and data source documentation can be referenced in the Spatial Surrogate Methodology document (CARB A. , Spatial Surrogate Methodology Document SNP20200323, 2020).

B. Point Sources

Each point source is allocated to grid cells using the latitude and longitude reported for each stack. If there are no stack latitude and longitude, the facility coordinates are used. Elevated point source release heights are calculated within the CALPUFF model. If any stack parameters are null or invalid (i.e. they fall outside of the acceptable range determined by SMOKE), then the fallback values in Table 8 are used.

Table 8. Fallback stack parameters

Stack Parameter	Units	Fallback Value
stack height	ft	121.4
stack diameter	ft	11.1
gas exit temperature	°F	699.5
gas exit velocity	ft/min	22.6

C. On-road Sources

The process for on-road toxics inventory spatial distribution is separated into two major steps. The first step is consolidating the highway links, and the second step is spatially allocating the emissions to the respective links or grid cells using Vehicle Miles Traveled as a surrogate (Atmospheric Modeling and Support Section, 2016).

Two CALPUFF link files are created using data from ITN (Geophysics, 2005). One file represents light-duty VMT, while the other file represents heavy-duty VMT. ITN combines data from travel demand models for all major Metropolitan Planning Organizations (MPOs) in California into one statewide master file. Where there is no MPO data available, data from the California Department of Transportation (Caltrans) California Statewide Travel Demand Model (CSTDM) (Caltrans, 2020) is used. Figure 2 shows an example statewide ITN file, while Table 9 contains all MPO information and the counties they cover throughout the statewide ITN file.

Figure 2. ITN statewide link file

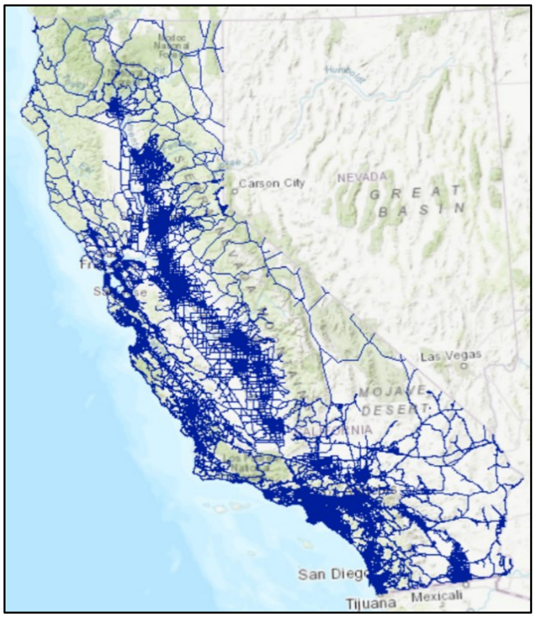


Table 9. MPO travel demand models

Metropolitan Planning Organization	County Name(s)
Association of Monterey Bay Area Governments (AMBAG)	Monterey, San Benito, Santa Cruz
Butte County Association of Governments (BCAG)	Butte
California Statewide Travel Demand Model (CSTDM)	Alpine, Amador, Calaveras, Colusa, Del Norte, Glenn, Humboldt, Inyo, Lake, Lassen, Mariposa, Mendocino, Modoc, Mono, Neveda, Plumas, Sierra, Siskiyou, Tehama, Trinity, Tuolumne, Ventura
Fresno Council of Governments (FCOG)	Fresno
Kings County Association of Governments (KCAG)	Kings
Kern Council of Governments (KCOG)	Kern
Merced County Association of Governments (MCAG)	Merced
Madera County Transportation Commission (MCTC)	Madera
Metropolitan Transportation Commission (MTC)	Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, Sonoma
Sacramento Area Council of Governments (SACOG)	El Dorado, Placer, Sacramento, Solano, Sutter, Yolo, Yuba
San Diego Association of Governments (SANDAG)	San Diego
Santa Barbara County Association of Governments (SBCAG)	Santa Barbara
Southern California Association of Governments (SCAG)	Imperial, Los Angeles, Orange, Riverside, San Bernardino, Ventura

Metropolitan Planning Organization	County Name(s)
San Joaquin Council of Governments (SJCOG)	San Joaquin
San Luis Obispo Council of Governments (SLOCOG)	San Luis Obispo
Shasta Regional Transportation Agency (SRTA)	Shasta
Stanislaus Council of Governments (StanCOG)	Stanislaus
Tulare County Association of Governments (TCAG)	Tulare
Tahoe Metropolitan Planning Organization (TMPO)	El Dorado, Placer

The statewide dataset contains various travel demand output variables including VMT, speed, number of lanes, and more. The variables from the MPO inputs are converted to a uniform format. MPO's provide different modeling years based on their own local regulations and studies. To create accurate spatial surrogates, the MPO travel demand model link data are converted to a uniform year. This is done by selecting the closest year of travel demand modeling data provided by a given MPO then applying EMFAC growth factors calculated by county.

Before ITN VMT is used for spatial allocation of CALPUFF emissions, the VMT totals are compared with EMFAC values. Table 10 shows an example comparison table. Manual inspection of the input MPO data is performed to remove any duplicate links and external centroid connectors that might exist.

Table 10. EMFAC VMT comparison

Link Name	EMFAC VMT	VMT	Percent Difference
ambag_links	19975467.61	17105445.56	14.37%
bcag_links	5274314.281	5454579.542	-3.42%
cofcg_links	26211826.97	25844465.59	1.40%
kcag_links	4344128.399	4047590.395	6.83%
kcog_links	28783275.55	27810881.67	3.38%
mctc_links	5115529.792	5348570.213	-4.56%
mtc_links	170358831	168741685.6	0.95%
noncoghdv_links		3260346.187	
noncogldv_links		19825564.12	
sacog_links	61073959.64	61883711.02	-1.33%
sandag_links	85008964.04	84645247.35	0.43%
sbcag_links	11259762.82	9422752.537	16.31%
scaghdv_links	37779533.42	34441389.2	8.84%
scagldv_links	440943254.9	421163310	4.49%
scrtpa_links	7616345.096	6407405.501	15.87%
slocog_links	7999881.128	8969367.627	-12.12%
tcag_links	11331869.44	11161948.5	1.50%

Link Name	EMFAC VMT	VMT	Percent Difference
threecounty_links	40854032.83	42411641.28	-3.81%
tmpo_links	2279176.721	1320345.956	42.07%

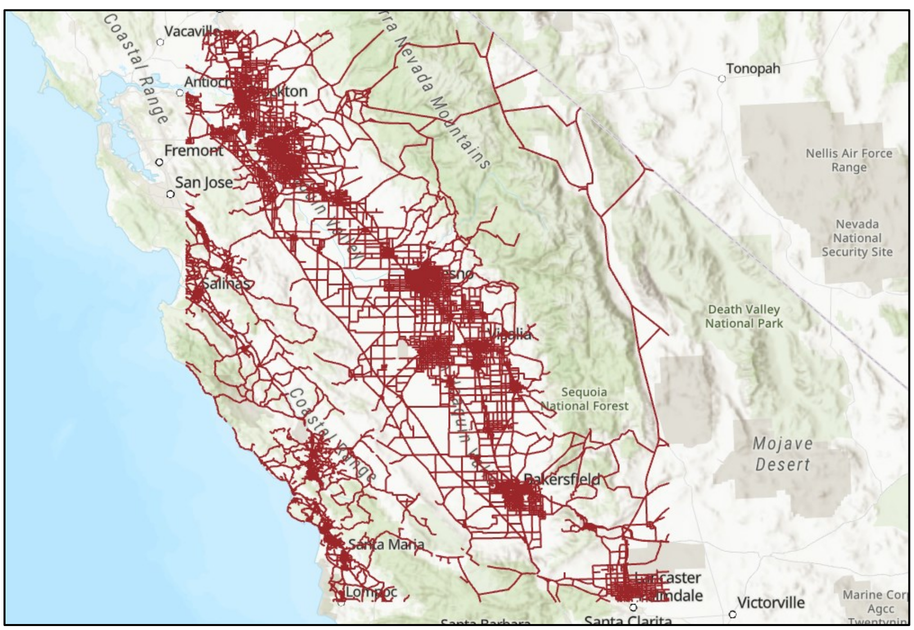
When possible, VMT data from the MPOs is split into light-duty and heavy-duty fractions. Table 11 shows which MPO data are currently being processed by vehicle type.

Table 11. MPO models that provide light-duty and heavy-duty splits

MPO	Vehicle Types Provided
CSTDM	Single Occupancy Vehicles, High-Occupancy Vehicles, Light Freight Trucks, Medium Freight Trucks
MTC	Trucks, Autos
SANDAG	Trucks, Autos, Buses
SCAG	Autos, Light Heavy-Duty Trucks, Medium Heavy-Duty Trucks, Heavy Heavy-Duty Trucks

VMT are aggregated by link into daily averaged VMT. All centroid links, which represent inter-Traffic Analysis Zone (TAZ) flow, are removed to avoid allocation to areas without roadways. Links are run through a gridding script which intersects links with a 50 m grid in order to be able to capture the bending of the roadways. Without this intersection, long traffic links can become oversimplified and do not represent the actual location of roadways. VMT totals are normalized by GAI and the weighting factor is assigned to each link accordingly. Figure 3 shows a sample CALPUFF link file for San Joaquin Valley.

Figure 3. San Joaquin Valley CALPUFF Link file



D. Ocean Going Vessels

OGV emissions are split into four different travel modes: Transit, At Berth, Maneuvering and At Anchorage. Emissions are spatially allocated differently to each of the travel modes as described below. Additionally, each travel mode is given a flat temporal profile meaning that the activity does not vary by month, day, or hour.

Transit: The emissions generated by ship travel between ports is assigned to transit activity. The overlay function within geopandas, which extends the datatypes used by pandas in python to allow spatial operations on geometric types, was used to clip the polygons that describe the ship lanes down to individual sub-domains and COABDIS regions. As a next step, the total transit line emissions were apportioned based on the updated transit link length. For CALPUFF, ship transit emissions are spatially allocated based on data from the Automated Instrumentation System (AIS) telemetry data collected in 2017. Using AIS data it was possible to separate the transit into three broad categories of vessels: cargo, tanker, and passenger. Emissions were assigned based on these vessel types.

At Berth: The at-berth locations were revised in August 2019 by the Atmospheric Modeling Support Section. The revised locations used online data sources for the individual ports where available. Plans or terminal maps were available for the ports of Oakland, San Diego, Benicia, Los Angeles, Long Beach, Humbolt, and Richmond. These sources were checked against aerial imagery available through ArcMAP and Google Earth. Finally, some berths were spot checked against data available from *https://www.marinetraffic.com/*. In some cases, it was possible to link specific vessel types to berths. Berths with a specific vessel type were only assigned emissions from that vessel type. If a vessel type could not be assigned to a specific berth, then emissions from all unassigned vessels were spread evenly across the berths.

Maneuvering: The maneuvering polygons were created in ArcMAP by linking transit lines with berth locations. A general rule of thumb was used to define the maneuvering lines as occurring at most 3 nautical miles away from the port location. The maneuvering polygons were shaped to follow the waterways leading into the ports and connecting directly to berths. Thus port-specific maneuvering emissions are allocated to these polygons. Maneuvering lines were constrained by the boundaries of the GAI boundary. No additional clipping was needed for any of these transit links as they all fall within their respective domains.

At Anchorage: The locations of the anchorage areas are defined using marine charts downloaded from NOAA (*https://nauticalcharts.noaa.gov/data/gis-data-and-services.html#seamless-raster-nautical-chart-services*). These locations were cross checked against data available from *https://www.marinetraffic.com/*. Emissions were distributed evenly across the area of the anchorages from all vessel types. The anchorage shapefiles were simplified to four polygons that could then be contained within the shape of the anchorage rasters from NOAA.

III. Chemical Speciation for CALPUFF

Emissions of TOG and PM for many sources can contain over hundreds of different chemical species. Speciation is the process of disaggregating these inventory pollutants into individual chemical species components or groups of species. CARB maintains and updates such speciation profiles for TOG and PM for a variety of source categories.

An organic gas (OG) speciation profile contains a list of organic compounds and the weight fraction that each compound comprises of the TOG emissions from a particular source type. The list includes all organic species that can become airborne (through evaporation, sublimation, as aerosols, etc.), excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides or carbonates and ammonium carbonate. The OG profiles are applied to TOG to estimate the amounts of various organic compound emissions, including toxic species.

The PM emissions are size fractionated by using PM size distribution profiles, which contain the total weight fraction for $PM_{2.5}$ and PM_{10} out of total PM. The fine and coarse PM chemical compositions are characterized by applying the PM chemical speciation profiles for each source type, which contains the weight fractions of each chemical species for $PM_{2.5}$, PM_{10} and total PM. PM chemical speciation profiles may also vary for different PM size fractions even for the same emission source. The chemical species in the PM profile consist of Particulate Elemental carbon (PEC), OC, sulfate, nitrate, ammonium, chloride, phosphorus, sulfur, chlorine, and metals including toxic species.

The OG profiles and PM profiles used for generating toxics emission inventories are available for download from CARB's web site at: *http://www.arb.ca.gov/ei/speciate/speciate.htm*. Each process or product category is keyed to one of the OG profiles and one of the PM profiles. Also available for download from CARB's web site is a cross-reference file that indicates which OG profile and PM profile are assigned to each category in the inventory. The inventory source categories are represented by an 8-digit source classification code (SCC) for point sources, or a 14-digit emission inventory code (EIC) for area and mobile sources. Some of the organic gas profiles and PM profiles related to motor vehicles, ocean going vessels, and fuel evaporative sources vary by the inventory year of interest, due to changes in fuel composition, vehicle fleet composition and diesel particulate filter (DPF) requirements over time. Details can be found in CARB's documentation of speciation profile development.

CARB has an ongoing effort to update speciation profiles as new data become available, such as through testing of emission sources or surveys of product formulations. New speciation data generally undergoes technical and peer review and updating of the profiles is coordinated with users of the data.

IV. Temporal Distribution for CALPUFF

The toxic emissions are adjusted temporally to represent variations by month, day of week and hour of day. Temporal data are stored in CARB's emission inventory database. Each local air district assigns temporal data for all processes at each facility in their district to represent when emissions at each process occur (Atmospheric Modeling and Support Section, 2016). Temporal profiles are generally consistent with those used for SIP modeling; more details can be found in (Atmospheric Modeling and Support Section, 2016).

V. Quality Assurance for CALPUFF

The emission allocation process is complex and therefore prone to errors. Quality assurance/quality control steps were conducted at the end of each step.

A. On-road Sources

While processing highway link activity, it was necessary to ensure that no activity within the domain was dropped. To verify this, the link and grid activity files are summarized. If the surrogate process is accurate, for each county/GAI that is contained within the domain, the VMT fractions should always sum to 1.0.

When allocating emissions, EMFAC totals by GAI are calculated. In addition, totals for each of the links and grid files are also calculated as part of spatial emissions allocation. Using VMT fractions, EMFAC totals (or the output file totals) are back calculated. This allows the direct comparison of emissions from EMFAC files and toxic inventory output files. In addition to emission totals, a set of shapefiles is routinely created for link sources and grid sources separately.

B. Area Sources

Since each Pre-CALPUFF text file is derived from two SMOKE reports (one with criteria pollutants and the other with PEC), daily totals are calculated from each file to ensure that mass is conserved. Since PM_{10} and $PM_{2.5}$ are determined by one set of split factors, the original total PM can be readily retrieved.

C. Point Sources

Spatial allocation of point sources is checked by geocoding each facility and verifying that the facilities fall within their assigned counties. There are several instances where assignments are flagged as incorrect. This information is relayed to the districts for further improvement of the inventory. Additional checks to point sources includes verifying that the PM_{10} or $PM_{2.5}$ does not exceed total PM are performed.

D. Ocean Going Vessels

Final toxic emission totals from the OGV sector are compared against the input totals. Since the top-down spatial allocation method —using spatial surrogates to distribute emissions spatially — is used for this sector, a simple county/GAI level comparison provides confidence in the results.

E. Residential Wood Combustion

Since RWC emissions reflect district-wide reduction programs, daily grid totals are generated for both the reduced pMEDs and the Pre-CALPUFF output text files. RWC consists of two EICs to represent woodstoves and fireplaces. These EICs have the same PM split profile and the same speciation profile. Using the PM and PEC split factors from the profiles, total PM can be derived from the Pre-CALPUFF files for comparison to the reduced pMEDs.

Emissions Processing for Photochemical Modeling (CMAQ)

CMAQ is a numerical air quality model that simulates the emission of pollutants and the subsequent chemistry and physical processes in the atmosphere that affect ambient pollution concentrations. The following sections describe the spatial, temporal, and chemical distribution of TACs for CMAQ modeling. As a general note, if a sector's methods are comparable to that of CALPUFF a reference will be made to that section in chapter 2. Repetition of noted sections was included for completeness in chapter 3.

I. Model Emissions Inventory Development for CMAQ

A. Area Sources

The CEPAM inventory version rf3043-snap20190508 is the current starting point for the area source toxics inventory. SMOKE4.0 runs through the usual speciation, gridding, and temporal allocation programs to form intermediate sparse matrices that in turn get merged to form gridded hourly netCDF files on the 2-km statewide modeling domain. CMAQ requires a single netCDF file for all aggregated area sources. A combine utility is used to merge all area source sectors into one gridded, hourly netCDF file.

Many emission sources are impacted by variability in meteorological factors such as temperature, wind and rain, therefore SMOKE's GenTpro utility is used as an option to specify meteorology based temporal profiles. The sectors that utilize SMOKE's GenTpro include: Residential Wood combustion, Agriculture Pesticides, Fertilizer and Livestock and Paved and Unpaved road dust. The emission processing of this sector follows the CALPUFF methodology and therefore can be referenced in sections 2.1.1.1 – 2.1.1.7 which describe the processing of these sectors in detail. The major difference between the two models is that CMAQ requires the output to be in a netCDF format whereas CALPUFF requires a SMOKE-report csv file. These are both available output formats from the SMOKE4.0 emission processor.

1. Residential Wood Combustion Sources

The meteorology-adjusted temporal profiles and districts emission reduction programs mentioned in section 2.1.1.1 for CALPUFF are also used in CMAQ processing. SMOKE netCDF output files on the 2-km statewide modeling domain were developed for use in CMAQ modeling.

2. Agriculture Pesticides and Fertilizers Processing

The meteorology-adjusted temporal profiles mentioned in section 2.1.1.2 for TOG and NH3 are also used in CMAQ processing. SMOKE netCDF output files on the 2-km statewide modeling domain were developed for use in CMAQ modeling.

3. Agriculture Livestock Processing

The meteorology-adjusted temporal profiles mentioned in section 2.1.1.3 for TOG and NH3 are also used in CMAQ processing. SMOKE netCDF output files on the 2-km statewide modeling domain were developed for use in CMAQ modeling.

4. Paved and Unpaved Road Dust Processing

The emission inventory and meteorology-adjusted temporal profiles mentioned in section 2.1.1.4 are also used in CMAQ processing. SMOKE netCDF output files on the 2-km statewide modeling domain were developed for use in CMAQ modeling.

5. Day-Specific Agburn Processing

The annual emission inventory and district reported day specific emissions mentioned in section 2.1.1.5 are also used in CMAQ processing. SMOKE netCDF output files on the 2-km statewide modeling domain were developed for use in CMAQ modeling.

6. Wildfires and Prescribed Burns

The annual emission inventory and district reported day specific emissions mentioned in section 2.1.1.6 are also used in CMAQ processing. SMOKE netCDF output inline files on the statewide domain were developed for use in CMAQ modeling instead of the diagnostic reports.

7. 3D Aircraft emissions processing with GATE

Commercial aircraft emissions are processed with the use of the Gridded Aircraft Trajectory Emissions Model in similar fashion as mentioned in section 2.1.1.7, however gridded netCDF files are created on the 2-km statewide modeling domain for use in CMAQ modeling.

B. Point Sources

CMAQ stationary source inventory is based on CEPMA2019 v1.01 for criteria pollutants and DPM. DPM is determined by classifying specific EICs from CEPAM that emit diesel particulate matter based on material used. The 2017 CTI inventory is used for toxic VOCs reported explicitly by a facility. Typically, CTI toxics emissions rely on voluntary reporting from individual in-state facilities once every four years. Since the CTI pollutants are explicit chemicals, they do not need speciation through SMOKE's spcmat program (Baek, 2015).

C. On-road Sources

The EMFAC emissions model as defined in section 2.1.3 is used by CARB to assess emissions from on-road vehicles. The estimates must be disaggregated spatially and temporally into gridded, hourly estimates for CMAQ modeling. The methodology used to disaggregate EMFAC emission estimates is to apply ESTA (emission spatial and temporal allocator) processing, which creates spatially and temporally distributed netCDF files for the CMAQ modeling system. The on-road spatial surrogates used for ESTA consist of heavy-duty link vehicle miles traveled (VMT), light-duty VMT, evaporative emissions, and idling surrogates.

Table 12 lists all ESTA surrogates used. More information regarding ESTA surrogate development can be found in section 3.2.3 and in the working reference document "ESTA_Surrogate_Doucmentation_03042021.docx". (CARB A. , ESTA Surrogate Documentation, 2021).

Table 12. ESTA spatial surrogate data sources.

Surrogate ID	Surrogate Description	Source
801-824	Hourly light-duty VMT	ITN
825-848	Hourly heavy-duty VMT	ITN
853	Idling 30%; Distribution 70%	EPA idling locations, Caltrans rest stops, distribution center parcels
859	Idling 90%; Distribution 10%	EPA idling locations, Caltrans rest stops, distribution center parcels
860	Diesel Motorcycles Evaporative Emissions	EMFAC fleet registration database
861	Gasoline Motorcycles Evaporative Emissions	EMFAC fleet registration database
862	Diesel Motorhomes/Buses Evaporative Emissions	EMFAC fleet registration database
863	Gasoline Motorhomes/Buses Evaporative Emissions	EMFAC fleet registration database
864	Diesel Passenger Evaporative Emissions	EMFAC fleet registration database
865	Gasoline Passenger Evaporative Emissions	EMFAC fleet registration database
866	Diesel Light/Medium Duty Trucks Evaporative Emissions	EMFAC fleet registration database
867	Gasoline Light/Medium Duty Trucks Evaporative Emissions	EMFAC fleet registration database
868	Diesel Heavy Duty Trucks Evaporative Emissions	EMFAC fleet registration database
869	Gasoline Heavy Duty Trucks Evaporative Emissions	EMFAC fleet registration database

D. Ocean Going Vessels Sources

OGV emissions for CMAQ toxics on the Statewide 2-km grid are based on SMOKE FF10 snapshot 20190508 derived from the CEPAM 2019 v1.01. It was split into three sectors: at-berth, ship lanes, and military.

At-berth emissions were classified based on the activity type "hoteling" characterized by EICSUB 9975 or 9995. Monthly/weekly temporal profiles were derived from Markit port activity data for 2016 covering Port of Los Angeles, Port of Long Beach, and berth locations in San Diego and Bay Areas. Berth locations were determined at grid centroids in the proximity of the ports. Model-ready netCDF files would be run as inline emissions.

Shiplane emissions cover three types of activities: transit, maneuvering, and anchorage. Flat temporal profiles were applied at all temporal intervals. The SMOKE output emissions were gridded at surface layer which were subsequently elevated to second and third layers evenly, with maximum altitude of 120m.

Military emissions cover all the activities. Without much information, flat temporal profiles were assumed. The subsequent SMOKE output emissions were also elevated to layers 2 and 3.

II. Spatial Distribution for CMAQ

CMAQ modeling is based on the CA state 2-km x 2-km grid domain and therefore has lower resolution than the CALPUFF dispersion model. The following sections describe the spatial allocation for each specific emission sector in CMAQ.

A. Area Sources

The process of spatially allocating area source emissions for CMAQ is like that of CALPUFF; more specific details are discussed in Section 2.2.1. The CA_State 2-km spatial surrogates were used to correspond with the modeling resolution needed in CMAQ.

B. Point Sources

The process of spatially allocating point source emissions to each grid cell for CMAQ is like that of CALPUFF; more detail can be found in section 2.2.2. However, elevated sources are treated differently in that they are processed within SMOKE. There are two types of point sources: elevated and non-elevated sources. Vertical distribution of emissions from elevated sources is accomplished using the plume rise algorithm in the SMOKE emissions processor, while non-elevated emissions are allocated to the first layer (Baek, 2015). Most stationary point sources with existing stacks are regarded as elevated sources. Those without physical stacks that provide only latitude/longitude, such as airports or landfills, are considered non-elevated.

C. On-Road Sources

On-road motor vehicle toxic emissions are spatially allocated in a similar fashion as for SIP modeling using the Integrated Transportation Network (ITN) and ESTA. Link-based VMT surrogates for light-duty and heavy-duty vehicles are split into hourly surrogates based on annually averaged VMT. All other ESTA surrogates have one surrogate for a representative day. VMT surrogate data comes from the Integrated Transportation Network (ITN), evaporative surrogates are sourced from the EMFAC fleet registration database, and idling surrogates are a combination of Caltrans rest stops, distribution centers, and EPA idling locations.

1. Processing Link-Based VMT Surrogates

ESTA's link based VMT surrogates are grouped into two sets of 24 surrogates. One set represents light-duty vehicles while the other is for heavy-duty vehicles. Each of the 24 surrogates in these sets represent hourly annually averaged VMT. The link based VMT surrogates were sourced from the current working version of ITN. Travel demand model output link data from all major California MPOs are combined into one statewide dataset. Where MPO data were unavailable, data from the California Department of Transportation

(Caltrans) California Statewide Travel Demand Model (CSTDM) were used. Table 9 shows each travel demand model used with a list of counties that each model represents.

This statewide dataset contains various travel demand output variables including VMT, speed, and number of lanes. To create accurate spatial surrogates, the MPO travel demand model link data are converted to a uniform year in a similar fashion that is done for CALPUFF modeling and described in section 2.2.3.

VMT, which is used to create the ESTA surrogates, is compared to EMFAC county VMT totals to ensure totals are accurate. Table 10 shows an example comparison table. Manual inspection of the input MPO data is performed to remove any duplicate links and external centroid connectors that might exist. When possible, VMT data from the MPOs is split into light-duty and heavy-duty fractions. Table 11 shows which MPO data are currently being processed by vehicle type.

All MPOs used in ITN provide VMT data for representative time periods relevant to the models' specific area of interest. However, these time periods aren't always congruent between bordering MPOs. To create uniform VMT totals, the period based VMT is converted to hourly VMT by dividing the total period VMT by the number of hours it is meant to represent. Table 13 shows the time period breakdown for each MPO used in ITN.

Table 13. MPO model time periods

MPO	Time Periods
AMBAG	-AM: 6:00 AM – 9:00 AM -Mid-Day: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -Evening/Night: 7:00 PM – 6:00 AM
BCAG	-AM: 6:00 AM – 9:00 AM -Midday: 9:00 AM - 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
CSTDM	-AM: 6:00 AM - 10:00 AM -Midday 9:00 AM – 3:00 PM -PM: 3:00 PM – 7:00 PM -Offpeak: 7:00 PM – 6:00 AM
FCOG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
KCAG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
KCOG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
MCAG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM

MPO	Time Periods
MCTC	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
MTC	-EA: 3:00 AM – 6:00 AM -AM: 6:00 AM – 10:00 AM -MD: 10:00 AM – 3:00 PM -PM: 3:00 PM – 7:00 PM -EV: 7:00 PM – 3:00 AM
SACOG	-A3: 7:00 AM – 10:00 AM -MD: 10:00 AM – 3:00 PM -P3: 3:00 PM – 6:00 PM -EV: 6:00 PM – 7:00 AM
SANDAG	-EA: 3:00 AM – 6:00 AM -AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 3:30 PM -PM: 3:30 PM – 7:00 PM -EV: 7:00 PM – 3:00 AM
SBCAG	-AM Peak: 7:00 AM – 9:00 AM -Late AM: 9:00 AM – 12:00 PM -Lunch: 12:00 PM – 2:00 PM -Early PM: 2:00 PM – 4:00 PM -PM Peak: 4:00 PM – 6:00 PM -Evening: 6:00 PM – 8:00 PM -Late Evening: 8:00 PM – 12:00 AM -Night: 12:00 AM – 7:00 AM
SCAG	-AM Peak: 6:00 AM – 9:00 AM -Midday: 9:00 AM – 3:00 PM -PM Peak: 3:00 PM – 7:00 PM -Evening: 7:00 PM – 9:00 PM -Night: 9:00 PM – 6:00 AM
SJCOG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
SLOCOG	-AM: 6:00 AM – 9:00 AM -Midday: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -Off Peak: 7:00 PM – 6:00 AM
SRTA	-AM: 6:00 AM – 9:00 AM -Midday: 9:00 AM – 3:00 PM -PM: 3:00 PM – 6:00 PM -OFF: 6:00 PM – 6:00 AM
StanCOG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM
TCAG	-AM: 6:00 AM – 9:00 AM -MD: 9:00 AM – 4:00 PM -PM: 4:00 PM – 7:00 PM -OFF: 7:00 PM – 6:00 AM

MPO	Time Periods
TMPO	-AM Peak: 7:00 AM – 10:00 AM -Midday: 10:00 AM – 4:00 PM -PM Peak: 4:00 PM – 7:00 PM -Late Night: 7:00 PM – 7:00 AM

The entire statewide link file from ITN is processed at once through the spatial allocator using hourly VMT as the weight field. The output surrogates are checked for artificial hotspots and/or missing areas.

2. Processing Idling and Distribution Surrogates

The idling and distribution surrogates are created by merging an idling surrogate with a distribution surrogate. Two separate surrogates are created using different weighting factors for the idling and distribution locations. The two idling/distribution splits are as follows: 90/10 and 30/70.

The distribution surrogate was created from Digital Map Products' large, high resolution, California-wide 2017 parcel data. In this shapefile, every parcel of land is covered by this dataset, and specific use code IDs were included to distinguish different types of land and registered information. Table 14 shows the use code descriptions used in the creation of the distribution surrogate.

Table 14. Parcel USE_CODE descriptions for Distribution Centers

USE_CODE ID	USE_CODE Description
2006	Grocery Stores/Super Markets
5003	Warehouse (Industrial)
5005	Food Packing / Packing Plant
5007	Food Processing
6002	Distribution Warehouse

The idling surrogate was created by combining the US EPA's idling locations dataset with the CalTrans rest stops dataset. The surrogate was weighted by truck count.

3. Processing Evaporative Surrogates:

Vehicle registration data were pulled from the EMFAC2017 fleet database. These data were grouped by vehicle and fuel type. Table 15 shows the vehicle type groupings that were used to create the evaporative emission surrogates for diesel and gas fuel types. These groups were summarized by Census block groups, which are defined as statistical divisions of census tracts that contain between 600 and 3,000 people. The summaries were assigned to the appropriate Census block group geometry.

To remove unpopulated areas from the census block group polygons, the data was intersected with Defense Meteorological Program Operational Linescan System Nighttime Lights (CARB, 2019) data set. An aggregate of 2016 through 2018 was used for the Nighttime lights dataset to capture any yearly variation that might occur. The shapefiles are then processed through the spatial allocator tool using vehicle count as the weighting field.

Table 15. EMFAC2017 fleet database vehicle group descriptions

Vehicle Groups	Vehicle Descriptions
Motorcycles	MC => Motorcycles
Motorhomes/Buses	BS => School buses BT => Urban buses B => All other buses MH => Motor Homes
Passenger Vehicles	P => Passenger Cars
Light/Medium Duty Trucks	T1 => Light-duty trucks (GVWR <6000 lbs, ETW ≤3750 lbs) T2 => Light-duty trucks (GVWR <6000 lbs, ETW 3751–5750 lbs) T3 => Medium-duty vehicles (GVWR 6000–8500 lbs) T4 => Light-heavy duty trucks (GVWR 8501–10,000 lbs)
Heavy Duty Trucks	T5 => Light-heavy duty trucks (GVWR 10,001–14,000 lbs) T6 => Medium-heavy duty trucks (GVWR 14,001–33,000 lbs) T7 => Heavy-heavy duty trucks (GVWR >33,000 lbs)

D. Ocean Going Vessels

There are three spatial approaches for OGV emissions modeling depending on the source type. The following sectors outline the spatial distribution for OGV at berth (inline), area (maneuvering, transit, and anchorage) as well as military OGV estimates.

1. OGV Inline

All OGV's in hoteling mode are assumed to be in port areas. Highly active ports have detailed arrivals and departures of vessels which aided in the development of port activity fractions. These fractions were used to disaggregate the original GAI emissions at each port location.

2. OGV Area and Military

A spatial cross-reference master file was created for OGV anchorage, maneuvering and transit activity modes. The file consists of three-tier spatial surrogate assignments, in order of relevance given availability of the first tier. The primary spatial distributions were derived based on AIS2017 data containing satellite-detected vessel traffic for 2017.

Existing spatial surrogates allocate OGV military emissions. Like OGV area, the resulting SMOKE-created surface emissions were elevated to second and third layers.

E. Residential Wood Combustion

The process of spatially allocating residential wood combustion emissions for CMAQ is like that of CALPUFF; more specific details are discussed in Section 2.1.5 (Baek, 2015).

III. Chemical Speciation for CMAQ

The process of chemical speciation for CMAQ is like that of CALPUFF; more specific details are discussed in Section 2.3 (Atmospheric Modeling and Support Section, 2016).

IV. Temporal Distribution for CMAQ

The process of temporally allocating emissions for CMAQ is like that of CALPUFF; more specific details are discussed in Section 2.4. However, point source temporal profiles for CMAQ are applied within SMOKE's temporal processing subroutine instead of appended to annual total reports.

References

Science Applications International Corporation. (2004). *EPA Spatial Allocator User Guide.* Raleigh, North Carolina.

Adelman, Z. (2010). *Proposal: Estimating Temporal Profiles for Residential Wood Combustion Emissions.*

Alpine Geophysics. (2020). *2017 Integrated Transportation Network (ITN).*

Anderson, C., Carlson, T., Dulla, R., & Wilkinson, J. (2012). *AGTOOL Agricultural sector emissions allocation model – Data development. CCOS_11-2.* Sacramento, CA: Sierra Research Inc.

ARB-MSEI. (2015). *Mobile Source Emissions Inventory -- Categories*. Retrieved 07 10, 2015, from http://www.arb.ca.gov/msei/categories.htm#emfac2014

ARB-PTSD. (2011, May). *Appendix D: Emissions estimation methodology for Ocean-going vessels.* Retrieved 7 10, 2015, from http://www.arb.ca.gov/regact/2011/ogv11/ogv11appd.pdf

Atmospheric Modeling and Support Section. (2016). *Emissions Inventory Protocol.* Unpublished.

Baek, B. a. (2015). *Final Summary Report: Development of SMOKE version 4.0. Under contract ITS:12-764.* Sacramento, CA.

Bay Area Air Quality Management District. (2009, December). *West Oakland Truck Study.* Retrieved from http://www.baaqmd.gov/~/media/files/planning-and-research/care-program/final-west-oakland-truck-survey-report-dec-2009.pdf

Board, T. R. (n.d.). Innovations in Freight Data. *TRANSPORTATION RESEARCH CIRCULAR E-C223.*

Boriboonsomsin, K., Sheckler, R., & Barth, M. (2012). Generating Heavy-Duty Truck Activity Data Inputs for MOVES Based on Large-Scale Truck Telematics Data. *Journal of the Transportation Research Board, 2270*, 49-58.

California Air Resources Board. (2006, September). *EMFAC Modeling Change Technical Memo: REDISTRIBUTION OF HEAVY-HEAVY DUTY DIESEL TRUCK VEHICLE.* Retrieved from https://www.arb.ca.gov/msei/onroad/techmemo/on-2006-05.pdf

California Air Resources Board. (2012, April 12). *Overview Diesel Exhaust and Health*. Retrieved from https://www.arb.ca.gov/research/diesel/diesel-health.htm

California Air Resources Board. (n.d.). *Mobile Source Emissions Inventory, Appendix G, Emissions Analysis 12 Methodology and Results: Medium-Heavy Duty Trucks, Heavy-Heavy Duty Diesel Trucks, and 13 Regulated Bus Categories.* Retrieved from https://www.arb.ca.gov/regact/2010/truckbus10/truckbusappg.pdf

Caltrans. (2020). *Transportiation Demand Modeling and Simulation*. Retrieved from California Department of Transportation : https://dot.ca.gov/programs/transportation-planning/multi-modal-system-planning/statewide-modeling

CARB. (2019). *Defense Meteorological Program Operational Linescan System Nighttime Lights.*

CARB, A. (2015). *Modeling emissions Data System (MEDS_Format).*

CARB, A. (2019). *Emissions Spatial and Temporal Allocation (ESTA)*. Retrieved from mmb-carb / ESTA_Documentation: https://github.com/mmb-carb/ESTA_Documentation

CARB, A. (2019). *Toxics Emission Inventory Protocol.*

CARB, A. (2020). *Alternative Onroad Processing Scripts.*

CARB, A. (2020). *AreaSector_EIC_rf3043.xlsx.*

CARB, A. (2020). *Spatial Surrogate Methodology Document SNP20200323.*

CARB, A. (2021). *ESTA Surrogate Documentation.*

CARB, A. (2021). *Gridded Aircraft Trajectory Emissions Model (GATE)*. Retrieved from mmb-carb/GATE: https://github.com/mmb-carb/GATE

CARB-MSEI. (2015). *Mobile Source Emissions Inventory -- Categories*. Retrieved 07 10, 2015, from http://www.arb.ca.gov/msei/categories.htm#emfac2014

Dun and Bradstreet. (2015). *Market insight database*. Retrieved June 15, 2015, from http://www.dnb.com

FactFinder, U. C. (2011). *"B25040: House heating fue. 2013 ACS 5-year estimatesl".* Retrieved March 2015, from http://factfinder2.census.gov

Funk, t., Stiefer, P., & Chinkin, L. (2001). *Development of gridded spatial allocation factors for the state of California.* San Joaquin Valley wide Air Pollution Study Agency and Environmental Protencion Agency- Air Resources Board. Sacramento: STI.

Garcia, C. B.-S. (2017). *AB-617 Nonvehicular air pollution: criteria air pollutants and toxic air contaminants.* Legislation, Sacramento.

Geophysics, A. (2005). *Development of Version Two of the California Integrated Transportation Network (ITN).*

Governments, S. C. (n.d.). *Transportation Models*. Retrieved from http://www.scag.ca.gov/DataAndTools/Pages/TransportationModels.aspx

Hernandez, S. (2014). Integration of Weigh-In-Motion and Inductive Signature Data for Truck Body Classification. *Ph.D. Dissertation*.

Hernandez, S., Tok, A., & Ritchie, S. (2013). *Transportation Research Board 93rd Annual Meeting and Transportation Research Record.* Retrieved from http://www.its.uci.edu/its/publications/papers/ITS/UCI-ITS-WP-13-3.pdf

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., . . . Wickham, J. a. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing, 81*(5), 345-354. Retrieved 7 10, 2015, from http://www.mrlc.gov/nlcd2011.php

Housten, D., M., K., & Winer, A. (2008). Diesel truck traffic in port-adjacent low-income and minority communities; environmental justice implications of near-roadway land use conflicts. 2008;. *J Transportation Research Board.*

Kwok, R. (2016). Meteorology-adjusted Temporal Profiles for Agricultural and Residential Wood Combustion Sectors Using Smoke Gentpro Utility Program. Sacramento: INTERNAL DRAFT CARB.

Mc. Rae, G. J. (1980). A simple procedure for calculating atmospheric water vapor concentration. *J. Air Pollution Control Association, 30*, 384.

Office of Air Quality Planning and Standards. (2002). *Example Application of Modeling Toxic Air Pollutants in Urban Areas.* Research Triangle Park, North Carolina: U.S. EPA.

Recker, W. (2017). *Operational Improvements to the California Vehicle Activity Database (CalVAD).* California Air Resources Board.

Reid, S., Penfold, B., & Chinkin, L. (2006). *Emission inventory for the Central California Ozone Study (CCOS) review of spatial variations of area, non-road mobile, and point sources of emissions.* Sonoma Technology Inc. Petaluma, CA: STI.

Scott, K. (2017). *Wild fire emissions and pre-gridded files (personal communication).*

Systems Applications Inc. (2001). *DTIM 4 User's Guide. Prepared for the California Department of Transportation.* Sacramento: Systems Applications, Inc.

Toxics Inventory and Special Projections Section, C. (2019). *2017 Stationary Toxics Inventory_CEIDARS_Final_102419.xlsx.*

Toxics Inventory and Special Projects Section, C. (2020). *DPM_SCC_SIC_EIC_Updated_20200307_tocsv.xlsx.*

U.S. E.P.A. (2017). *Air Quality Dispersion Modeling - Preferred and Recommended Models.* Retrieved from Support Center for Regulatory Atmospheric Modeing (SCRAM): https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models#main-content

U.S. E.P.A. (2020). *CMAQ User's Guide.* Retrieved from https://github.com/USEPA/CMAQ/tree/master/DOCS/Users_Guide

U.S. EPA. (2014). *Draft Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and RegionalHaze.* United States Environmental Protection Agency, Research Triangle Park. North Carollina: U.S. EPA. Retrieved July 10, 2015, from http://www.epa.gov/ttn/scram/guidance/guide/Draft_O3-PM-RH_Modeling_Guidance-2014.pdf

Appendix

Table A1: Comprehensive list of spatial surrogates used to allocate emissions

Surrogate Name	Surrogate Definition
AEROSPACE	Spatial distribution of businesses involved in aerospace
Airports	Spatial locations of all airports
All_PavedRds	Spatial distribution of road network (all paved roads)
AutobodyShops	Locations of auto-body repair and refinishing shops
CAFO	Spatial distribution of concentrated animal feeding operations
CANCOIL	Spatial distribution of businesses involved in can and coil operations
Cemeteries	Spatial locations of cemeteries
Comm_Airports	Spatial locations of commercial airports
COMPOST	Spatial distribution of composting
CONSTRUCTION_EQUIP	Spatial distribution of where construction equipment is used
Devplnd_HiDensity	Spatial distribution of developed land - low density, medium density and high density
Devplnd_LoDensity	Spatial distribution of developed land - open space (lowest density)
DREDGE	Locations of dredging
Drycleaners	Locations of dry cleaning facilities
DryLakeBeds	Locations of dry lake beds
Elev5000ft	Topological contours – areas above 5000 feet
Employ_Roads	Spatial distribution of total employment and road density (all paved roads)
FABRIC	Spatial distribution of businesses involved in fabric manufacturing
FERRIES	Locations of ferry ports and routes
FISHING_COMM	Locations of commercial fishing
Forestland	Spatial distribution of forest land
Fugitive_Dust	Spatial distribution of barren land
GAS_DISTRIBUTION	Location of gas pipelines
GAS_SEEP	Location of natural-occurring gas seeps
GasStations	Locations of gasoline service stations
GASWELL	Locations of gas wells
GolfCourses	Spatial locations of golf courses
HE_Sqft	Computed surrogate based on housing and employment (est. ft2 / person)
Hospitals	Spatial locations of hospitals
Housing	Spatial distribution of total housing
Housing_Autobody	Spatial distribution of housing and auto-body refinishing shops
Housing_Com_Emp	Spatial distribution of total housing and commercial employment
Housing_Restaurants	Spatial distribution of total housing and restaurants/bakeries
Surrogate Name	Surrogate Definition
INDUSTRIAL	Spatial distribution of industrial businesses where manufacturing occurs (SIC<4000)
Industrial_Emp	Spatial distribution of industrial employment
InlandShippingLanes	Spatial distribution of major shipping lanes within bays and inland areas

Surrogate Name	Surrogate Definition
Irr_Cropland	Spatial location of agricultural cropland
Lakes_Coastline	Locations of lakes, reservoirs, and coastline
LAKES_RIVERS_RECBOAT	Locations of lakes, rivers and reservoirs where recreational boats are used
LANDFILLS	Locations of landfills
LANDPREP	Spatial distribution of dust from land preparation operations (e.g. tilling)
LINEHAUL	Spatial distribution of Class I rail network
LiveStock	Spatial distribution of cattle ranches, feedlots, dairies, and poultry farms
MARINE	Spatial distribution of businesses involved in marine
METALFURN	Spatial distribution of businesses involved in metal furniture
METALPARTS	Spatial distribution of businesses involved in metal parts and products
Metrolink_Lines	Spatial distribution of metro-link network
MILITARY_AIRCRAFT	Locations of landing strips on military bases
MILITARY_SHIPS	Locations of military ship activity
MILITARY_TACTICAL	Military bases where tactical equipment are used
MiltaryBases	Locations of military bases
NON_PASTURE_AG	Spatial distribution of farmland
NonIrr_Pastureland	Spatial location of pasture land
NonRes_Chg	Computed surrogate based on spatial distribution of non-residential areas
OCEAN_RECBOAT	Locations of recreational boat activity that can occur on the ocean and SF Bay
OIL_SEEP	Location of naturally-occurring oil seeps
OILWELL	Locations of oil wells (both onshore and offshore)
OTHERCOAT	Spatial distribution of businesses with SIC<4000 not included in another category
PAPER	Spatial distribution of businesses involved in paper
PASTURE	Spatial distribution of grazing land
PEST_ME_BR	Spatial distribution of methyl bromide pesticides
PEST_NO_ME_BR	Spatial distribution of non-methyl bromide pesticides
PLASTIC	Spatial distribution of businesses involved in plastic
Pop_ComEmp_Hos	Spatial distribution of hospitals, population and commercial employment
Population	Spatial distribution of population
Ports	Locations of shipping ports
POTWs	Coordinate locations of POTWs
PrimaryRoads	Spatial distribution of road network (primary roads)
PRINT	Spatial distribution of print businesses
Raillines	Spatial distribution of railroad network
RailYards	Locations of rail yards
Rds_HE	Calculated surrogate based on road densities and housing/employment (est. ft2 / person)
RefinieriesTankFarms	Coordinate locations of refineries and tank farms
Res_NonRes_Chg	Computed surrogate based on spatial distribution of residential and non-residential areas
ResGasHeating	Spatial distribution of homes using gas supplied by a utility as primary source of heating

Surrogate Name	Surrogate Definition
Residential_Chg	Computed surrogate based on spatial distribution of residential areas
ResLPGHeat	Spatial distribution of homes using gas (bottled, tank or LP) as primary source of heating
ResNonResChg_IndEmp	Spatial distribution of industrial employment and residential/non-residential change
ResOilHeat	Spatial distribution of homes using fuel oil or kerosene as primary source of heating
Restaurants	Locations of restaurants
ResWoodHeating	Spatial distribution of homes using wood as primary source of heating
Fireplaces	Spatial distribution of residential wood heating actually being used by RWC woodstoves and fireplaces
Surrogate Name	Surrogate Definition
SandandGravelMines	Locations of sand/gravel excavation and mining
Schools	Spatial locations of schools
SecondaryPavedRds	Spatial distribution of road network (secondary roads)
SEMICONDUCT	Spatial distribution of businesses involved in semiconductors
Ser_ComEmp_Sch_GolfC_Cem	Spatial distribution of service and commercial employment, schools, cemeteries, golf courses
Service_Com_Emp	Spatial distribution of service and commercial employment
Shiplanes	Spatial distribution of major shipping lanes
SILAGE	Spatial distribution of silage operations
SingleHousingUnits	Spatial distribution of single dwelling units
TRU	Spatial distribution of transport refrigeration units
TUG_TOW	Spatial distribution of tug and tow boats
UnpavedRds	Spatial distribution of road network (unpaved roads)
Wineries	Locations of wineries
WOOD	Spatial distribution of businesses using wood
WOODFURN	Spatial distribution of businesses involved in wood furniture