



Technical Support Document (TSD): Preparation of Emissions Inventories for the 2021 North American Emissions Modeling Platform

Technical Support Document (TSD) Preparation of Emissions Inventories for the 2021 North American
Emissions Modeling Platform

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

Authors:

Alison Eyth (EPA/OAR)

Jeff Vukovich (EPA/OAR)

Caroline Farkas (EPA/OAR)

Janice Godfrey (EPA/OAR)

Karl Seltzer (EPA/OAR)

Yijia Dietrich (EPA/OAR)

TABLE OF CONTENTS

LIST OF TABLES.....	VII
LIST OF FIGURES.....	IX
ACRONYMS.....	X
1 INTRODUCTION	13
2 EMISSIONS INVENTORIES AND APPROACHES.....	15
2.1 POINT SOURCES (PTEGU, PT_OILGAS, PTNONIPM, AIRPORTS)	20
2.1.1 EGU sector (ptegu)	22
2.1.2 Point source oil and gas sector (pt_oilgas).....	23
2.1.3 Aircraft and ground support equipment (airports).....	25
2.1.4 Non-IPM sector (ptnonipm)	26
2.2 NONPOINT SOURCES (AFDUST, FERTILIZER, LIVESTOCK, NP_OILGAS, RWC, VCP, NONPT).....	26
2.2.1 Area fugitive dust sector (afdust)	27
2.2.2 Agricultural Livestock (livestock)	32
2.2.3 Agricultural Fertilizer (fertilizer).....	33
2.2.4 Nonpoint Oil and Gas Sector (np_oilgas).....	36
2.2.5 Residential Wood Combustion (rwc).....	40
2.2.6 Solvents (np_solvents)	41
2.2.7 Nonpoint (nonpt)	42
2.3 ONROAD MOBILE SOURCES (ONROAD)	43
2.3.1 Inventory Development using SMOKE-MOVES	44
2.3.2 Onroad Activity Data Development	46
2.3.3 MOVES Emission Factor Table Development	48
2.3.4 Onroad California Inventory Development (onroad_ca_adj).....	52
2.4 NONROAD MOBILE SOURCES (CMV, RAIL, NONROAD)	53
2.4.1 Category 1, Category 2 Commercial Marine Vessels (cmv_c1c2).....	53
2.4.2 Category 3 Commercial Marine Vessels (cmv_c3).....	60
2.4.3 Railway Locomotives (rail)	65
2.4.4 Nonroad Mobile Equipment (nonroad).....	70
2.5 FIRES (PTFIRE-RX, PTFIRE-WILD, PTAGFIRE)	73
2.5.1 Wild and Prescribed Fires (ptfire-rx, ptfire-wild).....	73
2.5.2 Point source Agriculture Fires (ptagfire).....	78
2.6 BIOGENIC SOURCES (BEIS).....	79
2.7 SOURCES OUTSIDE OF THE UNITED STATES.....	81
2.7.1 Point Sources in Canada and Mexico (canmex_point)	82
2.7.2 Fugitive Dust Sources in Canada (canada_afdust, canada_ptdust)	83
2.7.3 Agricultural Sources in Canada and Mexico (canmex_ag)	83
2.7.4 Surface-level Oil and Gas Sources in Canada (canada_og2D).....	83
2.7.5 Nonpoint and Nonroad Sources in Canada and Mexico (canmex_area).....	83
2.7.6 Onroad Sources in Canada and Mexico (canada_onroad, mexico_onroad).....	84
2.7.7 Fires in Canada and Mexico (ptfire_othna).....	84
2.7.8 Ocean Chlorine, Ocean Sea Salt, and Volcanic Mercury	84
3 EMISSIONS MODELING	85
3.1 EMISSIONS MODELING OVERVIEW	85
3.2 CHEMICAL SPECIATION	89
3.2.1 VOC speciation	94
3.2.2 PM speciation.....	99
3.2.2.1 Diesel PM	99
3.2.3 NO _x speciation	100
3.2.4 Sulfuric Acid Vapor (SULF).....	100
3.2.5 Speciation of Metals and Mercury	101
3.3 TEMPORAL ALLOCATION.....	103

3.3.1	<i>Use of FF10 format for finer than annual emissions</i>	105
3.3.2	<i>Temporal allocation for non-EGU sources (ptnonipm)</i>	105
3.3.3	<i>Electric Generating Utility temporal allocation (ptegu)</i>	105
3.3.4	<i>Airport Temporal allocation (airports)</i>	110
3.3.5	<i>Residential Wood Combustion Temporal allocation (rwc)</i>	113
3.3.6	<i>Agricultural Ammonia Temporal Profiles (livestock)</i>	117
3.3.7	<i>Oil and gas temporal allocation (np_oilgas)</i>	120
3.3.8	<i>Onroad mobile temporal allocation (onroad)</i>	120
3.3.9	<i>Nonroad mobile temporal allocation (nonroad)</i>	125
3.3.10	<i>Additional sector specific details (afdust, beis, cmv, rail, nonpt, ptfire-rx, ptfire-wild)</i>	126
3.4	SPATIAL ALLOCATION	128
3.4.1	<i>Spatial Surrogates for U.S. emissions</i>	128
3.4.2	<i>Allocation method for airport-related sources in the U.S.</i>	143
3.4.3	<i>Surrogates for Canada and Mexico emission inventories</i>	143
4	EMISSION SUMMARIES	154
5	REFERENCES	158

List of Tables

Table 2-1. Platform sectors used in the Emissions Modeling Process	16
Table 2-2. Point source oil and gas sector NAICS Codes	23
Table 2-3. Point source oil and gas sector emissions for 2021	24
Table 2-4. SCCs for the airports sector	25
Table 2-5. Afdust sector SCCs	27
Table 2-6. Total impact of 2021 fugitive dust adjustments to unadjusted inventory	28
Table 2-7. SCCs for the livestock sector	33
Table 2-8. Source of input variables for EPIC	35
Table 2-9. Nonpoint oil and gas emissions for 2021	36
Table 2-10. State emissions totals for year 2021 for Pipeline Blowdowns and Piggings sources.....	38
Table 2-11. State emissions totals for year 2021 for Abandoned Wells sources	39
Table 2-12. SCCs for the residential wood combustion sector	41
Table 2-13. Non-VCPy SCCs in the np_solvents sector	42
Table 2-14. MOVES vehicle (source) types.....	44
Table 2-15. The fraction of IHS vehicle populations retained for 2020 NEI and 2021 emissions modeling platform by model year	51
Table 2-16. SCCs for the cmv_c1c2 sector	54
Table 2-17. Vessel groups in the cmv_c1c2 sector	59
Table 2-18. SCCs for cmv_c3 sector.....	60
Table 2-19. SCCs for the Rail Sector	66
Table 2-20. 2020 R-1 Reported Locomotive Fuel Use for Class I Railroads	67
Table 2-21. 2020 Class II/III Line Haul Fleet by Tier Level	68
Table 2-22. SCCs included in the ptfire sector for the 2021 platform	73
Table 2-23. SCCs included in the ptagfire sector.....	78
Table 2-24. Meteorological variables required by BEIS4.....	80
Table 3-1. Key emissions modeling steps by sector.....	86
Table 3-2. Descriptions of the platform grids	88
Table 3-3. Emission model species produced for CB6R5_AE7 for CMAQ.....	89
Table 3-4. Additional HAP gaseous model species generated for toxics modeling	91
Table 3-5. Additional HAP particulate model species generated for toxics modeling	92
Table 3-6. PAH/POM pollutant groups.....	92
Table 3-7. Integration status for each platform sector.....	95
Table 3-8. Integrated species from MOVES sources	96
Table 3-9. Mobile Speciation Profile Updates	97
Table 3-10. Mobile NOx and HONO fractions	99
Table 3-11. NOx speciation profiles.....	100
Table 3-12. Sulfate Split Factor Computation.....	101
Table 3-13. SO2 speciation profiles	101
Table 3-14. Particle Size Speciation of Metals	102
Table 3-15. Mercury Speciation Profiles.....	103
Table 3-16. Temporal settings used for the platform sectors in SMOKE	104
Table 3-17. U.S. Surrogates available for the 2021 modeling platforms	131
Table 3-18. Shapefiles used to develop U.S. Surrogates.....	132
Table 3-19. Surrogates used to gapfill U.S. Surrogates	136
Table 3-20. Off-Network Mobile Source Surrogates	139
Table 3-21. Spatial Surrogates for Oil and Gas Sources	140
Table 3-22. Selected 2021 CAP emissions by sector for U.S. Surrogates (short tons in 12US1).....	141

Table 3-23. Canadian Spatial Surrogates	144
Table 3-24. Shapefiles and Attributes used to Compute Canadian Spatial Surrogates	145
Table 3-25. Shapefiles and Attributes used to Compute Mexican Spatial Surrogates	150
Table 3-26. 2021 CAP Emissions Allocated to Mexican and Canadian Spatial Surrogates for 12US1 (short tons)	150
Table 4-1. National by-sector CAP emissions for the 2021 platform, 12US1 grid (tons/yr)	155
Table 4-2. National by-sector VOC HAP emissions for the 2021 platform, 12US1 grid (tons/yr)	156
Table 4-3. National by-sector Diesel PM and metal emissions for the 2021 platform, 12US1 grid (tons/yr)	157

List of Figures

Figure 2-1. Fugitive dust emissions and impact of adjustments due to transport fraction, precipitation, and cumulative	30
Figure 2-2. “Bidi” modeling system used to compute emissions from fertilizer application	34
Figure 2-3. Map of Representative Counties.....	50
Figure 2-4. NEI Commercial Marine Vessel Boundaries and Automatic Identification System Request Boxes for the 2021 Emissions Modeling Platform.....	57
Figure 2-5. 2019 Class I Railroad Line Haul Activity	67
Figure 2-6. Class II and III Railroads in the United States	69
Figure 2-7. Amtrak National Rail Network.....	70
Figure 2-8. Processing flow for fire emission estimates in the 2021 inventory	76
Figure 2-9. Default fire type assignment by state and month where data are only from satellites	76
Figure 2-10. Blue Sky Modeling Pipeline.....	77
Figure 2-11. Annual biogenic VOC BEIS4 emissions for the 12US1 domain	81
Figure 3-1. Air quality modeling domains	88
Figure 3-2. Process of integrating HAPs and speciating VOC in a modeling platform.....	95
Figure 3-3. Eliminating unmeasured spikes in CEMS data	106
Figure 3-4. Regions used to Compute Temporal non-CEMS EGU Temporal Profiles	108
Figure 3-5. Example Daily Temporal Profiles for the LADCO Region and the Gas Fuel Type	109
Figure 3-6. Example Diurnal Temporal Profiles for the MANE-VU Region and the Coal Fuel Type	109
Figure 3-7. 2021 Airport Diurnal Profiles for PHX and state of Texas	111
Figure 3-8. 2021 Wisconsin and Atlanta annual-to-month profile for airport emissions	112
Figure 3-9. Alaska seaplane profile.....	113
Figure 3-10. Example of RWC temporal allocation using a 50 versus 60 °F threshold.....	114
Figure 3-11. Example of Annual-to-day temporal pattern of recreational wood burning emissions.....	115
Figure 3-12. RWC diurnal temporal profile	116
Figure 3-13. Data used to produce a diurnal profile for hydronic heaters	117
Figure 3-14. Monthly temporal profile for hydronic heaters	117
Figure 3-15. Examples of livestock temporal profiles in several parts of the country.....	119
Figure 3-16. Example of animal NH ₃ emissions temporal allocation approach (daily total emissions).....	119
Figure 3-17. TMAS Data: VMT Fraction by Hour of Day, Day of Week, and Month of Year	121
Figure 3-18. Example temporal variability of VMT compared to onroad NO _x emissions.....	124
Figure 3-19. Example Nonroad Day-of-week Temporal Profiles	125
Figure 3-20. Example Nonroad Diurnal Temporal Profiles.....	126
Figure 3-21. Agricultural burning diurnal temporal profile	127
Figure 3-22. Prescribed and Wildfire diurnal temporal profiles	128
Figure 3-23. 2020 Residential Wood Combustion Emissions using NLCD Low Intensity Surrogate	130
Figure 3-24. 2020 Residential Wood Combustion Emissions using ACS-based Surrogate	130

Acronyms

AADT	Annual average daily traffic
AE6	CMAQ Aerosol Module, version 6, introduced in CMAQ v5.0
AEO	Annual Energy Outlook
AERMOD	American Meteorological Society/Environmental Protection Agency Regulatory Model
AIS	Automated Identification System
APU	Auxiliary power unit
BEIS	Biogenic Emissions Inventory System
BELD	Biogenic Emissions Land use Database
BenMAP	Benefits Mapping and Analysis Program
BPS	Bulk Plant Storage
BSP	Blue Sky Pipeline
BTP	Bulk Terminal (Plant) to Pump
C1C2	Category 1 and 2 commercial marine vessels
C3	Category 3 (commercial marine vessels)
CAMD	EPA's Clean Air Markets Division
CAM_x	Comprehensive Air Quality Model with Extensions
CAP	Criteria Air Pollutant
CARB	California Air Resources Board
CB05	Carbon Bond 2005 chemical mechanism
CB6	Version 6 of the Carbon Bond mechanism
CBM	Coal-bed methane
CDB	County database (input to MOVES model)
CEMS	Continuous Emissions Monitoring System
CISWI	Commercial and Industrial Solid Waste Incinerators
CMAQ	Community Multiscale Air Quality
CMV	Commercial Marine Vessel
CNG	Compressed natural gas
CO	Carbon monoxide
CONUS	Continental United States
CoST	Control Strategy Tool
CRC	Coordinating Research Council
CSAPR	Cross-State Air Pollution Rule
E0, E10, E85	0%, 10% and 85% Ethanol blend gasoline, respectively
ECA	Emissions Control Area
ECCC	Environment and Climate Change Canada
EF	Emission Factor
EGU	Electric Generating Units
EIA	Energy Information Administration
EIS	Emissions Inventory System
EPA	Environmental Protection Agency
EMFAC	EMission FACtor (California's onroad mobile model)
EPIC	Environmental Policy Integrated Climate modeling system
FAA	Federal Aviation Administration

FCCS	Fuel Characteristic Classification System
FEST-C	Fertilizer Emission Scenario Tool for CMAQ
FF10	Flat File 2010
FINN	Fire Inventory from the National Center for Atmospheric Research
FIPS	Federal Information Processing Standards
FHWA	Federal Highway Administration
HAP	Hazardous Air Pollutant
HMS	Hazard Mapping System
HPMS	Highway Performance Monitoring System
ICI	Industrial/Commercial/Institutional (boilers and process heaters)
I/M	Inspection and Maintenance
IMO	International Marine Organization
IPM	Integrated Planning Model
LADCO	Lake Michigan Air Directors Consortium
LDV	Light-Duty Vehicle
LPG	Liquified Petroleum Gas
MACT	Maximum Achievable Control Technology
MARAMA	Mid-Atlantic Regional Air Management Association
MATS	Mercury and Air Toxics Standards
MCIP	Meteorology-Chemistry Interface Processor
MMS	Minerals Management Service (now known as the Bureau of Energy Management, Regulation and Enforcement (BOEMRE))
MOVES	Motor Vehicle Emissions Simulator
MSA	Metropolitan Statistical Area
MTBE	Methyl tert-butyl ether
MWC	Municipal waste combustor
MY	Model year
NAAQS	National Ambient Air Quality Standards
NAICS	North American Industry Classification System
NBAFM	Naphthalene, Benzene, Acetaldehyde, Formaldehyde and Methanol
NCAR	National Center for Atmospheric Research
NEEDS	National Electric Energy Database System
NEI	National Emission Inventory
NESCAUM	Northeast States for Coordinated Air Use Management
NH₃	Ammonia
NLCD	National Land Cover Database
NOAA	National Oceanic and Atmospheric Administration
NONROAD	OTAQ's model for estimation of nonroad mobile emissions
NO_x	Nitrogen oxides
NSPS	New Source Performance Standards
OHH	Outdoor Hydronic Heater
ONI	Off network idling
OTAQ	EPA's Office of Transportation and Air Quality
ORIS	Office of Regulatory Information System
ORD	EPA's Office of Research and Development
OSAT	Ozone Source Apportionment Technology

pcSOA	Potential combustion Secondary Organic Aerosol
PFC	Portable Fuel Container
PM_{2.5}	Particulate matter less than or equal to 2.5 microns
PM₁₀	Particulate matter less than or equal to 10 microns
POA	Primary Organic Aerosol
ppm	Parts per million
ppmv	Parts per million by volume
PSAT	Particulate Matter Source Apportionment Technology
RACT	Reasonably Available Control Technology
RBT	Refinery to Bulk Terminal
RIA	Regulatory Impact Analysis
RICE	Reciprocating Internal Combustion Engine
RWC	Residential Wood Combustion
RPD	Rate-per-vehicle (emission mode used in SMOKE-MOVES)
RPH	Rate-per-hour for hoteling (emission mode used in SMOKE-MOVES)
RPHO	Rate-per-hour for off-network idling (emission mode used in SMOKE-MOVES)
RPP	Rate-per-profile (emission mode used in SMOKE-MOVES)
RPS	Rate-per-start (emission mode used in SMOKE-MOVES)
RPV	Rate-per-vehicle (emission mode used in SMOKE-MOVES)
RVP	Reid Vapor Pressure
SCC	Source Classification Code
SMARTFIRE2	Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation version 2
SMOKE	Sparse Matrix Operator Kernel Emissions
SO₂	Sulfur dioxide
SOA	Secondary Organic Aerosol
SIP	State Implementation Plan
SPDPRO	Hourly Speed Profiles for weekday versus weekend
S/L/T	state, local, and tribal
TAF	Terminal Area Forecast
TCEQ	Texas Commission on Environmental Quality
TOG	Total Organic Gas
TSD	Technical support document
USDA	United States Department of Agriculture
VIIRS	Visible Infrared Imaging Radiometer Suite
VOC	Volatile organic compounds
VMT	Vehicle miles traveled
VPOP	Vehicle Population
WRAP	Western Regional Air Partnership
WRF	Weather Research and Forecasting Model
2014NEIv2	2014 National Emissions Inventory (NEI), version 2

1 Introduction

The U.S. Environmental Protection Agency (EPA) developed an air quality modeling platform for air toxics and criteria air pollutants that represents the year 2021. The platform is based on the 2020 National Emissions Inventory (2020 NEI) published in April 2023 (EPA, 2023) along with other data specific to the year 2021. The air quality modeling platform consists of all the emissions inventories and ancillary data files used for emissions modeling, as well as the meteorological, initial condition, and boundary condition files needed to run the air quality model. This document focuses on the emissions modeling component of the 2021 modeling platform, including the emission inventories, the ancillary data files, and the approaches used to transform inventories for use in air quality modeling.

The modeling platform includes all criteria air pollutants and precursors (CAPs), two groups of hazardous air pollutants (HAPs) and diesel particulate matter. The first group of HAPs are those explicitly used by the chemical mechanism in the Community Multiscale Air Quality (CMAQ) model (Appel, 2018) for ozone/particulate matter (PM): chlorine (Cl), hydrogen chloride (HCl), naphthalene, benzene, acetaldehyde, formaldehyde, and methanol (the last five are abbreviated as NBAFM in subsequent sections of the document). The second group of HAPs consists of 52 HAPs or HAP groups (such as polycyclic aromatic hydrocarbon groups) that are included in CMAQ for the purposes of air quality modeling for a HAP+CAP platform.

Emissions were prepared for the Community Multiscale Air Quality (CMAQ) model version 5.4,² which was used to model ozone (O₃) particulate matter (PM), and HAPs. CMAQ requires hourly and gridded emissions of the following inventory pollutants: carbon monoxide (CO), nitrogen oxides (NO_x), volatile organic compounds (VOC), sulfur dioxide (SO₂), ammonia (NH₃), particulate matter less than or equal to 10 microns (PM₁₀), and individual component species for particulate matter less than or equal to 2.5 microns (PM_{2.5}). In addition, the Carbon Bond mechanism version 6 (CB6) with chlorine chemistry within CMAQ allows for explicit treatment of the VOC HAPs naphthalene, benzene, acetaldehyde, formaldehyde and methanol (NBAFM), includes anthropogenic HAP emissions of HCl and Cl, and can model additional HAPs as described in Section 3. The short abbreviation for the modeling case name was “2021hb”, where 2021 is the year modeled, ‘h’ represents that it was based on the 2020 NEI, and ‘b’ represents that it was the second version of a 2020 NEI-based platform.

Emissions were also prepared for an air dispersion modeling system: American Meteorological Society/Environmental Protection Agency Regulatory Model (AERMOD) (EPA, 2018). AERMOD was run for 2021 for all NEI HAPs (about 130 more than covered by CMAQ) over the continental U.S. in a similar way as was done for the 2018 version of AirToxScreen (EPA, 2022a). This TSD focuses on the CMAQ aspects of the 2021 emissions modeling platform from which ozone and PM data were also developed for the Centers for Disease Control and Prevention. The effort to create the emission inputs for this study included development of emission inventories to represent emissions during the year of 2021, along with application of emissions modeling tools to convert the inventories into the format and resolution needed by CMAQ and AERMOD.

² CMAQ version 5.4: <https://zenodo.org/record/7218076>. CMAQ is also available from <https://www.epa.gov/cmaq> and the Community Modeling and Analysis System (CMAS) Center at: <https://www.cmascenter.org>.

The emissions modeling platform includes point sources, nonpoint sources, onroad mobile sources, nonroad mobile sources, biogenic emissions and fires for the U.S., Canada, and Mexico. Some platform categories use more disaggregated data than are made available in the NEI. For example, in the platform, onroad mobile source emissions are represented as hourly emissions by vehicle type, fuel type process and road type while the NEI emissions are aggregated to vehicle type/fuel type totals and annual temporal resolution. Emissions used in the CMAQ modeling from Canada are provided by Environment and Climate Change Canada (ECC) and Mexico are mostly provided by SEMARNAT and are not part of the NEI. Year-specific emissions were used for fires, biogenic sources, fertilizer, point sources, and onroad and nonroad mobile sources. Where available, hourly continuous emission monitoring system (CEMS) data were used for electric generating unit (EGU) emissions.

The primary emissions modeling tool used to create the CMAQ model-ready emissions was the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system. SMOKE version 5.0 was used to create CMAQ-ready emissions files for a 12-km grid covering the continental U.S. Additional information about SMOKE is available from <http://www.cmascenter.org/smoke>.

The gridded meteorological model used to provide input data for the emissions modeling was developed using the Weather Research and Forecasting Model (WRF, <https://ral.ucar.edu/solutions/products/weather-research-and-forecasting-model-wrf>) version 4.1.1, Advanced Research WRF core (Skamarock, et al., 2008). The WRF Model is a mesoscale numerical weather prediction system developed for both operational forecasting and atmospheric research applications. The WRF model was run for 2021 over a domain covering the continental U.S. at a 12km resolution with 35 vertical layers. The run for this platform included high resolution sea surface temperature data from the Group for High Resolution Sea Surface Temperature (GHR SST) (see <https://www.ghrsst.org/>) and is given the EPA meteorological case abbreviation “21k.” The full case abbreviation includes this suffix following the emissions portion of the case name to fully specify the abbreviation of the case as “2021hb_cb6_21k.”

In support of AirToxScreen, CMAQ and AERMOD were run with the prepared emissions for each of the four modeling domains. CMAQ outputs provide the overall mass, chemistry and formation for specific hazardous air pollutants (HAPs) formed secondarily in the atmosphere (e.g., formaldehyde, acetaldehyde, and acrolein), whereas AERMOD provides spatial granularity and more detailed source attribution. CMAQ also provided the biogenic and fire concentrations, as these sources were not run in AERMOD. Special steps were taken to estimate secondary HAPs, fire and biogenic emissions in these areas. The outputs from CMAQ and AERMOD were combined to provide spatially refined concentration estimates for HAPs, from which estimates of cancer and non-cancer risk were derived. Data files and summaries for this platform are available from this section of the air emissions modeling website <https://www.epa.gov/air-emissions-modeling/2021-emissions-modeling-platform>.

This document contains four additional sections. Section 2 describes the emission inventories input to SMOKE. Section 3 describes the emissions modeling and the ancillary files used to process the emission inventories into air quality model-ready inputs. Data summaries are provided in Section 4, and Section 5 provides references.

2 Emissions Inventories and Approaches

This section describes the emissions inventories created for input to SMOKE, which are based on the April 2023 version of the 2020 NEI with updates to reflect emissions in the year 2021. The NEI includes five main data categories: a) nonpoint sources; b) point sources; c) nonroad mobile sources; d) onroad mobile sources; and e) fires. For CAPs, the NEI data are largely compiled from data submitted by state, local and tribal (S/L/T) agencies. HAP emissions data are often augmented by EPA when they are not voluntarily submitted to the NEI by S/L/T agencies. The NEI was compiled using the Emissions Inventory System (EIS). EIS collects and stores facility inventory and emissions data for the NEI and includes hundreds of automated QA checks to improve data quality, and it also supports release point (stack) coordinates separately from facility coordinates. EPA collaboration with S/L/T agencies helped prevent duplication between point and nonpoint source categories such as industrial boilers. The 2020 NEI Technical Support Document describes in detail the development of the 2020 emission inventories and is available at <https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-technical-support-document-tsd> (EPA, 2023).

A complete set of emissions for all source categories is developed for the NEI every three years, with 2020 being the most recent year represented with a full “triennial” NEI. S/L/T agencies are required to submit all applicable point sources to the NEI in triennial years, including the year 2020. Because only point source emissions were submitted by S/L/T agencies to develop the NEI for 2021, emissions for any point sources not submitted for 2021, and not marked as shutdown, were pulled forward from the 2020 NEI. The SMARTFIRE2 system and the BlueSky Pipeline (<https://github.com/pnwairfire/bluesky>) emissions modeling system were used to develop the fire emissions. SMARTFIRE2 categorizes all fires as either prescribed burning or wildfire, and the BlueSky Pipeline system includes fuel loading, consumption and emission factor estimates for both types of fires. Onroad and nonroad mobile source emissions were developed for this project using MOVES4 (<https://www.epa.gov/moves>).

With the exception of fire emissions, Canadian emissions were provided by Environment Canada and Climate Change (ECCC) for the years 2020 and 2023 and most 2021 emissions were developed by interpolating between 2020 and 2023. For Mexico, inventories from the 2019 emissions modeling platform (EPA, 2022b) were used as the starting point with data for border states supplemented with data for 2018 developed by SEMARNAT in collaboration with U.S. EPA.

The emissions modeling process was performed using SMOKE v5.0. Through this process, the emissions inventories were apportioned into the grid cells used by CMAQ and temporally allocated into hourly values. In addition, the pollutants in the inventories (e.g., NO_x, PM and VOC) were split into the chemical species needed by CMAQ. For the purposes of preparing the CMAQ-ready emissions, the NEI emissions inventories by data category were split into emissions modeling platform “sectors”; and emissions from sources other than the NEI are added, such as the Canadian, Mexican, and offshore inventories. Emissions within the emissions modeling platform were separated into sectors for groups of related emissions source categories that were run through the appropriate SMOKE programs, except the final merge, independently from emissions categories in the other sectors. The final merge program called Mrggrid combines low-level sector-specific gridded, speciated and temporalized emissions to create the final CMAQ-ready emissions inputs. For biogenic and fertilizer emissions, the CMAQ model allows for these emissions to be included in the CMAQ-ready emissions inputs, or to be computed within CMAQ

itself (the “inline” option). This study used the option to compute biogenic emissions within the model and the CMAQ bidirectional ammonia process to compute the fertilizer emissions.

Table 2-1 presents the sectors in the emissions modeling platform used to develop the year 2021 emissions for this project. The sector abbreviations are provided in italics; these abbreviations are used in the SMOKE modeling scripts, the inventory file names, and throughout the remainder of this section.

Table 2-1. Platform sectors used in the Emissions Modeling Process

Platform Sector: <i>abbreviation</i>	NEI Data Category	Description and resolution of the data input to SMOKE
EGU units: <i>Ptegu</i>	Point	2021 NEI point source EGUs, replaced with hourly Continuous Emissions Monitoring System (CEMS) values for NO _x and SO ₂ , and the remaining pollutants temporally allocated according to CEMS heat input where the units are matched to the NEI. Emissions for all sources not matched to CEMS data come from 2021 NEI point inventory. Annual resolution for sources not matched to CEMS data, hourly for CEMS sources. EGUs closed in 2021 are not part of the inventory.
Point source oil and gas: <i>pt_oilgas</i>	Point	2021 NEI point sources that include oil and gas production emissions processes for facilities with North American Industry Classification System (NAICS) codes related to Oil and Gas Extraction, Natural Gas Distribution, Drilling Oil and Gas Wells, Support Activities for Oil and Gas Operations, Pipeline Transportation of Crude Oil, and Pipeline Transportation of Natural Gas. Includes U.S. offshore oil production.
Aircraft and ground support equipment: <i>airports</i>	Point	2021 NEI point source emissions from airports, including aircraft and airport ground support emissions projected to 2021 based on the 2022 Terminal Area Forecast (TAF). Annual resolution.
Remaining non-EGU point: <i>Ptnonipm</i>	Point	All 2021 NEI point source records not matched to the airports, ptegu, or pt_oilgas sectors. Includes 2020 NEI rail yard emissions projected to 2021. Annual resolution.
Livestock: <i>Livestock</i>	Nonpoint	2021 nonpoint livestock emissions developed using a similar method to 2020 NEI but with adjusted animal counts and using 2021 meteorology. Livestock includes ammonia and other pollutants (except PM _{2.5}). County and annual resolution.
Agricultural Fertilizer: <i>fertilizer</i>	Nonpoint	2021 agricultural fertilizer ammonia emissions computed inline within CMAQ.

Platform Sector: <i>abbreviation</i>	NEI Data Category	Description and resolution of the data input to SMOKE
Area fugitive dust: <i>afdust_adj</i>	Nonpoint	PM ₁₀ and PM _{2.5} fugitive dust sources from the 2020 NEI nonpoint inventory; including building construction, road construction, agricultural dust, and paved and unpaved road dust where paved and unpaved road dust were adjusted to 2021 based on VMT differences. The emissions modeling system applies a transportable fraction reduction and zero-out adjustments based on the year-specific gridded hourly meteorology (precipitation and snow/ice cover). Emissions are county and annual resolution.
Biogenic: <i>beis</i>	Nonpoint	Year 2021 emissions from biogenic sources. These were left out of the CMAQ-ready merged emissions, in favor of inline biogenic emissions produced during the CMAQ model run itself. Version 4 of the Biogenic Emissions Inventory System (BEIS) was used with Version 6 of the Biogenic Emissions Landuse Database (BELD6). The CMAQ-generated emissions are similar to the 2021 biogenic emissions generated through running SMOKE, but they are not exactly the same.
Category 1, 2 CMV: <i>cmv_c1c2</i>	Nonpoint	2021 Category 1 (C1) and Category 2 (C2), commercial marine vessel (CMV) emissions based on 2021 Automatic Identification System (AIS) data categorized using SCCs specific to ship type. Point and hourly resolution.
Category 3 CMV: <i>cmv_c3</i>	Nonpoint	2021 Category 3 (C3) commercial marine vessel (CMV) emissions based on 2021 AIS data categorized using SCCs specific to ship type. Point and hourly resolution.
Locomotives : <i>rail</i>	Nonpoint	Line haul rail locomotives emissions from 2020 NEI projected to 2021 using 5 percent growth based on Annual Energy Outlook (AEO) changes from 2020 to 2021. County and annual resolution.
Nonpoint source oil and gas: <i>np_oilgas</i>	Nonpoint	Nonpoint emissions from oil and gas-related processes for 2021 computed using activity data for 2021. County and annual resolution.
Residential Wood Combustion: <i>rwc</i>	Nonpoint	2020 NEI nonpoint sources with residential wood combustion (RWC) processes, projected to 2021 with state-level adjustment factors derived from the State Energy Data System (SEDS). County and annual resolution.
Solvents: <i>np_solvents</i>	Nonpoint	Emissions of solvents for 2021 based on methods used for the 2020 NEI (Seltzer, 2021). Includes household cleaners, personal care products, adhesives, architectural and aerosol coatings, printing inks, and pesticides. Annual and county resolution.
Remaining nonpoint: <i>nonpt</i>	Nonpoint	2020 NEI nonpoint sources not included in other platform sectors. County and annual resolution.

Platform Sector: <i>abbreviation</i>	NEI Data Category	Description and resolution of the data input to SMOKE
Nonroad: <i>nonroad</i>	Nonroad	2021 nonroad equipment emissions developed with MOVES4, including the updates made to spatial apportionment that were developed with the 2016v1 platform. MOVES4 was used for all states except California, which submitted their own emissions for 2020 and 2023 from which an interpolation to 2021 was performed. County and monthly resolution.
Onroad: <i>onroad</i>	Onroad	Onroad mobile source gasoline and diesel vehicles from parking lots and moving vehicles for 2021 developed using VMT data from 2020 NEI projected to 2021 using factors based on FHWA VM-2 data. Includes the following emission processes: exhaust, extended idle, auxiliary power units, evaporative, permeation, refueling, vehicle starts, off network idling, long-haul truck hoteling, and brake and tire wear. MOVES4 was run for 2021 to generate year-specific emission factors.
Onroad California: <i>onroad_ca_adj</i>	Onroad	California-provided 2020 and 2023 CAPs that were interpolated to 2021. HAPs speciated from CAPs. Onroad mobile source gasoline and diesel vehicles from parking lots and moving vehicles based on Emission Factor (EMFAC), gridded and temporalized based on outputs from MOVES4.
Point source agricultural fires: <i>ptagfire</i>	Nonpoint	Agricultural fire sources for 2021 developed by EPA as point and day-specific emissions. ³ Only EPA-developed data were used in this study, thus 2020 NEI state submissions are not included. Agricultural fires are in the nonpoint data category of the NEI, but in the modeling platform, they are treated as day-specific point sources. Updated HAP-augmentation factors were applied.
Point source prescribed fires: <i>ptfire-rx</i>	Nonpoint	Point source day-specific prescribed fires for 2021 computed using SMARTFIRE 2 and BlueSky Pipeline. The ptfire emissions were run as two separate sectors: ptfire-rx (prescribed, including Flint Hills / grasslands) and ptfire-wild.
Point source wildfires: <i>ptfire-wild</i>	Nonpoint	Point source day-specific wildfires for 2021 computed using SMARTFIRE 2 and BlueSky Pipeline.
Non-US. Fires: <i>ptfire_othna</i>	N/A	Point source day-specific wildfires and agricultural fires outside of the U.S. for 2021. Canadian fires were computed using SMARTFIRE 2 and BlueSky Pipeline. Mexico, Caribbean, Central American, and other international fires, are from v2.5 of the Fire INventory (FINN) from National Center for Atmospheric Research (Wiedinmyer, C., 2023).
Canada Area Fugitive dust sources: <i>canada_afdust</i>	N/A	Area fugitive dust sources from ECCC for 2021 (interpolated between provided 2020 and 2023 emisissions) with transport fraction and snow/ice adjustments based on 2021 meteorological data. Annual and province resolution.

³ Only EPA-developed agricultural fire data were included in this study; data submitted by states to the NEI were excluded.

Platform Sector: <i>abbreviation</i>	NEI Data Category	Description and resolution of the data input to SMOKE
Canada Point Fugitive dust sources: <i>canada_ptdust</i>	N/A	Point source fugitive dust sources from ECCC for 2021 (interpolated between provided 2020 and 2023 emissions) with transport fraction and snow/ice adjustments based on 2021 meteorological data. Monthly and province resolution.
Canada and Mexico stationary point sources: <i>canmex_point</i>	N/A	Canada and Mexico point source emissions not included in other sectors. Canada point sources were provided by ECCC for 2020 and 2023, and interpolated to 2021. Mexico point source emissions for border states represent 2018 and were developed by SEMARNAT in collaboration with EPA, while emissions for all other states were carried forward from 2019ge (EPA, 2022b). Annual and monthly resolution.
Canada and Mexico agricultural sources: <i>canmex_ag</i>	N/A	Canada and Mexico agricultural emissions. Canada emissions were provided by ECCC for 2020 and 2023, and interpolated to 2021. Mexico agricultural emissions were provided by SEMARNAT and include updated emissions for border states representing 2018 developed by SEMARNAT in collaboration with EPAT, while emissions for all other states were carried forward from 2019ge. Annual resolution.
Canada low-level oil and gas sources: <i>canada_og2D</i>	N/A	Canada emissions from upstream oil and gas, provided by ECCC for 2020 and 2023, and interpolated to 2021. This sector contains the portion of oil and gas emissions which are not subject to plume rise. The rest of the Canada oil and gas emissions are in the canmex_point sector. Annual resolution.
Canada and Mexico nonpoint and nonroad sources: <i>canmex_area</i>	N/A	Canada and Mexico nonpoint source emissions not included in other sectors. Canada: ECCC provided surrogates and 2020 and 2023 inventories, that were interpolated to 2021. Mexico: include updated emissions for border states representing 2018 developed by SEMARNAT in collaboration with EPA, while emissions for all other states were carried forward from 2019ge. Annual and monthly resolution.
Canada onroad sources: <i>canada_onroad</i>	N/A	Canada onroad emissions. 2020 and 2023 Canada inventories provided by ECCC and interpolated to 2021; processed using updated surrogates. Province and monthly resolution.
Mexico onroad sources: <i>mexico_onroad</i>	N/A	Mexico onroad emissions. 2020 and 2023 emissions output from MOVES-Mexico were interpolated to 2021. Municipio and monthly resolution.

Ocean chlorine emissions were also merged in with the above sectors. The ocean chlorine gas emission estimates are based on the build-up of molecular chlorine (Cl₂) concentrations in oceanic air masses (Bullock and Brehme, 2002). Ocean chlorine data at 12 km resolution were available from earlier studies

and were not modified other than the name “CHLORINE” was changed to “CL2” because that is the name required by the CMAQ model.

The emission inventories in SMOKE input formats for the platform are available from EPA’s Air Emissions Modeling website: <https://www.epa.gov/air-emissions-modeling/2021-emissions-modeling-platform>. The platform informational text file indicates the zipped files associated with each platform sector. Some emissions data summaries are available with the data files for the 2021 platform. The types of reports include state summaries of inventory pollutants and model species by modeling platform sector and county annual totals by modeling platform sector.

2.1 Point sources (*ptegu, pt_oilgas, ptnonipm, airports*)

Point sources are sources of emissions for which specific geographic coordinates (e.g., latitude/longitude) are specified, as in the case of an individual facility. A facility may have multiple emission release points that may be characterized as units such as boilers, reactors, spray booths, kilns, etc. A unit may have multiple processes (e.g., a boiler that sometimes burns residual oil and sometimes burns natural gas). With a couple of minor exceptions, this section describes only NEI point sources within the contiguous U.S. The offshore oil platform (*pt_oilgas* sector) and CMV emissions (*cmv_c1c2* and *cmv_c3* sectors) are processed by SMOKE as point source inventories and are discussed later in this section. A complete NEI is developed every three years. At the time of this writing, 2020 is the most recently finished complete NEI. A comprehensive description about the development of the 2020 NEI is available in the 2020 NEI TSD (EPA, 2023). Point inventories are also available in EIS for non-triennial NEI years such as 2019 and 2021. In the interim year point inventories, states are required to update large sources with the emissions that occurred in that year, while sources not updated by states for the interim year were either carried forward from the most recent triennial NEI or marked as closed and removed.

In preparation for modeling, the complete set of point sources in the NEI was exported from EIS for the year 2021 into the Flat File 2010 (FF10) format that is compatible with SMOKE (see <https://cmascenr.org/smoke/documentation/4.9/html/ch06s02s08.html>) and was then split into several sectors for modeling. For both flat files, sources without specific locations (i.e., the FIPS code ends in 777) were dropped and inventories for the other point source sectors were created from the remaining point sources. The point sectors are: EGUs (*ptegu*), point source oil and gas extraction-related sources (*pt_oilgas*), airport emissions (*airports*), and the remaining non-EGUs (*ptnonipm*). The EGU emissions were split out from the other sources to facilitate the use of distinct SMOKE temporal processing and future-year projection techniques. The oil and gas sector emissions (*pt_oilgas*) and airport emissions (*airports*) were processed separately for the purposes of developing emissions summaries and due to distinct projection techniques from the remaining non-EGU emissions (*ptnonipm*), although this study does not include emissions projected to other years.

In some cases, data about facility or unit closures are entered into EIS after the inventory modeling inventory flat files have been extracted. Prior to processing through SMOKE, submitted facility and unit closures were reviewed and where closed sources were found in the inventory, those were removed.

For the 2021 platform, an analysis of point source stack parameters (e.g., stack height, diameter, temperature, and velocity) was performed due to the presence of unrealistic and repeated stack

parameters. The defaulted values were noticed in data submissions for the states of Illinois, Louisiana, Michigan, Pennsylvania, Texas, and Wisconsin. Where these defaults were detected and deemed to be unreasonable for the specific process, the affected stack parameters were replaced by values from the PSTK file that is input to SMOKE. PSTK contains default stack parameters by source classification code (SCC). These updates impacted the ptnonipm and pt_oilgas inventories.

The inventory pollutants processed through SMOKE for input to CMAQ for the ptegu, pt_oilgas, ptnonipm, and airports sectors included: CO, NO_x, VOC, SO₂, NH₃, PM₁₀, and PM_{2.5} and the following HAPs: HCl (pollutant code = 7647010), Cl (code = 7782505), and several dozen other HAPs listed in Section 3. NBAFM pollutants from the point sectors were utilized. For AERMOD, additional HAPs were included as described in the 2021 AirToxScreen TSD.

The ptnonipm, pt_oilgas, and airports sector emissions were provided to SMOKE as annual emissions. For sources in the ptegu sector that could be matched to 2021 CEMS data, hourly CEMS NO_x and SO₂ emissions for 2021 from EPA's Acid Rain Program were used rather than annual inventory emissions. For all other pollutants (e.g., VOC, PM_{2.5}, HCl), annual emissions were used as-is from the annual inventory but were allocated to hourly values using heat input from the CEMS data. For the unmatched units in the ptegu sector, annual emissions were allocated to daily values using IPM region- and pollutant-specific profiles, and similarly, region- and pollutant-specific diurnal profiles were applied to create hourly emissions.

The non-EGU stationary point source (ptnonipm) emissions were used as inputs to SMOKE as annual emissions. The full description of how the NEI emissions were developed is provided in the NEI documentation - a brief summary of their development follows:

- a. CAP and HAP data were provided by States, locals and tribes under the Air Emissions Reporting Rule (AERR) [the reporting size threshold is larger for inventory years between the triennial inventory years of 2011, 2014, 2017, 2020, ...].
- b. EPA corrected known issues and filled PM data gaps.
- c. EPA added HAP data from the Toxic Release Inventory (TRI) where corresponding data was not already provided by states/locals.
- d. EPA stored and applied matches of the point source units to units with CEMS data and also for all EGU units modeled by EPA's Integrated Planning Model (IPM).
- e. Data for airports and rail yards were incorporated.
- f. Off-shore platform data were added from the Bureau of Ocean Energy Management (BOEM).

The changes made to the NEI point sources prior to modeling with SMOKE are as follows:

- The tribal data, which do not use state/county Federal Information Processing Standards (FIPS) codes in the NEI, but rather use the tribal code, were assigned a state/county FIPS code of 88XXX, where XXX is the 3-digit tribal code in the NEI. This change was made because SMOKE requires all sources to have a state/county FIPS code.

- Sources that did not have specific counties assigned (i.e., the county code ends in 777) were not included in the modeling because it was only possible to know the state in which the sources resided, but no more specific details related to the location of the sources were available.

Each of the point sectors is processed separately through SMOKE as described in the following subsections.

2.1.1 EGU sector (ptegu)

The ptegu sector contains emissions from EGUs in the 2021 point source inventory that could be matched to units found in the National Electric Energy Database System (NEEDS) v6 that is used by the Integrated Planning Model (IPM) to develop projected EGU emissions. It was necessary to put these EGUs into a separate sector in the platform because EGUs use different temporal profiles than other sources in the point sector and it is useful to segregate these emissions from the rest of the point sources to facilitate summaries of the data. Sources not matched to units found in NEEDS were placed into the pt_oilgas or ptnonipm sectors. For studies that include analytic years, the sources in the ptegu sector are fully replaced with the emissions output from IPM. It is therefore important that the matching between the NEI and NEEDS database be as complete as possible because there can be double-counting of emissions in analytic year modeling scenarios if emissions for units projected by IPM are not properly matched to the units in the base year point source inventory.

The 2021 ptegu emissions inventory is a subset of the point source flat file exported from the Emissions Inventory System (EIS). In the point source flat file, emission records for sources that have been matched to the NEEDS database have a value filled into the IPM_YN column based on the matches stored within EIS. Thus, unit-level emissions were split into a separate EGU flat file for units that have a populated (non-null) ipm_yn field. A populated ipm_yn field indicates that a match was found for the EIS unit in the NEEDS v6 database. Updates were made to the flat file output from EIS as follows:

- ORIS facility and unit identifiers were updated based on additional matches in a cross-platform spreadsheet, based on state comments, and using the EIS alternate identifiers table as described later in this section.

Some units in the ptegu sector are matched to Continuous Emissions Monitoring System (CEMS) data via Office of Regulatory Information System (ORIS) facility codes and boiler IDs. For the matched units, the annual emissions of NO_x and SO₂ in the flat file were replaced with the hourly CEMS emissions in base year modeling. For other pollutants at matched units, the hourly CEMS heat input data were used to allocate the NEI annual emissions to hourly values. All stack parameters, stack locations, and Source Classification Codes (SCC) for these sources come from the flat file. If CEMS data exists for a unit, but the unit is not matched to the NEI, the CEMS data for that unit were not used in the modeling platform. However, if the source exists in the NEI and is not matched to a CEMS unit, the emissions from that source are still modeled using the annual emission value in the NEI temporally allocated to hourly values.

EIS stores many matches from NEI units to the ORIS facility codes and boiler IDs used to reference the CEMS data. In the flat file, emission records for point sources matched to CEMS data have values filled into the ORIS_FACILITY_CODE and ORIS_BOILER_ID columns. The CEMS data are available at

<https://campd.epa.gov/data>. Many smaller emitters in the CEMS program cannot be matched to the NEI due to differences in the way a unit is defined between the NEI and CEMS datasets, or due to uncertainties in source identification such as inconsistent plant names in the two data systems. In addition, the NEEDS database of units modeled by IPM includes many smaller emitting EGUs that do not have CEMS. Therefore, there will be more units in the ptegu sector than have CEMS data.

Matches from the NEI to ORIS codes and the NEEDS database were improved in the platform where applicable. In some cases, NEI units in EIS match to many CAMD units. In these cases, a new entry was made in the flat file with a “_M_” in the ipm_yn field of the flat file to indicate that there are “multiple” ORIS IDs that match that unit. This helps facilitate appropriate temporal allocation of the emissions by SMOKE. Temporal allocation for EGUs is discussed in more detail in the Ancillary Data section below.

The EGU flat file was split into two flat files: those that have unit-level matches to CEMS data using the oris_facility_code and oris_boiler_id fields and those that do not so that different temporal profiles could be applied. In addition, the hourly CEMS data were processed through v2.1 of the CEMCorrect tool to mitigate the impact of unmeasured values in the data.

2.1.2 Point source oil and gas sector (pt_oilgas)

The pt_oilgas sector was separated from the ptnonipm sector by selecting sources with specific North American Industry Classification System (NAICS) codes shown in Table 2-2. The emissions and other source characteristics in the pt_oilgas sector are submitted by states, while EPA developed a dataset of nonpoint oil and gas emissions for each county in the U.S. with oil and gas activity that was available for states to use. Nonpoint oil and gas emissions can be found in the np_oilgas sector. The pt_oilgas sector includes emissions from offshore oil platforms. Where available, the point source emissions submitted as part of the 2021 NEI process were used. Sources without data submitted for 2021 were projected to 2021 from 2020 NEI emissions.

Table 2-2. Point source oil and gas sector NAICS Codes

NAICS	NAICS description
2111	Oil and Gas Extraction
211112	Natural Gas Liquid Extraction (although no emissions for this NAICS code exist in the 2021 inventory)
21112	Crude Petroleum Extraction
211120	Crude Petroleum Extraction
21113	Natural Gas Extraction
211130	Natural Gas Extraction
213111	Drilling Oil and Gas Wells
213112	Support Activities for Oil and Gas Operations
2212	Natural Gas Distribution
22121	Natural Gas Distribution
221210	Natural Gas Distribution
237120	Oil and Gas Pipeline and Related Structures Construction
4861	Pipeline Transportation of Crude Oil
48611	Pipeline Transportation of Crude Oil

NAICS	NAICS description
486110	Pipeline Transportation of Crude Oil
4862	Pipeline Transportation of Natural Gas
48621	Pipeline Transportation of Natural Gas
486210	Pipeline Transportation of Natural Gas

Information on the development of the 2020 NEI oil and gas emissions can be found in Section 13 of the 2020 NEI TSD. The point oil and gas emissions for 2021 by state are shown in Table 2-3.

Table 2-3. Point source oil and gas sector emissions for 2021

State	2021 NO_x	2021 VOC
Alabama	9,694	1,054
Alaska	38,833	1,608
Arizona	2,460	183
Arkansas	3,235	262
California	2,623	2,529
Colorado	14,162	11,714
Connecticut	62	40
Delaware	6	1
Florida	5,409	623
Georgia	2,763	436
Idaho	1,057	30
Illinois	3,925	1,034
Indiana	1,601	185
Iowa	3,804	208
Kansas	15,641	2,982
Kentucky	10,000	1,398
Louisiana	28,478	8,149
Maine	30	58
Maryland	175	176
Massachusetts	183	64
Michigan	9,207	1,076
Minnesota	2,175	139
Mississippi	20,986	1,926
Missouri	1,847	78
Montana	714	972
Nebraska	2,597	218
Nevada	233	19
New Jersey	83	98
New Mexico	32,396	52,754
New York	905	264
North Carolina	1,776	220

State	2021 NO _x	2021 VOC
North Dakota	4,556	2,799
Ohio	9,714	1,481
Oklahoma	34,400	25,119
Oregon	792	63
Pennsylvania	3,798	1,022
Puerto Rico	17	0
Rhode Island	50	22
South Carolina	279	104
South Dakota	358	10
Tennessee	6,251	514
Texas	44,710	19,906
Utah	2,359	488
Virginia	2,621	406
Washington	688	45
West Virginia	8,274	3,230
Wisconsin	376	223
Wyoming	13,327	49,362
Offshore	34,660	31,406
Tribal Data	7,909	2,254

2.1.3 Aircraft and ground support equipment (airports)

Emissions at airports were separated from other sources in the point inventory based on sources that have the facility source type of 100 (airports). The airports sector includes all aircraft types used for public, private, and military purposes and aircraft ground support equipment. The Federal Aviation Administration's (FAA) Aviation Environmental Design Tool (AEDT) is used to estimate emissions for this sector. Additional information about aircraft emission estimates can be found in section 3 of the 2020 NEI TSD. Airport emissions were projected from the 2020 NEI to 2021 using factors derived from the 2023 Terminal Area Forecast (TAF)⁴ data. EPA used airport-specific factors where available. Airport emissions were spread out into multiple 12km grid cells when the airport runways were determined to overlap multiple grid cells. Otherwise, airport emissions for a specific airport are confined to one air quality model grid cell. The SCCs included in the airport sector are shown in Table 2-4.

Table 2-4. SCCs for the airports sector

SCC	Tier 1 description	Tier 2 description	Tier 3 description	Tier 4 description
2275001000	Mobile Sources	Aircraft	Military Aircraft	Total
2275020000	Mobile Sources	Aircraft	Commercial Aircraft	Total: All Types
2275050011	Mobile Sources	Aircraft	General Aviation	Piston
2275050012	Mobile Sources	Aircraft	General Aviation	Turbine
2275060011	Mobile Sources	Aircraft	Air Taxi	Piston

⁴ See https://www.faa.gov/data_research/aviation/taf for 2023 TAF released in January 2024

SCC	Tier 1 description	Tier 2 description	Tier 3 description	Tier 4 description
2275060012	Mobile Sources	Aircraft	Air Taxi	Turbine

2.1.4 Non-IPM sector (ptnonipm)

With some exceptions, the ptnonipm sector contains the point sources that are not in the ptegu, pt_oilgas, or airports sectors. For the most part, the ptnonipm sector reflects non-EGU emissions sources and rail yards. However, it is possible that some low-emitting EGUs not matched to units the NEEDS database or to CEMS data are in the ptnonipm sector.

The ptnonipm sector contains a small amount of fugitive dust PM emissions from vehicular traffic on paved or unpaved roads at industrial facilities, coal handling at coal mines, and grain elevators. Sources with state/county FIPS code ending with “777” are in the NEI but are not included in any modeling sectors. These sources typically represent mobile (temporary) asphalt plants that are only reported for some states and are generally in a fixed location for only a part of the year and are therefore difficult to allocate to specific places and days as is needed for modeling. Therefore, these sources are dropped from the point-based sectors in the modeling platform.

The ptnonipm sources (i.e., not EGUs and non -oil and gas sources) were used as-is from the 2021 NEI point inventory. Solvent emissions from point sources were removed from the np_solvents sector to prevent double-counting, so that all point sources can be retained in the modeling as point sources rather than as area sources. The modeling was based the point flat file exported from EIS on January 28, 2023 with edits made through April 14, 2023 that included corrections to how the selection was implemented in EIS, updates from the state/local review, and updates specific to ethylene oxide. The np_solvents sector is described in more detail in Section 2.2.6.

Emissions from rail yards are included in the ptnonipm sector. Railyards are from the 2020 NEI railyard inventory with a projection factor applied. Additional information about railyard estimates can be found in section 3 of the 2020 NEI TSD.

2.2 Nonpoint sources (*afdust, fertilizer, livestock, np_oilgas, rwc, vcp, nonpt*)

This section describes the *stationary* nonpoint sources in the NEI nonpoint data category. Locomotives, C1 and C2 CMV, and C3 CMV are included in the NEI nonpoint data category but are mobile sources that are described in Section 2.4. The 2020 NEI TSD includes documentation for the nonpoint data.

Nonpoint tribal emissions submitted to the NEI are dropped during spatial processing with SMOKE due to the configuration of the spatial surrogates. Part of the reason for this is to prevent possible double-counting with county-level emissions and also because spatial surrogates for tribal data are not currently available. These omissions are not expected to have an impact on the results of the air quality modeling at the 12-km resolution used for this platform.

The following subsections describe how the sources in the NEI nonpoint inventory were separated into modeling platform sectors, along with any data that were updated (replaced) with non-NEI data.

2.2.1 Area fugitive dust sector (afdust)

The area-source fugitive dust (afdust) sector contains PM₁₀ and PM_{2.5} emission estimates for nonpoint SCCs identified by EPA as dust sources. Categories included in the afdust sector are paved roads, unpaved roads and airstrips, construction (residential, industrial, road and total), agriculture production, and mining and quarrying. It does not include fugitive dust from grain elevators, coal handling at coal mines, or vehicular traffic on paved or unpaved roads at industrial facilities because these are treated as point sources so they are properly located. Table 2-5 is a listing of the Source Classification Codes (SCCs) in the afdust sector.

Table 2-5. Afdust sector SCCs

SCC	Tier 1 description	Tier 2 description	Tier 3 description	Tier 4 description
2294000000	Mobile Sources	Paved Roads	All Paved Roads	Total: Fugitives
2296000000	Mobile Sources	Unpaved Roads	All Unpaved Roads	Total: Fugitives
2311010000	Industrial Processes	Construction: SIC 15 - 17	Residential	Total
2311020000	Industrial Processes	Construction: SIC 15 - 17	Industrial/Commercial/ Institutional	Total
2311030000	Industrial Processes	Construction: SIC 15 - 17	Road Construction	Total
2325000000	Industrial Processes	Mining and Quarrying: SIC 14	All Processes	Total
2325020000	Industrial Processes	Mining and Quarrying: SIC 14	Crushed and Broken Stone	Total
2325030000	Industrial Processes	Mining and Quarrying: SIC 14	Sand and Gravel	Total
2325060000	Industrial Processes	Mining and Quarrying: SIC 10	Lead Ore Mining and Milling	Total
2801000000	Miscellaneous Area Sources	Ag. Production - Crops	Agriculture – Crops	Total
2801000003	Miscellaneous Area Sources	Ag. Production - Crops	Agriculture – Crops	Tilling
2801000005	Miscellaneous Area Sources	Ag. Production - Crops	Agriculture – Crops	Harvesting
2801000008	Miscellaneous Area Sources	Ag. Production - Crops	Agriculture - Crops	Transport
2805100010	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Beef cattle - finishing operations on feedlots (drylots)
2805100020	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Dairy Cattle
2805100030	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Broilers
2805100040	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Layers
2805100050	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Swine
2805100060	Miscellaneous Area Sources	Ag. Production – Livestock	Dust kicked up by Livestock	Turkeys

Area Fugitive Dust Transport Fraction

The afdust sector was separated from other nonpoint sectors to allow for the application of a “transport fraction,” and meteorological/precipitation reductions. These adjustments were applied using a script that applies land use-based gridded transport fractions based on landscape roughness, followed by another script that zeroes out emissions for days on which at least 0.01 inches of precipitation occurs or there is snow cover on the ground. The land use data used to reduce the NEI emissions determines the amount of emissions that are subject to transport. This methodology is discussed in Pouliot, et al., 2010, and in “Fugitive Dust Modeling for the 2008 Emissions Modeling Platform” (Adelman, 2012). Both the transport fraction and meteorological adjustments are based on the gridded resolution of the platform (i.e., 12km grid cells); therefore, different emissions will result if the process were applied to different grid resolutions. A limitation of the transport fraction approach is the lack of monthly variability that would be expected with seasonal changes in vegetative cover. While wind speed and direction were not accounted for in the emissions processing, the hourly variability due to soil moisture, snow cover and precipitation were accounted for in the subsequent meteorological adjustment.

Paved and unpaved road dust emissions were from the 2020 NEI adjusted to 2021 levels based on changes between 2020 and 2021 VMT. For the fugitive dust emissions compiled into the 2020 NEI, meteorological adjustments were applied to paved and unpaved road SCCs but not transport adjustments. This is because the modeling platform applies meteorological adjustments and transport adjustments based on unadjusted NEI values. For the 2021 platform, the meteorological adjustments that were applied in the NEI to paved and unpaved road SCCs were backed out and reapplied in SMOKE at an hourly resolution for each grid cell. The FF10 that is run through SMOKE consists of 100% unadjusted emissions, and after SMOKE all afdust sources have both transport and meteorological adjustments applied according to year 2021 meteorology. The total impacts of the transport fraction and meteorological adjustments are shown in Table 2-6.

Table 2-6. Total impact of 2021 fugitive dust adjustments to unadjusted inventory

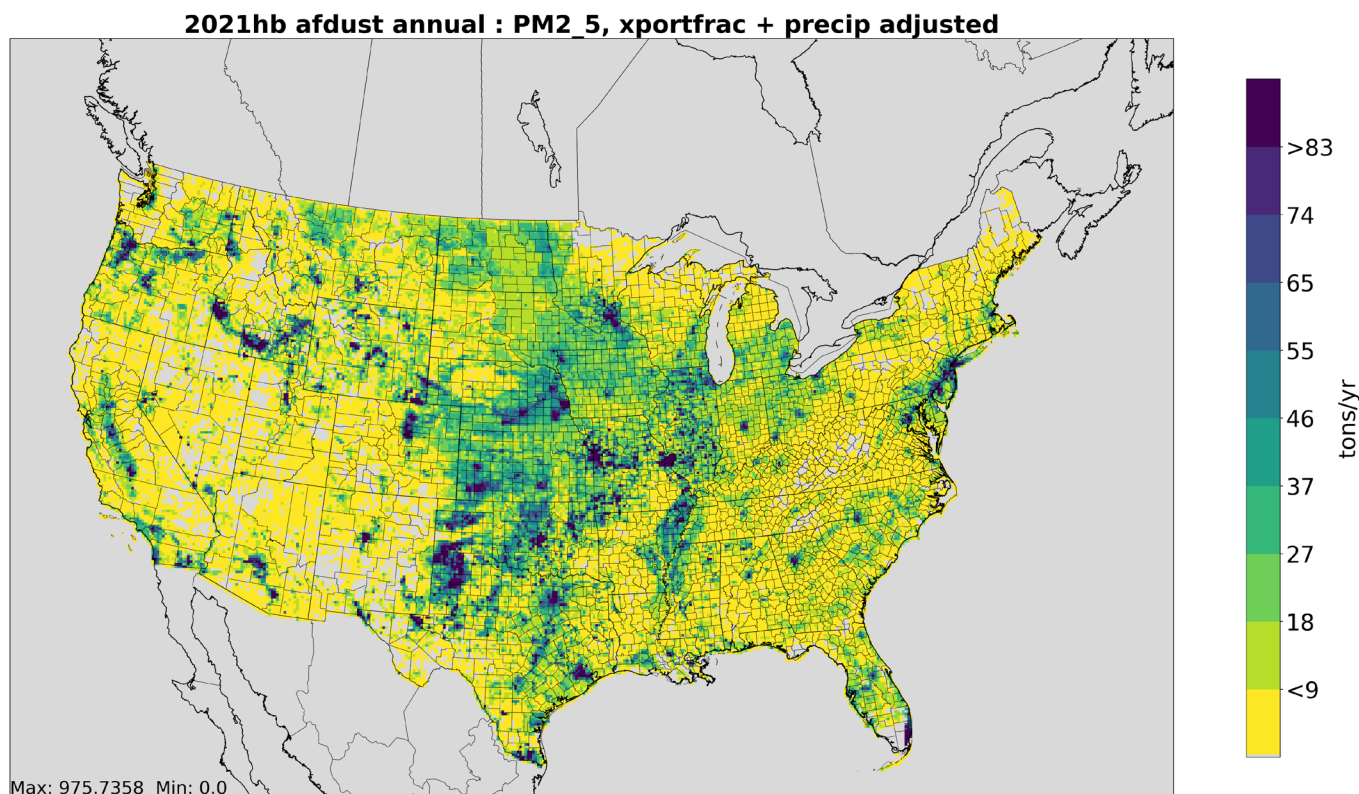
State	Unadjusted PM₁₀	Unadjusted PM_{2.5}	Change in PM₁₀	Change in PM_{2.5}	PM₁₀ Reduction	PM_{2.5} Reduction
Alabama	286,292	36,777	-222,367	-28,460	78%	77%
Arizona	161,676	21,687	-60,138	-7,889	37%	36%
Arkansas	427,784	58,382	-319,511	-42,773	75%	73%
California	349,822	44,855	-149,476	-18,577	43%	41%
Colorado	292,813	41,223	-152,704	-20,416	52%	50%
Connecticut	22,679	3,535	-16,774	-2,595	74%	73%
Delaware	16,664	2,594	-9,530	-1,476	57%	57%
District of Columbia	3,464	470	-2,210	-299	64%	64%
Florida	217,858	34,421	-118,661	-18,283	54%	53%
Georgia	304,157	42,745	-234,361	-32,658	77%	76%
Idaho	531,962	61,840	-299,851	-33,490	56%	54%
Illinois	742,277	94,786	-485,169	-61,670	65%	65%
Indiana	155,315	28,593	-111,213	-20,539	72%	72%

State	Unadjusted PM ₁₀	Unadjusted PM _{2.5}	Change in PM ₁₀	Change in PM _{2.5}	PM ₁₀ Reduction	PM _{2.5} Reduction
Iowa	388,755	56,597	-232,197	-33,773	60%	60%
Kansas	635,604	85,521	-294,147	-39,002	46%	46%
Kentucky	181,268	29,313	-144,359	-23,309	80%	80%
Louisiana	206,623	30,708	-149,591	-22,157	72%	72%
Maine	44,767	6,183	-35,375	-4,897	79%	79%
Maryland	62,084	8,943	-39,187	-5,659	63%	63%
Massachusetts	65,930	8,755	-48,611	-6,315	74%	72%
Michigan	316,401	41,154	-223,319	-28,937	71%	70%
Minnesota	573,650	76,027	-365,089	-47,744	64%	63%
Mississippi	453,862	54,475	-353,342	-42,033	78%	77%
Missouri	1,566,937	177,571	-1,121,332	-126,505	72%	71%
Montana	539,661	70,598	-324,321	-40,283	60%	57%
Nebraska	546,482	74,164	-246,367	-32,652	45%	44%
Nevada	128,743	16,635	-44,929	-5,813	35%	35%
New Hampshire	16,230	3,319	-13,075	-2,662	81%	80%
New Jersey	140,214	17,371	-93,221	-11,439	66%	66%
New Mexico	192,717	24,285	-80,669	-10,015	42%	41%
New York	268,443	37,903	-210,393	-29,366	78%	77%
North Carolina	272,110	36,412	-201,347	-26,809	74%	74%
North Dakota	370,295	56,700	-188,627	-28,550	51%	50%
Ohio	285,183	44,047	-211,889	-32,724	74%	74%
Oklahoma	587,438	80,122	-315,498	-42,187	54%	53%
Oregon	842,361	93,017	-628,197	-67,489	75%	73%
Pennsylvania	149,717	26,243	-110,784	-19,704	74%	75%
Rhode Island	6,070	1,017	-4,144	-684	68%	67%
South Carolina	198,821	25,976	-147,357	-19,110	74%	74%
South Dakota	215,000	37,848	-109,221	-18,927	51%	50%
Tennessee	141,974	26,071	-106,932	-19,649	75%	75%
Texas	1,629,045	224,169	-827,471	-112,813	51%	50%
Utah	150,567	18,876	-87,685	-10,722	58%	57%
Vermont	63,537	7,034	-55,955	-6,180	88%	88%
Virginia	141,234	22,127	-109,282	-17,226	77%	78%
Washington	182,602	22,573	-99,565	-12,396	55%	55%
West Virginia	71,178	9,957	-64,115	-8,945	90%	90%
Wisconsin	207,935	34,793	-138,880	-23,227	67%	67%
Wyoming	674,782	71,920	-377,702	-39,839	56%	55%
Domain Total (12km CONUS)	16,030,982	2,130,333	-9,986,143	-1,308,867	62%	61%

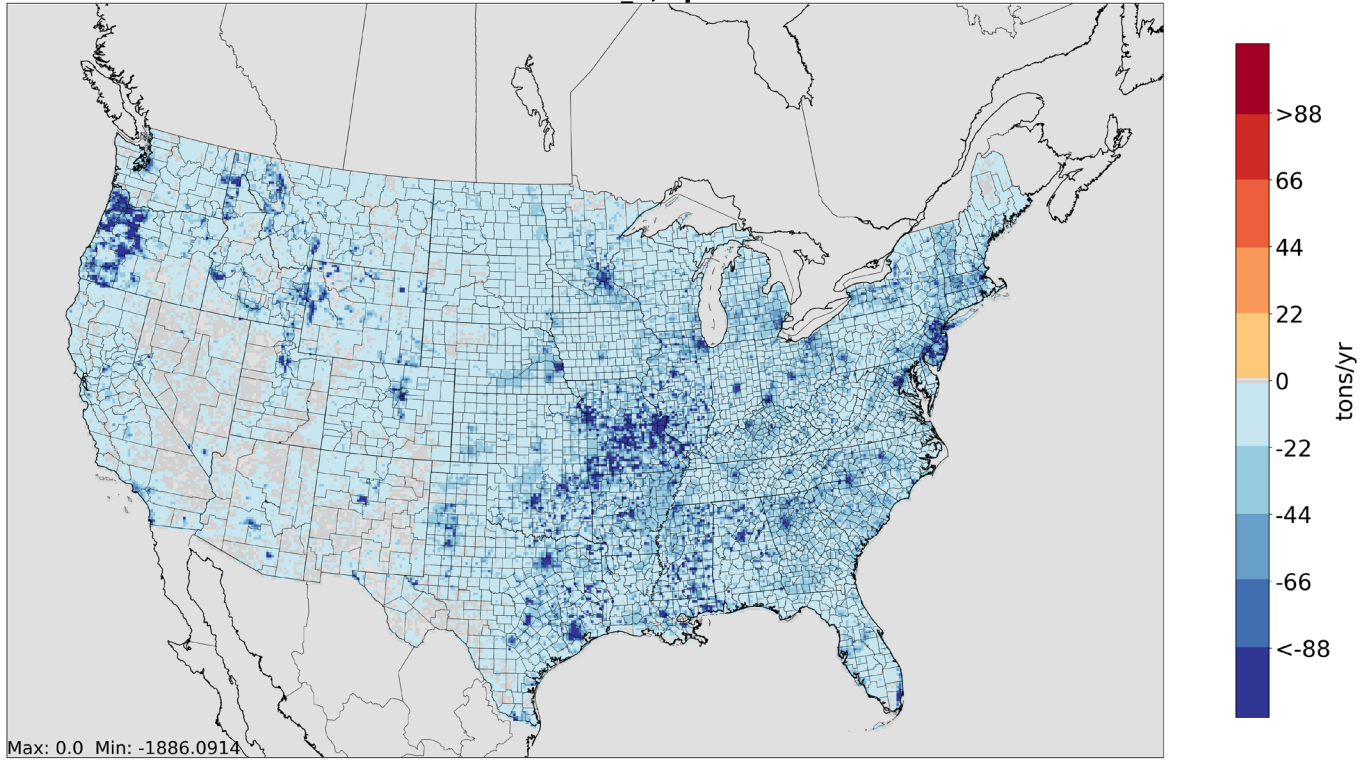
For categories other than paved and unpaved roads, where states submitted afdust data it was assumed that the state-submitted data were not met-adjusted and therefore the meteorological adjustments were applied. Thus, if states submitted data that were met-adjusted for sources other than paved and unpaved roads, these sources would have been adjusted for meteorology twice. Even with that possibility, air quality modeling shows that, in general, dust is frequently overestimated in the air quality modeling results.

Figure 2-1 illustrates the impact of each step of the adjustment. The reductions due to the transport fraction adjustments alone are shown at the top of the figure. The reductions due to the precipitation adjustments alone are shown in the middle of the figure. The cumulative emission reductions after both transport fraction and meteorological adjustments are shown at the bottom of the figure. The top plot shows how the transport fraction has a larger reduction effect in the east, where forested areas are more effective at reducing PM transport than in many western areas. The middle plot shows how the meteorological impacts of precipitation, along with snow cover in the north, further reduce the dust emissions.

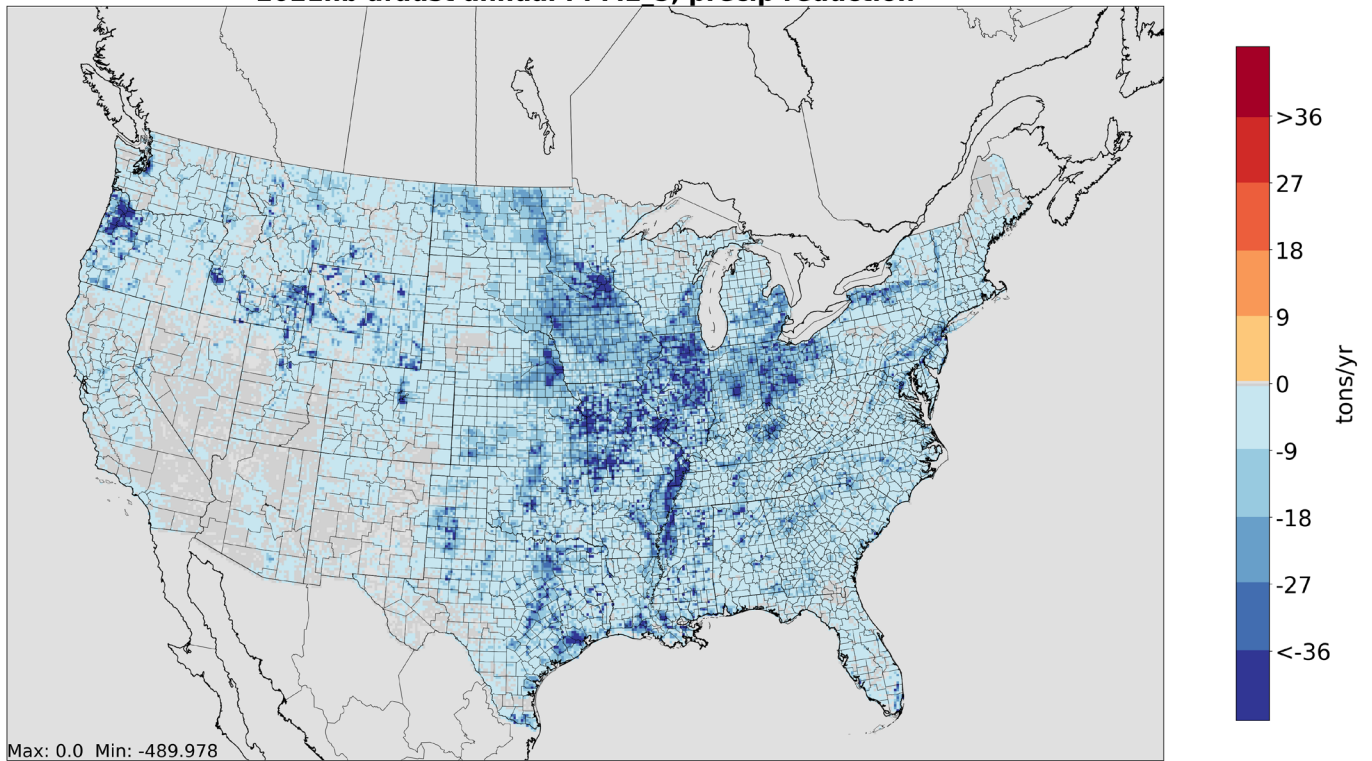
Figure 2-1. Fugitive dust emissions and impact of adjustments due to transport fraction, precipitation, and cumulative



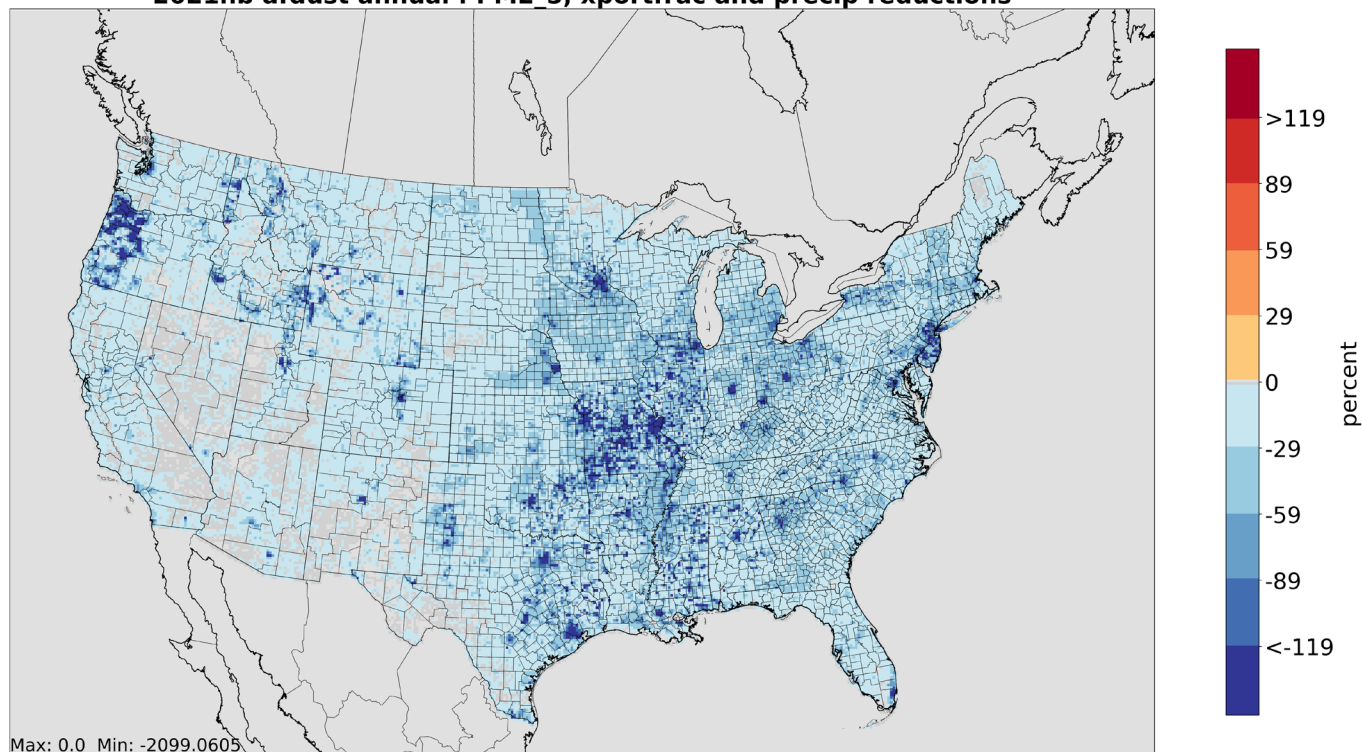
2021hb afdust annual : PM2_5, xportfrac reduction



2021hb afdust annual : PM2_5, precip reduction



2021hb afdust annual : PM2_5, xportfrac and precip reductions



2.2.2 Agricultural Livestock (livestock)

The livestock SCCs are shown in Table 2-7. The livestock emissions are related to beef and dairy cattle, poultry production and waste, swine production, waste from horses and ponies, and production and waste for sheep, lambs, and goats. The sector does not include quite all of the livestock NH_3 emissions, as there is a very small amount of NH_3 emissions from livestock in the ptnonipm inventory (as point sources). In addition to NH_3 , the sector includes livestock emissions from all pollutants other than $\text{PM}_{2.5}$. $\text{PM}_{2.5}$ from livestock are in the afdust sector.

Agricultural livestock emissions in the 2021 platform were developed using methods similar to those used to develop the 2020 NEI, which is a mix of state-submitted data and EPA estimates. The 2020 NEI approach for livestock utilizes daily emission factors by animal and county from a model developed by Carnegie Mellon University (CMU) (Pinder, 2004, McQuilling, 2015) and 2020 U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) survey. Details on the approach used to develop livestock emissions for the 2020 NEI are provided in Section 10 of the 2020 NEI TSD. VOC livestock emissions, new for this sector, were estimated by multiplying a national VOC/ NH_3 emissions ratio by the county NH_3 emissions. Animal populations used for estimating livestock emissions in the 2021 platform were adjusted from 2020 NEI levels to 2021 levels based on 2021 USDA survey data (see QuickStats at <https://quickstats.nass.usda.gov>) for the available counties. The adjustment factors ranged from 0.8 to 1.2. For other counties, the animal populations remained the same as 2020. The FEM model was run for 2021 using meteorological data for 2021.

Table 2-7. SCCs for the livestock sector

SCC	Tier 1 description	Tier 2 description	Tier 3 description	Tier 4 description
2805002000	Miscellaneous Area Sources	Ag. Production – Livestock	Beef cattle production composite	Not Elsewhere Classified
2805007100	Miscellaneous Area Sources	Ag. Production – Livestock	Poultry production - layers with dry manure management systems	Confinement
2805009100	Miscellaneous Area Sources	Ag. Production – Livestock	Poultry production - broilers	Confinement
2805010100	Miscellaneous Area Sources	Ag. Production – Livestock	Poultry production - turkeys	Confinement
2805018000	Miscellaneous Area Sources	Ag. Production – Livestock	Dairy cattle composite	Not Elsewhere Classified
2805025000	Miscellaneous Area Sources	Ag. Production – Livestock	Swine production composite	Not Elsewhere Classified (see also 28-05-039, -047, -053)
2805035000	Miscellaneous Area Sources	Ag. Production – Livestock	Horses and Ponies Waste Emissions	Not Elsewhere Classified
2805040000	Miscellaneous Area Sources	Ag. Production – Livestock	Sheep and Lambs Waste Emissions	Total
2805045000	Miscellaneous Area Sources	Ag. Production – Livestock	Goats Waste Emissions	Not Elsewhere Classified

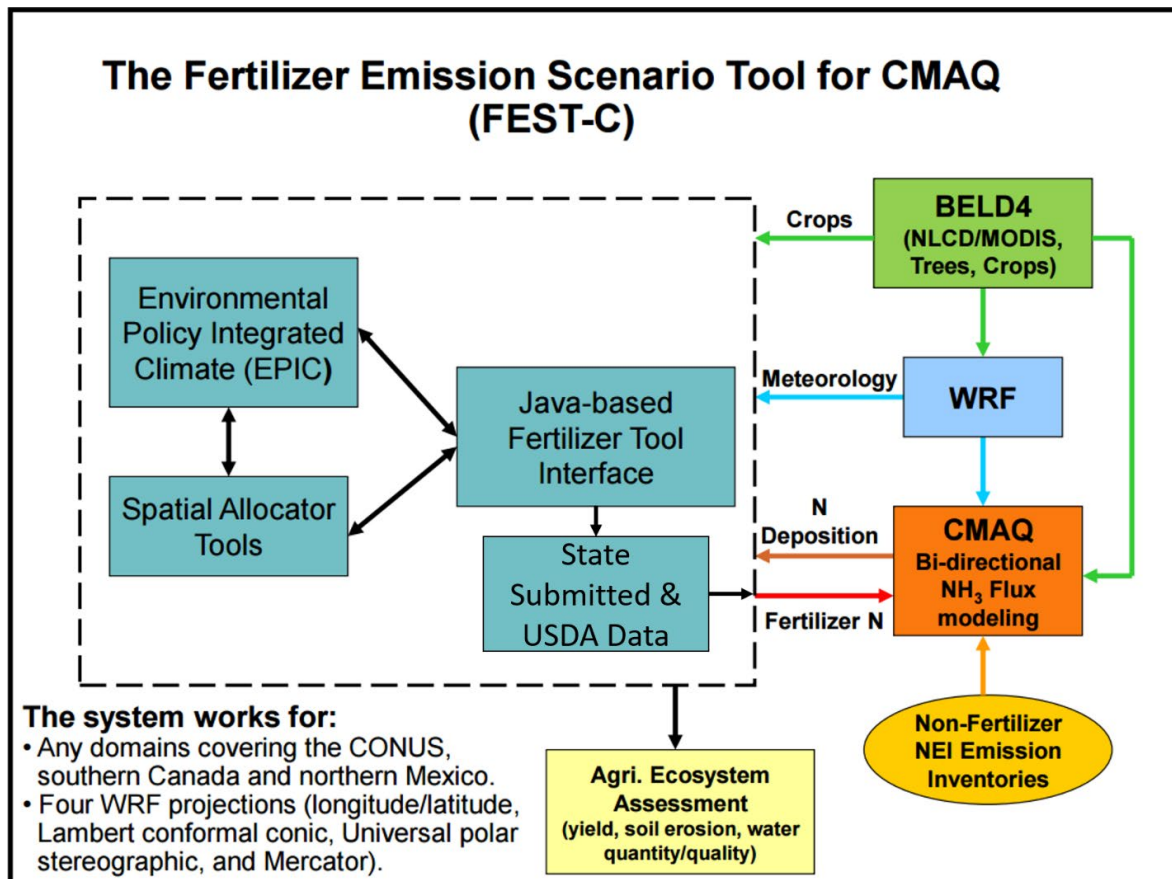
2.2.3 Agricultural Fertilizer (fertilizer)

As described in the 2020 NEI TSD, fertilizer emissions were based on the FEST-C model (<https://www.cmascenter.org/fest-c/>). Unlike most of the other emissions input to the CMAQ model, fertilizer emissions are computed during a run of CMAQ in bi-directional mode and are output during the model run. The bidirectional version of CMAQ (v5.3) and the Fertilizer Emissions Scenario Tool for CMAQ FEST-C (v1.3) were used to estimate ammonia (NH₃) emissions from agricultural soils. The computed emissions were saved during the CMAQ run so they can be included in emissions summaries and in other model runs that do not use the bidirectional method.

FEST-C is the software program that processes land use and agricultural activity data to develop inputs for the CMAQ model when run with bidirectional exchange. FEST-C reads land use data from the Biogenic Emissions Landuse Dataset (BELD), meteorological variables from the Weather Research and Forecasting (WRF) model, and nitrogen deposition data from a previous or historical average CMAQ simulation. FEST-C, then uses the Environmental Policy Integrated Climate (EPIC) modeling system (<https://epicapex.tamu.edu/epic/>) to simulate the agricultural practices and soil biogeochemistry and provides information regarding fertilizer timing, composition, application method and amount.

An iterative calculation was applied to estimate fertilizer emissions. First, fertilizer application by crop type was estimated using FEST-C modeled data. To develop the NEI emissions, CMAQ v5.4 was run with the Surface Tiled Aerosol and Gaseous Exchange (STAGE) deposition option along with bidirectional exchange to estimate fertilizer and biogenic NH₃ emissions. However, for this study, the M3DRY option was used to develop the fertilizer emissions.

Figure 2-2. “Bidi” modeling system used to compute emissions from fertilizer application



Fertilizer Activity Data

The following activity parameters were input into the EPIC model:

- Grid cell meteorological variables from WRF
- Initial soil profiles/soil selection
- Presence of 21 major crops: irrigated and rain fed hay, alfalfa, grass, barley, beans, grain corn, silage corn, cotton, oats, peanuts, potatoes, rice, rye, grain sorghum, silage sorghum, soybeans, spring wheat, winter wheat, canola, and other crops (e.g., lettuce, tomatoes, etc.)
- Fertilizer sales to establish the type/composition of nutrients applied
- Management scenarios for the 10 USDA production regions. These include irrigation, tile drainage, intervals between forage harvest, fertilizer application method (injected versus surface applied), and equipment commonly used in these production regions.

The WRF meteorological model was used to provide grid cell meteorological parameters for the year 2021 using a national 12-km rectangular grid covering the continental U.S. The meteorological parameters in Table 2-8 were used as EPIC model inputs.

Table 2-8. Source of input variables for EPIC

EPIC input variable	Variable Source
Daily Total Radiation (MJ/m ²)	WRF
Daily Maximum 2-m Temperature (C)	WRF
Daily minimum 2-m temperature (C)	WRF
Daily Total Precipitation (mm)	WRF
Daily Average Relative Humidity (unitless)	WRF
Daily Average 10-m Wind Speed (m s ⁻¹)	WRF
Daily Total Wet Deposition Oxidized N (g/ha)	CMAQ
Daily Total Wet Deposition Reduced N (g/ha)	CMAQ
Daily Total Dry Deposition Oxidized N (g/ha)	CMAQ
Daily Total Dry Deposition Reduced N (g/ha)	CMAQ
Daily Total Wet Deposition Organic N (g/ha)	CMAQ

Initial soil nutrient and pH conditions in EPIC were based on the 1992 USDA Soil Conservation Service (CSC) Soils-5 survey. The EPIC model then was run for 25 years using current fertilization and agricultural cropping techniques to estimate soil nutrient content and pH for the 2017 EPIC/WRF/CMAQ simulation.

The presence of crops in each model grid cell was determined using USDA Census of Agriculture data (2012) and USGS National Land Cover data (2011). These two data sources were used to compute the fraction of agricultural land in a model grid cell and the mix of crops grown on that land.

Fertilizer sales data and the 6-month period in which they were sold were extracted from the 2014 Association of American Plant Food Control Officials (AAPFCO, <http://www.aapfco.org/publications.html>). AAPFCO data were used to identify the composition (e.g., urea, nitrate, organic) of the fertilizer used, and the amount applied was estimated using the modeled crop demand. These data were useful in making a reasonable assignment of what kind of fertilizer was applied to which crops.

Management activity data refers to data used to estimate representative crop management schemes. The USDA Agricultural Resource Management Survey (ARMS, [https://www.nass.usda.gov/Surveys/Guide to NASS Surveys/Ag Resource Management/](https://www.nass.usda.gov/Surveys/Guide%20to%20NASS%20Surveys/Ag%20Resource%20Management/)) was used to provide management activity data. These data cover 10 USDA production regions and provide management schemes for irrigated and rain fed hay, alfalfa, grass, barley, beans, grain corn, silage corn, cotton, oats, peanuts, potatoes, rice, rye, grain sorghum, silage sorghum, soybeans, spring wheat, winter wheat, canola, and other crops (e.g., lettuce, tomatoes, etc.).

2.2.4 Nonpoint Oil and Gas Sector (np_oilgas)

The nonpoint oil and gas (np_oilgas) sector includes onshore and offshore oil and gas emissions. The EPA estimated emissions for all counties with 2021 oil and gas activity data using the Oil and Gas Tool. The types of sources covered include drill rigs, workover rigs, artificial lift, hydraulic fracturing engines, pneumatic pumps and other devices, storage tanks, flares, truck loading, compressor engines, and dehydrators. Because of the importance of emissions from this sector, special consideration was given to the speciation, spatial allocation, and monthly temporalization of nonpoint oil and gas emissions, instead of relying on older, more generalized profiles.

The 2020 NEI version of the Nonpoint Oil and Gas Emission Estimation Tool (i.e., the “NEI oil and gas tool”) populated with 2021-specific activity data was used to estimate 2021. Year 2021 oil and gas activity data were obtained from Enverus’ activity database (www.enverus.com) and supplied by some state air agencies. The NEI oil and gas tool is an Access database that utilizes county-level activity data (e.g., oil production and well counts), operational characteristics (types and sizes of equipment), and emission factors to estimate emissions. The tool was used to create a CSV-formatted emissions dataset covering all national nonpoint oil and gas emissions. This dataset was converted to the FF10 format for use in SMOKE modeling. More details on the inputs for and running of the tool for 2020 are provided in the 2020 NEI TSD. Table 2-9 shows the nonpoint oil and gas NO_x and VOC emissions for 2021 by state. The Colorado emissions in this table include emissions submitted to the NEI within the Southern Ute reservation. For spatial allocation purposes, the Southern Ute oil and gas emissions – totaling 11,663 tons/yr of NO_x and 879 tons/yr of VOC - were allocated to Colorado counties, with 95% of the emissions in La Plata County (FIPS 08067) and 5% of the emissions in Archuleta County (FIPS 08007).

Table 2-9. Nonpoint oil and gas emissions for 2021

State	2021 NO _x	2021 VOC
Alabama	3,830	9,596
Alaska	1,583	12,596
Arizona	8	113
Arkansas	4,540	8,686
California	1,438	16,827
Colorado	26,960	68,984
Florida	27	614
Georgia	-	0.04
Idaho	6	61
Illinois	13,558	56,158
Indiana	2,457	14,200
Iowa	-	0.01
Kansas	25,960	70,263
Kentucky	10,792	44,542
Louisiana	19,947	55,400
Maryland	1	2
Michigan	11,609	15,072
Minnesota	-	0.01

State	2021 NO _x	2021 VOC
Mississippi	3,323	8,018
Missouri	432	1,031
Montana	2,258	30,151
Nebraska	256	2,035
Nevada	4	180
New Mexico	120,851	295,051
New York	695	6,505
North Carolina	-	0.003
North Dakota	51,202	233,638
Ohio	2,450	34,243
Oklahoma	43,946	166,564
Oregon	6	31
Pennsylvania	65,308	197,669
South Dakota	175	1,213
Tennessee	874	3,529
Texas	273,305	1,289,662
Utah	13,940	67,149
Virginia	3,967	8,464
Washington	-	3
West Virginia	22,431	162,678
Wyoming	2,105	8,146

A new source was added to the oil and gas sector for the 2020 NEI. Pipeline Blowdowns and Pigging (SCC= 2310021801) emissions were estimated using US EPA Greenhouse Gas Reporting Program (GHGRP) data. These Pipeline Blowdowns and Pigging emissions included county-level estimates of VOC, benzene, toluene, ethylbenzene, and xylene (BTEX). These emissions estimates were calculated outside of the Oil and Gas Tool and submitted to EIS separately from the Oil and Gas Tool emissions. These emissions were considered EPA default emissions and SLTs had the opportunity to submit their own Pipeline Blowdowns and Pigging (e.g., Utah) emissions and/or accept/omit these emissions using the Nonpoint Survey. Unfortunately, these EPA default Pipeline Blowdowns and Pigging emissions did not get into the 2020 NEI release for the states that accepted these emissions due to EIS tagging issues. These emissions were included in this 2021 Emissions Modeling Platform. Table 2-10 shows the emissions totals by state for Pipeline Blowdowns and Pigging sources.

An additional new source was added to the oil and gas sector for this 2021 modeling platform. This new source were Abandoned oil and gas wells in the USA. The term "abandoned wells" encompasses various types of wells:

- Wells with no recent production, and not plugged. Common terms (such as those used in state databases) might include: inactive, temporarily abandoned, shut-in, dormant, and idle.
- Wells with no recent production and no responsible operator. Common terms might include: orphaned, deserted, long-term idle, and abandoned.
- Wells that have been plugged to prevent migration of gas or fluids.

As of year 2020, there are approximately 3.7 million abandoned wells in the U.S., with around 2.3 million abandoned oil wells, 0.6 million abandoned gas wells, and 0.8 million abandoned dry wells (may be oil or gas wells). Abandoned wells may emit CH₄, CO₂, VOC, and various HAP.

Estimates of greenhouse gas (GHG) emissions (CH₄ and CO₂) from abandoned wells have been estimated as part of the Inventory of U.S. Greenhouse Gas Emissions and Sinks since 2018. Currently, a draft version of the inventory from 1990 – 2021 is available for public review. The GHG inventory (GHGI) methodology and estimates of emissions from abandoned wells served as the starting point for development of the VOC and HAP emissions inventory for abandoned wells used in this year 2021 modeling platform. Year 2020 estimates of VOC and BTEX were estimated and used in this year 2021 modeling platform. Table 2-11 shows the emissions totals by state for Pipeline Blowdowns and Pigging sources.

Table 2-10. State emissions totals for year 2021 for Pipeline Blowdowns and Pigging sources

State	VOC (tpy)	Benzene (tpy)	Ethylbenzene (tpy)	Toluene (tpy)	Xylene (tpy)
Alabama	363	1.57	0.088	1.38	0.40
Alaska	11	0.05	0.003	0.05	0.01
Arizona	77	0.35	0.020	0.31	0.09
Arkansas	24	0.01	0.000	0.00	0.00
California	236	1.08	0.061	0.96	0.27
Colorado	3,413	8.89	0.449	10.85	3.38
Florida	1	0.00	0.000	0.00	0.00
Illinois	239	0.88	0.050	0.79	0.22
Indiana	237	1.00	0.056	0.89	0.25
Kansas	4,046	6.86	0.789	6.53	2.79
Kentucky	661	2.99	0.169	2.67	0.76
Louisiana	453	3.26	0.003	0.36	0.57
Maryland	0	0.00	0.000	0.00	0.00
Michigan	241	1.10	0.062	0.98	0.28
Mississippi	1,002	1.55	0.034	0.61	0.50
Missouri	22	0.02	0.000	0.02	0.01
Montana	222	1.01	0.057	0.90	0.26
Nebraska	24	0.06	0.003	0.07	0.02
New Mexico	1,753	0.00	0.000	0.00	0.00
New York	192	0.87	0.049	0.78	0.22
North Dakota	15	0.07	0.004	0.06	0.02
Ohio	476	2.16	0.122	1.93	0.55
Oklahoma	3,019	2.53	0.086	1.46	1.06
Oregon	14	0.06	0.003	0.05	0.02
Pennsylvania	1,366	6.20	0.351	5.53	1.57
South Dakota	4	0.02	0.001	0.02	0.00
Tennessee	19	0.09	0.005	0.08	0.02

State	VOC (tpy)	Benzene (tpy)	Ethylbenzene (tpy)	Toluene (tpy)	Xylene (tpy)
Texas	25,237	29.05	0.642	11.33	9.59
Utah	25	0.13	0.007	0.12	0.05
Virginia	294	1.33	0.076	1.19	0.34
West Virginia	1,240	5.62	0.318	5.01	1.42
Wyoming	236	1.58	0.112	0.63	0.49
US Total	45,161	80.38	3.619	55.53	25.15

Table 2-11. State emissions totals for year 2021 for Abandoned Wells sources

State	2021 VOC (tpy)
Alabama	185
Alaska	60
Arizona	10
Arkansas	909
California	5,239
Colorado	349
Florida	40
Georgia	0
Idaho	0
Illinois	6,726
Indiana	3,214
Iowa	0
Kansas	7,146
Kentucky	11,762
Louisiana	2,375
Maryland	1
Michigan	433
Minnesota	0
Mississippi	725
Missouri	118
Montana	598
Nebraska	137
Nevada	14
New Mexico	391
New York	658
North Carolina	0
North Dakota	418
Ohio	21,812
Oklahoma	7,426
Oregon	3
Pennsylvania	69,837

State	2021 VOC (tpy)
South Dakota	27
Tennessee	1,279
Texas	31,921
Utah	173
Virginia	67
Washington	3
West Virginia	2,098
Wyoming	514
US Total	176,670

Lastly, EPA and the state of New Mexico worked together to exercise the point source subtraction step in the Oil and Gas Tool during the 2020 NEI development period. This point source subtraction step was used for New Mexico because additional oil and gas point sources submitted by New Mexico that were the same processes estimated in the Oil and Gas Tool (non-point sources). This point source subtraction step is a processed used to eliminate possible double counting of sources in the Oil and Gas Tool that are already defined in the point source inventory. Unfortunately, the resulting non-point emissions from the point source subtraction step for New Mexico did not get into the 2020 NEI release due to EIS tagging issues. New Mexico non-point oil and gas emissions are overestimated in the 2020 NEI as a result. This overestimation was corrected for this 2021 Emissions Modeling Platform.

2.2.5 Residential Wood Combustion (rwc)

The residential wood combustion (rwc) sector includes residential wood burning devices such as fireplaces, fireplaces with inserts (inserts), free standing woodstoves, pellet stoves, outdoor hydronic heaters (also known as outdoor wood boilers), indoor furnaces, and outdoor burning in firepots and chimeneas. Free standing woodstoves and inserts are further differentiated into three categories: 1) conventional (not EPA certified); 2) EPA certified, catalytic; and 3) EPA certified, noncatalytic. Generally speaking, the conventional units were constructed prior to 1988. Units constructed after 1988 have to meet EPA emission standards and they are either catalytic or non-catalytic. As with the other nonpoint categories, a mix of S/L and EPA estimates were used. The EPA's estimates use updated methodologies for activity data and some changes to emission factors. The source classification codes (SCCs) in the rwc sector are listed in Table 2-12.

The 2021 platform RWC emissions are unchanged from the data in the 2020 NEI and include some improvements to RWC emissions estimates developed as part of the 2020 NEI process. The EPA, along with the Commission on Environmental Cooperation (CEC), the Northeast States for Coordinated Air Use Management (NESCAUM), and Abt Associates, conducted a national survey of wood-burning activity in 2018. The results of this survey were used to estimate county-level burning activity data. The activity data for RWC processes is the amount of wood burned in each county, which is based on data from the CEC survey on the fraction of homes in each county that use each wood-burning appliance and the average amount of wood burned in each appliance. These assumptions were used with the number of occupied homes in each county to estimate the total amount of wood burned in each county, in cords for cordwood appliances and tons for pellet appliances. Cords of wood were converted to tons using

county-level density factors from the U.S. Forest Service. RWC emissions were calculated by multiplying the tons of wood burned by emissions factors. For more information on the development of the residential wood combustion emissions, see Section 27 of the 2020 NEI TSD.

Table 2-12. SCCs for the residential wood combustion sector

SCC	Tier 1 Description	Tier 2 Description	Tier 3 Description	Tier 4 Description
2104008100	Stationary Source Fuel Combustion	Residential	Wood	Fireplace: general
2104008210	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: fireplace inserts; non-EPA certified
2104008220	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: fireplace inserts; EPA certified; non-catalytic
2104008230	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: fireplace inserts; EPA certified; catalytic
2104008300	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: freestanding, general
2104008310	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: freestanding, non-EPA certified
2104008320	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: freestanding, EPA certified, non-catalytic
2104008330	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: freestanding, EPA certified, catalytic
2104008400	Stationary Source Fuel Combustion	Residential	Wood	Woodstove: pellet-fired, general (freestanding or FP insert)
2104008510	Stationary Source Fuel Combustion	Residential	Wood	Furnace: Indoor, cordwood-fired, non-EPA certified
2104008530	Stationary Source Fuel Combustion	Residential	Wood	Furnace: Indoor, pellet-fired, general
2104008610	Stationary Source Fuel Combustion	Residential	Wood	Hydronic heater: outdoor
2104008620	Stationary Source Fuel Combustion	Residential	Wood	Hydronic heater: indoor
2104008630	Stationary Source Fuel Combustion	Residential	Wood	Hydronic heater: pellet-fired
2104008700	Stationary Source Fuel Combustion	Residential	Wood	Outdoor wood burning device, NEC (fire-pits, chimeneas, etc)
2104009000	Stationary Source Fuel Combustion	Residential	Firelog	Total: All Combustor Types

2.2.6 Solvents (np_solvents)

The np_solvents sector is a diverse collection of emission sources for which emissions are driven by evaporation. Included in this sector are everyday items, such as cleaners, personal care products, adhesives, architectural and aerosol coatings, printing inks, and pesticides. These sources exclusively emit organic gases and feature origins spanning residential, commercial, institutional, and industrial

settings. The organic gases that evaporate from these sources often fulfill other functions than acting as a traditional solvent (e.g., propellants, fragrances, emollients). For this reason, the solvents sector is often referred to as “volatile chemical products.” The base methodology used to estimate these emissions are unchanged from the 2020 NEI, which is described in [Section 32 of the 2020 NEI TSD](#).

For 2021, in many instances, emissions within the 2020 NEI were carried forward. This includes pollutant/county/SCC emission submissions (including Hazardous Air Pollutants) from State, Locality, and Tribal partners, and all asphalt related emissions. The np_solvents sector also includes emissions from SCCs included in the 2020 NEI but not covered by VCPy, the model used to estimate most nonpoint emissions in the solvent sector (Seltzer, et al., 2021). These emissions come from State, Locality, and Tribal emission submissions for select SCCs, all of which are listed in Table 2-13. All other np_solvents emissions were estimated using the same methods used to generate the 2020 NEI and are specific to 2021-year usage.

Table 2-13. Non-VCPy SCCs in the np_solvents sector

SCC	Description
2401050000	Solvent Utilization;Surface Coating;Miscellaneous Finished Metals: SIC 34 - (341 + 3498);Total: All Solvent Types
2440020000	Solvent Utilization;Miscellaneous Industrial;Adhesive (Industrial) Application;Total: All Solvent Types
2461021000	Solvent Utilization;Miscellaneous Non-industrial: Commercial;Cutback Asphalt;Total: All Solvent Types
2461022000	Solvent Utilization;Miscellaneous Non-industrial: Commercial;Emulsified Asphalt;Total: All Solvent Types
2461023000	Solvent Utilization;Miscellaneous Non-industrial: Commercial;Asphalt Roofing;Total: All Solvent Types
2461025100	Solvent Utilization;Miscellaneous Non-industrial: Commercial; Asphalt Paving: Hot and Warm Mix;Hot Mix Total: All Solvent Types
2461025200	Solvent Utilization;Miscellaneous Non-industrial: Commercial; Asphalt Paving: Hot and Warm Mix;Warm Mix Total: All Solvent Types
2461800001	Solvent Utilization;Miscellaneous Non-industrial: Commercial;Pesticide Application: All Processes;Surface Application

2.2.7 Nonpoint (nonpt)

The 2021 platform nonpt sector inventory is unchanged from the April 2023 version of the 2020 NEI. Stationary nonpoint sources that were not subdivided into the afdust, livestock, fertilizer, np_oilgas, rwc or np_solvents sectors were assigned to the “nonpt” sector. Locomotives and CMV mobile sources from the 2020 NEI nonpoint inventory are described with the mobile sources. The types of sources in the nonpt sector include:

- stationary source fuel combustion, including industrial, commercial, and residential and orchard heaters;
- chemical manufacturing;

- industrial processes such as commercial cooking, metal production, mineral processes, petroleum refining, wood products, fabricated metals, and refrigeration;
- storage and transport of petroleum for uses such as portable gas cans, bulk terminals, gasoline service stations, aviation, and marine vessels;
- storage and transport of chemicals;
- waste disposal, treatment, and recovery via incineration, open burning, landfills, and composting; and
- miscellaneous area sources such as cremation, hospitals, lamp breakage, and automotive repair shops.

The nonpt sector includes emission estimates for Portable Fuel Containers (PFCs), also known as “gas cans.” The PFC inventory consists of three distinct sources of PFC emissions, further distinguished by residential or commercial use. The three sources are: (1) displacement of the vapor within the can; (2) emissions due to evaporation (i.e., diurnal emissions); and (3) emissions due to permeation. Note that spillage and vapor displacement associated with using PFCs to refuel nonroad equipment are included in the nonroad inventory.

2.3 Onroad Mobile sources (onroad)

Onroad mobile source include emissions from motorized vehicles operating on public roadways. These include passenger cars, motorcycles, minivans, sport-utility vehicles, light-duty trucks, heavy-duty trucks, and buses. The sources are further divided by the fuel they use, including diesel, gasoline, E-85, compressed natural gas (CNG), or electric vehicles. The sector characterizes emissions from parked vehicle processes (e.g., starts, hot soak, and extended idle) as well as from on-network processes (i.e., from vehicles as they move along the roads). For more details on the approach and for a summary of the MOVES inputs submitted by states, see section 5 of the 2020 NEI TSD.

For the 2021 emissions modeling platform activity data (i.e., VMT, VPOP) were based on state submitted CDBs for 2020, as well as data from Federal Highway Administration (FHWA) annual VMT at the county level. VMT were projected from 2020 to 2021 using state VM-2 data from FHWA. VPOP was held constant at 2020 levels. A new MOVES run for 2021 was done using MOVES4 to obtain year-specific emission factors.

Except for California, all onroad emissions were generated using the SMOKE-MOVES emissions modeling framework that leverages MOVES-generated emission factors (<https://www.epa.gov/moves>), county and SCC-specific activity data, and hourly 2021 meteorological data. Specifically, EPA used projected vehicle miles traveled (VMT), and a factor of 2021 VMT/2020 VMT was applied to the 2020 hoteling hours; along with tools that integrated the MOVES model with SMOKE. In this way, it was possible to take advantage of the gridded hourly temperature data available from meteorological modeling that are also used for air quality modeling. The onroad source classification codes (SCCs) in the emissions modeling platform are more finely resolved than those in the National Emissions Inventory (NEI). The NEI SCCs distinguish vehicles and fuels. The SCCs used in the model platform also distinguish between emissions processes (i.e., off-network, on-network, and extended idle), and road types. EPA elected to keep 2020 NEI fuel splits (derived from MOVES3) and not upgrade to MOVES4 fuels.

MOVES4 includes the following updates from MOVES3:

- Incorporates updates to vehicle populations, fuel supply, travel activity, and emission rates.
- Accounts for the emission impacts of the EPA heavy-duty low NO_x rule for model years 2027 and later and the light-duty greenhouse gas rule for model years 2023 and later.
- Adds the ability to model heavy-duty battery-electric and fuel-cell vehicles, as well as CNG long-haul combination trucks.
- Improves the modeling of light-duty electric vehicles.
- Improves the user interface to make the model easier to use and updating the platform for compatibility with newer software.

2.3.1 Inventory Development using SMOKE-MOVES

Except for California, onroad emissions were computed with SMOKE-MOVES by multiplying specific types of vehicle activity data by the appropriate emission factors. This section includes discussions of the activity data and the emission factor development. The vehicles (aka source types) for which MOVES computes emissions are shown in Table 2-14. SMOKE-MOVES was run for specific modeling grids. Emissions for the contiguous U.S. states and Washington, D.C., were computed for a grid covering those areas. Emissions for Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands were computed by running SMOKE-MOVES for distinct grids covering each of those regions and are included in the onroad_nonconus sector. In some summary reports these non-CONUS emissions are aggregated with emissions from the onroad sector.

Table 2-14. MOVES vehicle (source) types

MOVES vehicle type	Description	HPMS vehicle type
11	Motorcycle	10
21	Passenger Car	25
31	Passenger Truck	25
32	Light Commercial Truck	25
41	Other Bus	40
42	Transit Bus	40
43	School Bus	40
51	Refuse Truck	50
52	Single Unit Short-haul Truck	50
53	Single Unit Long-haul Truck	50
54	Motor Home	50
61	Combination Short-haul Truck	60
62	Combination Long-haul Truck	60

SMOKE-MOVES makes use of emission rate “lookup” tables generated by MOVES that differentiate emissions by process (i.e., running, start, vapor venting, etc.), vehicle type, road type, temperature, speed, hour of day, etc., to generate the MOVES emission rates that could be applied across the U.S., EPA used an automated process to run MOVES to produce year 2021-specific emission factors by temperature and speed for a series of “representative counties,” to which every other county was mapped. The representative counties for which emission factors were generated were selected

according to their state, elevation, fuels, age distribution, ramp fraction, and inspection and maintenance programs. Each county was then mapped to a representative county based on its similarity to the representative county with respect to those attributes. For this study, there are 254 representative counties in the continental U.S. and a total of 293 including the non-CONUS areas. The only difference between 2020 and 2021 being a change in Alaska county equivalents which removed one borough (county ID 2261, Valdez-Cordova Census Area) which in 2019 split into two areas (county ID 2063, Chugach Census Area; and county ID 2066, Copper River Census Area).

Once representative counties were identified, emission factors were generated with MOVES for each representative county and for two “fuel months” – January to represent winter months, and July to represent summer months – due to the different types of fuels used. SMOKE selected the appropriate MOVES emissions rates for each county, hourly temperature, SCC, and speed bin and then multiplied the emission rate by appropriate activity data. For on-roadway emissions, vehicle miles travelled (VMT) is the activity data; off-network processes use vehicle population (VPOP), vehicle starts, and hours of off-network idling (ONI); and hoteling hours are used to develop emissions for extended idling of combination long-haul trucks. These calculations were done for every county and grid cell in the continental U.S. for each hour of the year.

The SMOKE-MOVES process for creating the model-ready emissions consists of the following steps:

- 1) Determine which counties will be used to represent other counties in the MOVES runs.
- 2) Determine which months will be used to represent other month’s fuel characteristics.
- 3) Create inputs needed only by MOVES. MOVES requires county-specific information on vehicle populations, age distributions, and inspection-maintenance programs for each of the representative counties.
- 4) Create inputs needed both by MOVES and by SMOKE, including temperatures and activity data.
- 5) Run MOVES to create emission factor tables for the temperatures found in each county.
- 6) Run SMOKE to apply the emission factors to activity data (VMT, VPOP, STARTS, off-network idling, and HOTELING) to calculate emissions based on the gridded hourly temperatures in the meteorological data.
- 7) Aggregate the results to the county-SCC level for summaries and quality assurance.

The onroad emissions were processed in six processing streams that were then merged together into the onroad sector emissions after each of the six streams have been processed:

- rate-per-distance (RPD) uses VMT as the activity data plus speed and speed profile information to compute on-network emissions from exhaust, evaporative, permeation, refueling, and brake and tire wear processes;
- rate-per-vehicle (RPV) uses VPOP activity data to compute off-network emissions from exhaust, evaporative, permeation, and refueling processes;
- rate-per-profile (RPS) uses STARTS activity data to compute off-network emissions from vehicles starts;

- rate-per-profile (RPP) uses VPOP activity data to compute off-network emissions from evaporative fuel vapor venting, including hot soak (immediately after a trip) and diurnal (vehicle parked for a long period) emissions;
- rate-per-hour (RPH) uses hoteling hours activity data to compute off-network emissions for idling of long-haul trucks from extended idling and auxiliary power unit process; and
- rate-per-hour off-network idling (RPHO) uses off network idling hours activity data to compute off-network idling emissions for all types of vehicles.

The onroad emissions inputs to MOVES for the 2021 emissions modeling platform are based on the 2020 NEI, described in more detail in Section 5 of the 2020 NEI TSD. These inputs include:

- Key parameters in the MOVES County databases (CDBs) including Low Emission Vehicle (LEV) table
- Fuel months
- Activity data (e.g., VMT, VPOP, speed, HOTELING)

Fuel months, age distributions, and other inputs were consistent with those used to compute the 2020 NEI. Activity data submitted by states and development of the EPA default activity data sets for VMT, VPOP, and hoteling hours are described in detail in the 2020 NEI TSD and supporting documents. Hoteling hours activity were used to calculate emissions from extended idling and auxiliary power units (APUs) by combination long-haul trucks. Hoteling hours for 2021 were calculated by applying a 2021 restricted road VMT/2020 restricted road VMT factor was applied to the 2020 hoteling hours.

2.3.2 Onroad Activity Data Development

SMOKE-MOVES uses vehicle miles traveled (VMT), vehicle population (VPOP), vehicle starts, hours of off-network idling (ONI), and hours of hoteling, to calculate emissions. These datasets are collectively known as “activity data”. For each of these activity datasets, first a national dataset was developed; this national dataset is called the “EPA default” dataset. The default dataset started with the 2020 NEI activity data, which was supplemented with data submitted by state and local agencies. EPA grew the 2020 VMT to 2021 using factors derived from FHWA’s VM-2 data. VPOP was held constant at 2020 levels, as were the starts and fuel splits. EPA default activity was used for California, but the emissions were scaled to California-supplied values during the emissions processing.

Vehicle Miles Traveled (VMT) and Vehicle Population (VPOP)

Activity data submitted by states and development of the EPA default activity data sets for VMT, VPOP, and hoteling hours are described in detail in the 2020 NEI TSD (EPA, 2023) and supporting documents. A growth factor, derived from FHWA data, was applied to the VMT to make it representative of 2021. VPOP was held constant with the 2020 NEI VPOP.

Speed Activity (SPDIST)

In SMOKE 4.7, SMOKE-MOVES was updated to use speed distributions similarly to how they are used when running MOVES in inventory mode. The speed distribution file, called SPDIST, specifies the amount

of time spent in each MOVES speed bin for each county, vehicle (aka source) type, road type, weekday/weekend, and hour of day. This file contains the same information at the same resolution as the Speed Distribution table used by MOVES but is reformatted for SMOKE.

Due to pandemic effects on most of the 2020 data, for the 2021 emissions modeling platform, the best estimate for year 2021 speeds available were the speeds from December 2020. Speed data from the Streetlight dataset were used to generate hourly speed profiles by county, SCC, and month. The SPDIST files for the 2021 emissions modeling platform are based on a combination of the Streetlight project data and 2020 NEI MOVES CDBs. More information can be found in the 2020 NEI TSD (EPA, 2023) and supporting documents.

Hoteling Hours (HOTELING)

Hoteling hours were based on the hoteling hours for the 2020 NEI. County-specific factors of 2021 restricted road VMT/2020 restricted road VMT for combination long haul trucks were applied to the 2020 hoteling hours to make them more representative of 2021. Hoteling hours were capped by county at a theoretical maximum and any excess hours of the maximum were reduced. For calculating reductions, a dataset of truck stop parking space availability was used, which includes a total number of parking spaces per county. This same dataset is used to develop the spatial surrogate for allocating county-total hoteling emissions to model grid cells. The parking space dataset includes several recent updates based on new truck stops opening and other new information. There are 8,784 hours in the year 2020; therefore, the maximum number of possible hoteling hours in a particular county is equal to $8,784 \times$ the number of parking spaces in that county. Hoteling hours were capped at that theoretical maximum value for 2020 in all counties, with some exceptions. This cap was only evaluated for the data reflecting 2020 and not re-evaluated for 2021.

Because the truck stop parking space dataset may be incomplete in some areas, and trucks may sometimes idle in areas other than designated spaces, it was assumed that every county has at least 12 parking spaces, even if fewer parking spaces are found in the parking space dataset. Therefore, hoteling hours were never reduced below 105,408 hours for the year in any county. If the unreduced hoteling hours were already below that maximum, the hours were left unchanged; in other words, hoteling activity were never increased in this analysis. For the 2020 NEI, four states requested that no reductions be applied to the hoteling activity based on parking space availability: CO, ME, NJ, and NY. For these states, reductions based on parking space availability were not applied.

The final step related to hoteling activity is to split county totals into separate values for extended idling (SCC 2202620153) and Auxiliary Power Units (APUs) (SCC 2202620191). For 2021 modeling with MOVES4, an 8.2% APU split is used nationwide, meaning that during 8.2% of the hoteling hours auxiliary power units are assumed to be running.

Starts

Onroad “start” emissions are the instantaneous exhaust emissions that occur at the engine start (e.g., due to the fuel rich conditions in the cylinder to initiate combustion) as well as the additional running exhaust emissions that occur because the engine and emission control systems have not yet stabilized at

the running operating temperature. Operationally, start emissions are defined as the difference in emissions between an exhaust emissions test with an ambient temperature start and the same test with the engine and emission control systems already at operating temperature. As such, the units for start emission rates are instantaneous grams/start.

MOVES4 uses vehicle population information to sort the vehicle population into source bins defined by vehicle source type, fuel type (gas, diesel, etc.), regulatory class, model year and age. The model uses default data from instrumented vehicles (or user-provided values) to estimate the number of starts for each source bin and to allocate them among eight operating mode bins defined by the amount of time parked (“soak time”) prior to the start. Thus, MOVES4 accounts for different amounts of cooling of the engine and emission control systems. Each source bin and operating mode has an associated g/start emission rate. Start emissions are also adjusted to account for fuel characteristics, LD inspection and maintenance (I/M) programs, and ambient temperatures. Starts were held constant from 2020 to 2021, however, new monthly profiles were applied for 2021.

Off-network Idling Hours

After creating VMT inputs for SMOKE-MOVES, Off-network idle (ONI) activity data were also needed. ONI is defined in MOVES as time during which a vehicle engine is running idle and the vehicle is somewhere other than on the road, such as in a parking lot, a driveway, or at the side of the road. This engine activity contributes to total mobile source emissions but does not take place on the road network. Examples of ONI activity include:

- light duty passenger vehicles idling while waiting to pick up children at school or to pick up passengers at the airport or train station,
- single unit and combination trucks idling while loading or unloading cargo or making deliveries, and
- vehicles idling at drive-through restaurants.

Note that ONI does not include idling that occurs on the road, such as idling at traffic signals, stop signs, and in traffic—these emissions are included as part of the running and crankcase running exhaust processes on the other road types. ONI also does not include long-duration idling by long-haul combination trucks (hoteling/extended idle), as that type of long duration idling is accounted for in other MOVES processes.

ONI activity hours were calculated based on VMT. For each representative county, the ratio of ONI hours to onroad VMT (on all road types) was calculated using the MOVES ONI Tool by source type, fuel type, and month. These ratios are then multiplied by each county’s total VMT (aggregated by source type, fuel type, and month) to get hours of ONI activity.

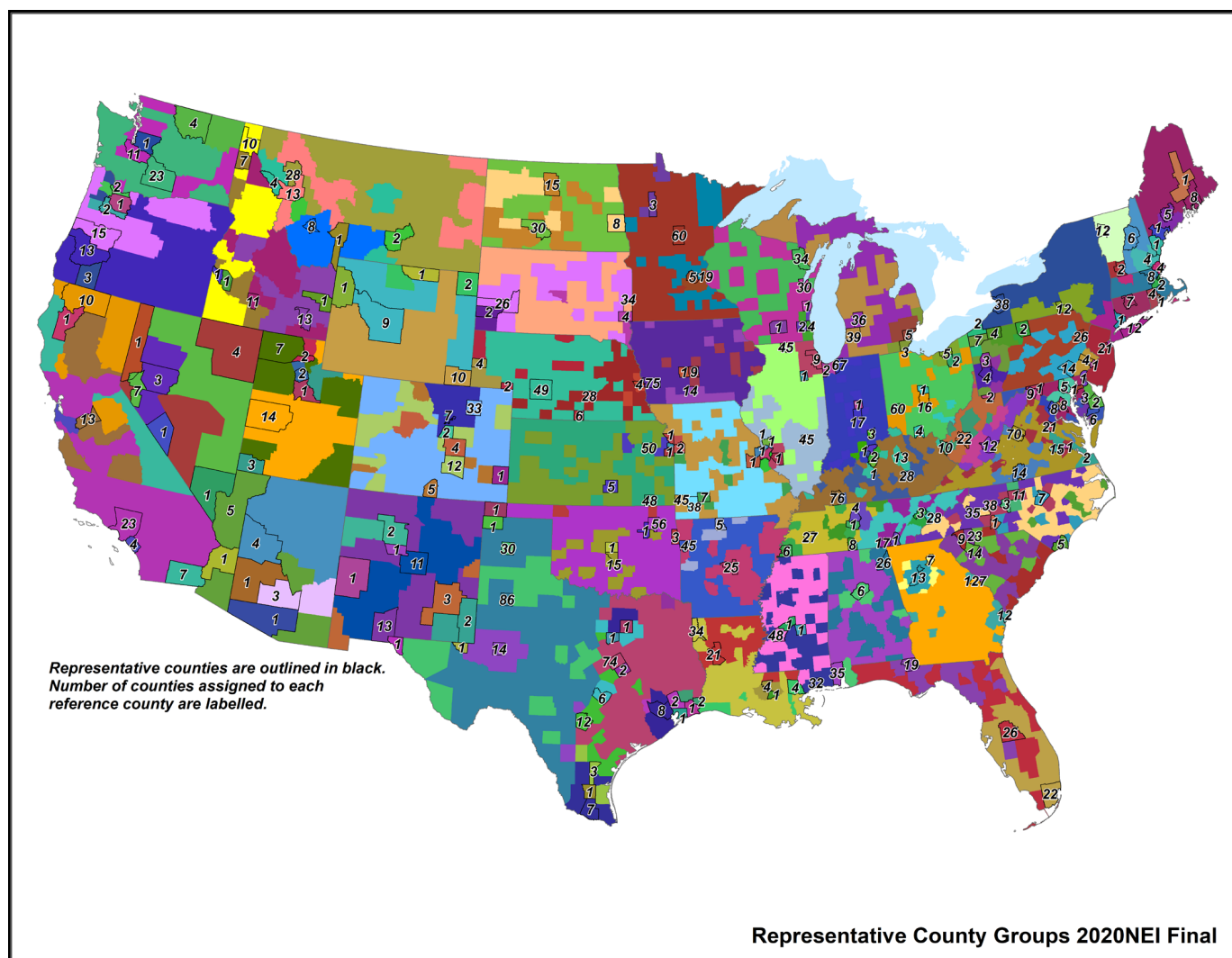
2.3.3 MOVES Emission Factor Table Development

MOVES4 was run in emission rate mode to create emission factor tables for 2021, for all representative counties and fuel months. The county databases used to run MOVES to develop the emission factor tables included the state-specific control measures such as the California LEV program, and fuels

representing the year 2021. The range of temperatures run along with the average humidities used were specific to the year 2021. The remaining settings for the CDBs are documented in the 2020 NEI TSD. To create the emission factors, MOVES was run separately for each representative county and fuel month for each temperature bin needed for the calendar year 2021. The MOVES results were post-processed into CSV-formatted emission factor tables that can be read by SMOKE-MOVES. Additionally, MOVES was run for all counties in Alaska, Hawaii, and Virgin Islands, and for a single representative county in Puerto Rico.

The county databases (CDBs) used to run MOVES to develop the emission factor tables were based on those used for the 2020 NEI. The 2021 emissions modeling platform development included an extensive review of the various tables including speed distributions. Each county in the continental U.S. was classified according to its state, altitude (high or low), fuel region, the presence of I/M programs, and the mean light-duty age. A binning algorithm was executed to identify “like counties. The result was 254 representative counties for the CONUS shown in Figure 2-3 along with 39 for Alaska, Hawaii, Puerto Rico, and the US Virgin Islands. The CONUS representation counties for 2021 are the same as those used for 2020 NEI.

Figure 2-3. Map of Representative Counties



Age distributions are a key input to MOVES in determining emission rates. Age distributions were held constant from 2020 for the 2021 emissions modeling platform. The age distributions for 2020 were updated based on vehicle registration data obtained from IHS Markit, subject to reductions for older vehicles. One of the findings of CRC project A-115 is that IHS data contain higher vehicle populations than state agency analyses of the same Department of Motor Vehicles data, and the discrepancies tend to increase with increasing vehicle age (i.e., there are more older vehicles in the IHS data) and appropriate decreases in older vehicles were applied when the age distributions were computed for 2020 as follows.

Although 33 S/L/T agencies participated in the data submittal process for 2020 NEI onroad mobile sources, only 15 provided both LDV populations (MOVES `SourceTypeYear` table) and age distributions (MOVES `SourceTypeAgeDistribution` table) based on 2020 registration data, which was a requirement for comparison with the 2020 IHS data. Other agencies were excluded from the adjustment factor analysis because they provided only one type of local data (e.g., population but no age distribution) or data with outdated (e.g., year 2013) or unknown registration data draw dates. For the 15 areas that

could be included in the analysis, EPA first combined the populations of passenger cars (source type 21) and light-duty trucks (source types 31 and 32) at the county level to remove the uncertainty of VIN decoding personal passenger vehicles as cars vs. light-duty trucks. EPA then allocated each county's LDV total source type population to vehicle model years for comparison with IHS and found that the IHS populations for 2020 were higher than the state data by 10.8 percent. Similar to prior years' comparisons, EPA again found that the discrepancies in the 2020 data between IHS and states are larger for older vehicles. Table 2-15 shows the adjustments EPA made to the 2020 IHS data prior to its use in the NEI and in the 2021 emissions modeling platform.

EPA calculated the adjustment factors representing the fraction of population remaining in every model year, with two exceptions. Model years from 2011 to 2020 received no adjustment and the model year 1990 received a capped adjustment that equals the adjustment for model year 1991. The adjustment factors in Table 2-15 were applied to the 2020 IHS data to create the EPA Default set of population and age distributions for the NEI.

Table 2-15. The fraction of IHS vehicle populations retained for 2020 NEI and 2021 emissions modeling platform by model year

Model Year	LDV Adjustment Factor
pre-1991	0.722
1991	0.722
1992	0.728
1993	0.742
1994	0.754
1995	0.766
1996	0.774
1997	0.790
1998	0.787
1999	0.798
2000	0.796
2001	0.806
2002	0.808
2003	0.828
2004	0.844
2005	0.857
2006	0.874
2007	0.892
2008	0.905
2009	0.919
2010	0.929
2011 - 2020	1

EPA also removed the county-specific fractions of antique license plate vehicles present in the registration data from IHS, based on the assumption that antique vehicles are operated significantly less

than average. States without any CDB submittals received EPA Default populations and age distributions based on the adjusted IHS data, and some states with submittals were overridden, decided on a case-by-case basis.

In addition to removing the older and antique plate vehicles from the IHS data, 28 counties found to be outliers because their fleet age was significantly younger than in typical counties. The outlier review was limited to LDV source types 21, 31, and 32. Many rural counties have outliers for low-population source types such as Transit Bus and Refuse Truck due to small sample sizes, but these do not have much of an impact on the inventory overall and reflect sparse data in low-population areas and therefore do not require correction.

The most extreme examples of LDV outliers were Light Commercial Truck age distributions where over 85 percent of the population in the entire county is 0 and 1 years old. These sorts of young fleets can happen if the headquarters of a leasing or rental company is the owner/entity of a relatively large number of vehicles relative to the county-wide population. While the business owner of thousands of new vehicles may reside in a single county, the vehicles likely operate in broader areas without being registered where they drive.

In areas where submitted vehicle population data were accepted for the 2020 NEI, the relative populations of cars vs. light-duty trucks were reapportioned (while retaining the magnitude of the light-duty vehicles from the submittals) using the county-specific percentages from the IHS data. In this way, the categorization of cars versus light trucks is consistent from state to state. The county total light-duty vehicle populations were preserved through this process.

To create the emission factors, MOVES was run separately for each representative county and fuel month and for each temperature bin needed for calendar year 2021. The CDBs used to run MOVES include the state-specific control measures such as the California low emission vehicle (LEV) program. In addition, the range of temperatures and the average humidities used in the CDBs were specific to the year 2021. The MOVES results were post-processed into CSV-formatted emission factor tables that can be read by SMOKE-MOVES.

2.3.4 Onroad California Inventory Development (onroad_ca_adj)

California uses their own emission model, EMFAC, to develop onroad emissions inventories and provides those inventories to EPA. EMFAC uses emission inventory codes (EICs) to characterize the emission processes instead of SCCs. The EPA and California worked together to develop a code mapping to better match EMFAC's EICs to EPA MOVES' detailed set of SCCs that distinguish between off-network and on-network and brake and tire wear emissions. This detail is needed for modeling but not for the NEI. California submitted onroad emissions for the 2020 NEI, as well as the year 2023. An interpolation between 2020 and 2023 was used for the 2021 modeling. Since California's 2023 inventory did not contain HAPs, VOC-based speciation factors were used to estimate VOC HAPs for 2021, and PM_{2.5}-based speciation factors to estimate metal HAPs. Other HAPs (e.g., PAHs) were computed using MOVES. The EPA added NH₃ to the CARB inventory by using the state total NH₃ from MOVES and allocating it at the county level based on CO. Refueling emissions were taken from MOVES for California.

California onroad mobile source emissions were created through a hybrid approach of combining state-supplied annual emissions with EPA-developed SMOKE-MOVES runs. Through this approach, the platform was able to reflect the California-developed emissions, while leveraging the more detailed SCCs and the highly resolved spatial patterns, temporal patterns, and speciation from SMOKE-MOVES. The basic steps involved in temporally allocating onroad emissions from California based on SMOKE-MOVES results were:

- 1) Run CA using EPA inputs through SMOKE-MOVES to produce hourly emissions hereafter known as “EPA estimates.” These EPA estimates for CA were run in a separate sector called “onroad_ca.”
- 2) Calculate ratios between state-supplied emissions and EPA estimates. The ratios were calculated for each county/SCC/pollutant combination based on the California onroad emissions inventory. The 2020 California data did not separate off and on-network emissions or extended idling, and also did not include information for vehicles fueled by E-85, so these differentiations were obtained using MOVES.
- 3) Create an adjustment factor file (CFPRO) that includes EPA-to-state estimate ratios.
- 4) Rerun CA through SMOKE-MOVES using EPA inputs and the new adjustment factor file.

Through this process, adjusted model-ready files were created that sum to annual totals from California, but have the temporal and spatial patterns reflecting the highly resolved meteorology and SMOKE-MOVES. After adjusting the emissions, this sector is called “onroad_ca_adj.” Note that in emission summaries, the emissions from the “onroad” and “onroad_ca_adj” sectors were summed and designated as the emissions for the onroad sector.

2.4 Nonroad Mobile sources (cmv, rail, nonroad)

The nonroad mobile source emission modeling sectors consist of nonroad equipment emissions (nonroad), locomotive (rail), and commercial marine CMV emissions (cmv_c1c2 and cmv_c3).

2.4.1 Category 1, Category 2 Commercial Marine Vessels (cmv_c1c2)

The cmv_c1c2 sector contains Category 1 and 2 (C1C2) CMV emissions. Category 1 and 2 vessels use diesel fuel. All emissions in this sector are annual and at county-SCC resolution; however, in the emissions modeling platform they are provided at the sub-county level (i.e., port shape ids) and by SCC and emission type (e.g., hoteling, maneuvering). For the 2021 emissions modeling platform EPA expanded the list of SCCs. SCCs are now further resolved based on ship type than they were for the 2020 NEI. A list of SCCs for the C1C2 sector can be seen in Table 2-16. For more information on CMV emissions development, see the supplemental documentation.⁵ C1C2 emissions that occur outside of state waters are not assigned to states. For this modeling platform, all CMV emissions in the cmv_c1c2 sector are treated as hourly gridded point sources with stack parameters that should result in them being placed in layer 1.

Sulfur dioxide (SO₂) emissions reflect rules that reduced sulfur emissions for CMV that took effect in the year 2015. The cmv_c1c2 inventory sector contains small to medium-size engine CMV emissions.

⁵ See https://gaftp.epa.gov/Air/emismod/2021/reports/C1C2_Documentation_2021_Final.pdf.

Category 1 and Category 2 marine diesel engines typically range in size from about 700 to 11,000 hp. These engines are used to provide propulsion power on many kinds of vessels including tugboats, towboats, supply vessels, fishing vessels, and other commercial vessels in and around ports. They are also used as stand-alone generators for auxiliary electrical power on many types of vessels. Category 1 represents engines up to 7 liters per cylinder displacement. Category 2 includes engines from 7 to 30 liters per cylinder.

The cmv_c1c2 inventory sector contains sources that traverse state and federal waters along with emissions from surrounding areas of Canada, Mexico, and international waters. The cmv_c1c2 sources are modeled as point sources but using plume rise parameters that cause the emissions to be released in the ground layer of the air quality model.

The cmv_c1c2 sources within state waters are identified in the inventory with the Federal Information Processing Standard (FIPS) county code for the state and county in which the vessel is registered. The cmv_c1c2 sources that operate outside of state waters but within the Emissions Control Area (ECA) are encoded with a state FIPS code of 85. The ECA areas include parts of the Gulf of Mexico, and parts of the Atlantic and Pacific coasts. The cmv_c1c2 sources are categorized as operating either in-port or underway and as main and auxiliary engines are encoded using the SCCs listed in Table 2-16. Level 1 and Level 2 descriptions for all of the entries are “Mobile Sources”, and “Marine Vessels, Commercial”, respectively.

Table 2-16. SCCs for the cmv_c1c2 sector

SCC	Level 3 Description	Level 4 Description
2280201113	Diesel Barge	C1C2 Port Emissions: Main Engine
2280202113	Diesel Offshore support	C1C2 Port Emissions: Main Engine
2280203113	Diesel Bulk Carrier	C1C2 Port Emissions: Main Engine
2280204113	Diesel Commercial Fishing	C1C2 Port Emissions: Main Engine
2280205113	Diesel Container Ship	C1C2 Port Emissions: Main Engine
2280206113	Diesel Ferry	C1C2 Port Emissions: Main Engine
2280207113	Diesel General Cargo	C1C2 Port Emissions: Main Engine
2280208113	Diesel Government	C1C2 Port Emissions: Main Engine
2280209113	Diesel Miscellaneous	C1C2 Port Emissions: Main Engine
2280210113	Diesel RollOn RollOff	C1C2 Port Emissions: Main Engine
2280211113	Diesel Tanker	C1C2 Port Emissions: Main Engine
2280212113	Diesel Tour Boat	C1C2 Port Emissions: Main Engine
2280213113	Diesel Tug	C1C2 Port Emissions: Main Engine
2280214113	Diesel Refrigerated	C1C2 Port Emissions: Main Engine
2280215113	Diesel Cruise	C1C2 Port Emissions: Main Engine
2280216113	Diesel Passenger Other	C1C2 Port Emissions: Main Engine
2280201114	Diesel Barge	C1C2 Port Emissions: Auxiliary Engine
2280202114	Diesel Offshore support	C1C2 Port Emissions: Auxiliary Engine
2280203114	Diesel Bulk Carrier	C1C2 Port Emissions: Auxiliary Engine
2280204114	Diesel Commercial Fishing	C1C2 Port Emissions: Auxiliary Engine
2280205114	Diesel Container Ship	C1C2 Port Emissions: Auxiliary Engine

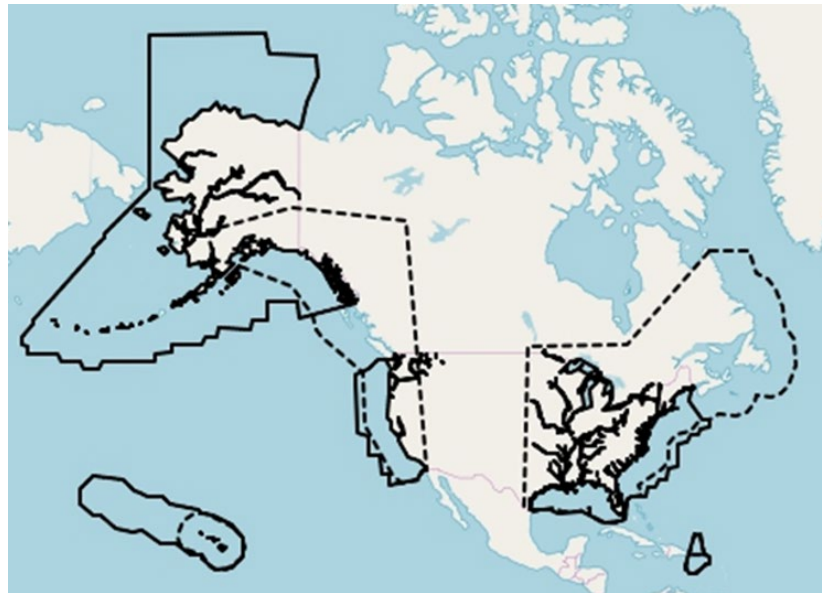
SCC	Level 3 Description	Level 4 Description
2280206114	Diesel Ferry	C1C2 Port Emissions: Auxiliary Engine
2280207114	Diesel General Cargo	C1C2 Port Emissions: Auxiliary Engine
2280208114	Diesel Government	C1C2 Port Emissions: Auxiliary Engine
2280209114	Diesel Miscellaneous	C1C2 Port Emissions: Auxiliary Engine
2280210114	Diesel RollOn RollOff	C1C2 Port Emissions: Auxiliary Engine
2280211114	Diesel Tanker	C1C2 Port Emissions: Auxiliary Engine
2280212114	Diesel Tour Boat	C1C2 Port Emissions: Auxiliary Engine
2280213114	Diesel Tug	C1C2 Port Emissions: Auxiliary Engine
2280214114	Diesel Refrigerated	C1C2 Port Emissions: Auxiliary Engine
2280215114	Diesel Cruise	C1C2 Port Emissions: Auxiliary Engine
2280216114	Diesel Passenger Other	C1C2 Port Emissions: Auxiliary Engine
2280201123	Diesel Barge	C1C2 Underway emissions: Main Engine
2280202123	Diesel Offshore support	C1C2 Underway emissions: Main Engine
2280203123	Diesel Bulk Carrier	C1C2 Underway emissions: Main Engine
2280204123	Diesel Commercial Fishing	C1C2 Underway emissions: Main Engine
2280205123	Diesel Container Ship	C1C2 Underway emissions: Main Engine
2280206123	Diesel Ferry	C1C2 Underway emissions: Main Engine
2280207123	Diesel General Cargo	C1C2 Underway emissions: Main Engine
2280208123	Diesel Government	C1C2 Underway emissions: Main Engine
2280209123	Diesel Miscellaneous	C1C2 Underway emissions: Main Engine
2280210123	Diesel RollOn RollOff	C1C2 Underway emissions: Main Engine
2280211123	Diesel Tanker	C1C2 Underway emissions: Main Engine
2280212123	Diesel Tour Boat	C1C2 Underway emissions: Main Engine
2280213123	Diesel Tug	C1C2 Underway emissions: Main Engine
2280214123	Diesel Refrigerated	C1C2 Underway emissions: Main Engine
2280215123	Diesel Cruise	C1C2 Underway emissions: Main Engine
2280216123	Diesel Passenger Other	C1C2 Underway emissions: Main Engine
2280201124	Diesel Barge	C1C2 Underway emissions: Auxiliary Engine
2280202124	Diesel Offshore support	C1C2 Underway emissions: Auxiliary Engine
2280203124	Diesel Bulk Carrier	C1C2 Underway emissions: Auxiliary Engine
2280204124	Diesel Commercial Fishing	C1C2 Underway emissions: Auxiliary Engine
2280205124	Diesel Container Ship	C1C2 Underway emissions: Auxiliary Engine
2280206124	Diesel Ferry	C1C2 Underway emissions: Auxiliary Engine
2280207124	Diesel General Cargo	C1C2 Underway emissions: Auxiliary Engine
2280208124	Diesel Government	C1C2 Underway emissions: Auxiliary Engine
2280209124	Diesel Miscellaneous	C1C2 Underway emissions: Auxiliary Engine
2280210124	Diesel RollOn RollOff	C1C2 Underway emissions: Auxiliary Engine
2280211124	Diesel Tanker	C1C2 Underway emissions: Auxiliary Engine
2280212124	Diesel Tour Boat	C1C2 Underway emissions: Auxiliary Engine
2280213124	Diesel Tug	C1C2 Underway emissions: Auxiliary Engine
2280214124	Diesel Refrigerated	C1C2 Underway emissions: Auxiliary Engine
2280215124	Diesel Cruise	C1C2 Underway emissions: Auxiliary Engine

SCC	Level 3 Description	Level 4 Description
2280216124	Diesel Passenger Other	C1C2 Underway emissions: Auxiliary Engine

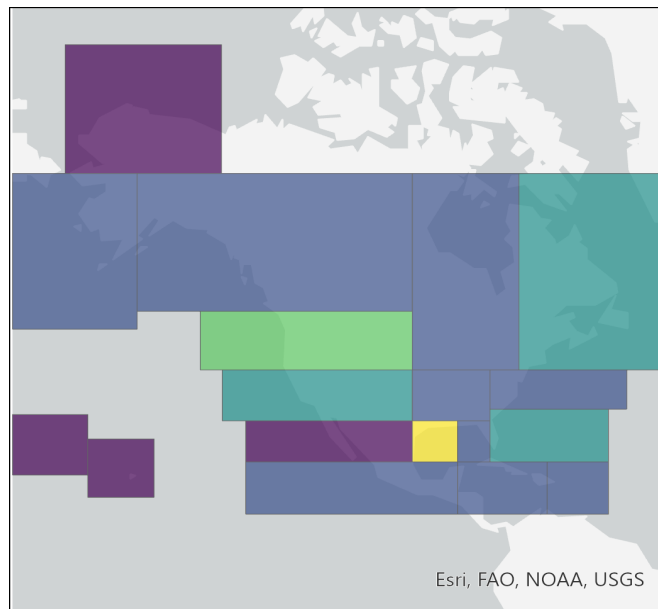
Category 1 and 2 CMV emissions were developed for the 2021 platform and were not based on 2020 NEI although the methods used to develop the emissions were similar. The emissions were developed based on signals from Automated Identification System (AIS) transmitters. AIS is a tracking system used by vessels to enhance navigation and avoid collision with other AIS transmitting vessels. The USEPA Office of Transportation and Air Quality received AIS data from the U.S. Coast Guard (USCG) to quantify all ship activity which occurred between January 1 and December 31, 2021. To ensure coverage for all of the areas needed by the NEI, the requested and provided AIS data extend beyond 200 nautical miles from the U.S. coast. The area covered by the NEI is shown in Figure 2-4 (a). This boundary is roughly equivalent to the border of the U.S Exclusive Economic Zone and the North American ECA, although some non-ECA activity are captured as well. Two types of AIS data were received: satellite (S-AIS) and terrestrial (T-AIS). The counts of data received for S-AIS and T-AIS for the 2021 emissions modeling platform are shown in Figure 2-4 (b).

Figure 2-4. NEI Commercial Marine Vessel Boundaries and Automatic Identification System Request Boxes for the 2021 Emissions Modeling Platform

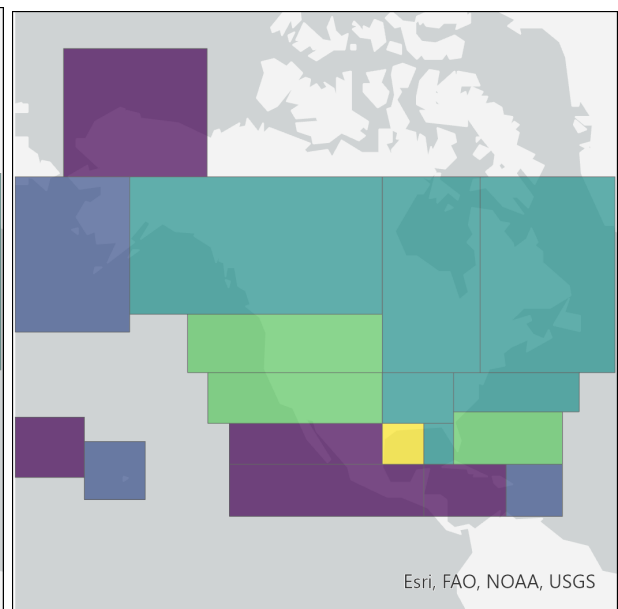
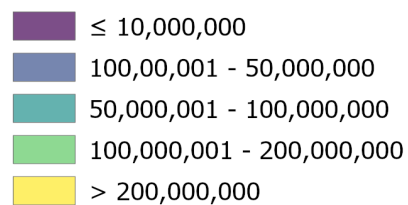
a) NEI (solid) and ECA (dashed) geographical extent



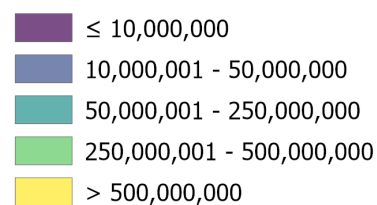
b) Areas of AIS request boxes and amount of data received



Num. Rows in S-AIS 2021



Num. Rows in T-AIS 2021



The AIS data were compiled into five-minute intervals by the USCG, providing a reasonably refined assessment of a vessel's movement. For example, using a five-minute average, a vessel traveling at 25 knots would be captured every two nautical miles that the vessel travels. For slower moving vessels, the distance between transmissions would be less. The ability to track vessel movements through AIS data and link them to attribute data, has allowed for the development of an inventory of very accurate emission estimates. These AIS data were used to define the locations of individual vessel movements, estimate hours of operation, and quantify propulsion engine loads. The compiled AIS data also included the vessel's International Marine Organization (IMO) number and Maritime Mobile Service Identifier (MMSI); which allowed each vessel to be matched to their characteristics obtained from the Clarksons ship registry (Clarksons, 2021).

The engine bore and stroke data were used to calculate cylinder volume. Any vessel that had a calculated cylinder volume greater than 30 liters was incorporated into the USEPA's new Category 3 Commercial Marine Vessel (C3CMV) model. The remaining records were assumed to represent Category 1 and 2 (C1C2) or non-ship activity. The C1C2 AIS data were quality assured including the removal of duplicate messages, signals from pleasure craft, and signals that were not from CMV vessels (e.g., buoys, helicopters, and vessels that are not self-propelled).

The emissions were calculated for each time interval between consecutive AIS messages for each vessel and allocated to the location of the message following to the interval. Emissions were calculated according to Equation 2-1.

$$Emissions_{interval} = Time (hr)_{interval} \times Power(kW) \times EF\left(\frac{g}{kWh}\right) \times LLAF \quad \text{Equation 2-1}$$

Power was calculated for the propulsive (main), auxiliary, and auxiliary boiler engines for each interval and emission factor (EF) reflects the assigned emission factors for each engine, as described below. LLAF represents the low load adjustment factor, a unitless factor which reflects increasing propulsive emissions during low load operations. Time indicates the activity duration time between consecutive intervals.

Next, vessels were identified to determine their vessel type, and thus their vessel group, power rating, and engine tier information which are required for the emissions calculations. See the 2021 C1 C2 CMV development documentation for more details on this process. Following the identification, 236 different vessel types were matched to the C1C2 vessels. Vessel attribute data were not available for all these vessel types, so the vessel types were aggregated into 13 different vessel groups for which surrogate data were available as shown in Table 2-17. 19,106 vessels were directly identified by their ship and cargo number. The remaining group of miscellaneous ships represent 1.5 percent of the AIS vessels (excluding recreational vessels) for which a specific vessel type could not be assigned.

Table 2-17. Vessel groups in the cmv_c1c2 sector

Vessel Group	2021 Area Ship Count
Bulk Carrier	46
Commercial Fishing	5,826
Container Ship	11
Ferry Excursion	849
General Cargo	3,190
Government	1,179
Miscellaneous	291
Offshore support	1,416
Pilot	15
Reefer	12
Ro Ro	219
Tanker	591
Tug	5,299
Work Boat	162
Total in Inventory:	19,106

As shown in Equation 2-1, power is an important component of the emissions computation. Vessel-specific installed propulsive power ratings and service speeds were pulled from Clarksons ship registry and adopted from the Global Fishing Watch (GFW) dataset when available. However, there is limited vessel specific attribute data for most of the C1C2 fleet. This necessitated the use of surrogate engine power and load factors, which were computed for each vessel group. In addition to the power required by propulsive engines, power needs for auxiliary engines were also computed for each vessel group. Emissions from main and auxiliary engines are inventoried with different SCCs as shown in Table 2-16.

The final components of the emissions computation equation are the emission factors and the low load adjustment factor. The emission factors used in this inventory take into consideration the EPA's marine vessel fuel regulations as well as exhaust standards that are based on the year that the vessel was manufactured to determine the appropriate regulatory tier. Emission factors in g/kWhr by tier for NO_x, PM₁₀, PM_{2.5}, CO, CO₂, SO₂ and VOC were developed using Tables 3-7 through 3-10 in USEPA's (2008) Regulatory Impact Analysis on engines less than 30 liters per cylinder. To compile these emissions factors, population-weighted average emission factors were calculated per tier based on C1C2 population distributions grouped by engine displacement. Boiler emission factors were obtained from an earlier Swedish Environmental Protection Agency study (Swedish EPA, 2004). If the year of manufacture was unknown then it was assumed that the vessel was Tier 0, such that actual emissions may be less than those estimated in this inventory. Without more specific data, the magnitude of this emissions difference cannot be estimated.

Propulsive emissions from low-load operations were adjusted to account for elevated emission rates associated with activities outside the engines' optimal operating range. The emission factor adjustments were applied by load and pollutant, based on the data compiled for the Port Everglades 2015 Emission

Inventory.⁶ Hazardous air pollutants and ammonia were added to the inventory according to multiplicative factors applied either to VOC or PM_{2.5}.

The stack parameters used for cmv_c1c2 are a stack height of 1 ft, stack diameter of 1 ft, stack temperature of 70°F, and a stack velocity of 0.1 ft/s. These parameters force emissions into layer 1.

For more information on the emission computations for 2021, see the supporting documentation for the development of the 2021 C1C2 CMV emissions. The cmv_c1c2 emissions were aggregated to total hourly values in each grid cell and run through SMOKE as point sources. SMOKE requires an annual inventory file to go along with the hourly data and this file was generated for 2021.

2.4.2 Category 3 Commercial Marine Vessels (cmv_c3)

The cmv_c3 sector contains large engine CMV emissions. Category 3 (C3) marine diesel engines at or above 30 liters per cylinder. Typically, these are the largest CMV engines and are rated at 3,000 to 100,000 hp. C3 engines are typically used for propulsion on ocean-going vessels including container ships, oil tankers, bulk carriers, and cruise ships. Emissions control technologies for C3 CMV sources are limited due to the nature of the residual fuel used by these vessels.⁷ The cmv_c3 sector contains sources that traverse state and federal waters; along with sources in waters not covered by the NEI in surrounding areas of Canada, Mexico, and international waters. For more information on development of C3 emissions refer to Section 11 of the 2020 NEI TSD for information on the C3 methodology.

The cmv_c3 sources that operate outside of state waters but within the federal Emissions Control Area (ECA) are encoded with a FIPS state code of 85, with the “county code” digits representing broad regions such as the Atlantic, Gulf of Mexico, and Pacific. The ECA areas include parts of the Gulf of Mexico, and parts of the Atlantic and Pacific coasts. CMV C3 sources around Puerto Rico, Hawaii and Alaska, which are outside the ECA areas, are included in the inventory but are in separate files from the emissions around the continental United States (CONUS). The cmv_c3 sources in the inventory are categorized as operating either in-port or underway and are encoded using the SCCs listed in Table 2-18 and distinguish between diesel and residual fuel, in port areas versus underway, and main and auxiliary engines. The Level 1 and Level 2 descriptions for each of the SCCs are “Mobile Sources” and “Marine Vessels, Commercial”, respectively.

Table 2-18. SCCs for cmv_c3 sector

SCC	Level 3 Description	Level 4 Description
2280201313	Diesel Barge	C3 Port Emissions: Main Engine
2280202313	Diesel Offshore support	C3 Port Emissions: Main Engine
2280203313	Diesel Bulk Carrier	C3 Port Emissions: Main Engine
2280204313	Diesel Commercial Fishing	C3 Port Emissions: Main Engine

⁶ USEPA. EPA and Port Everglades Partnership: Emission Inventories and Reduction Strategies. US Environmental Protection Agency, Office of Transportation and Air Quality, June 2018.

<https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100UKV8.pdf>.

⁷ <https://www.epa.gov/regulations-emissions-vehicles-and-engines/regulations-emissions-marine-vessels>.

SCC	Level 3 Description	Level 4 Description
2280205313	Diesel Container Ship	C3 Port Emissions: Main Engine
2280206313	Diesel Ferry	C3 Port Emissions: Main Engine
2280207313	Diesel General Cargo	C3 Port Emissions: Main Engine
2280208313	Diesel Government	C3 Port Emissions: Main Engine
2280209313	Diesel Miscellaneous	C3 Port Emissions: Main Engine
2280210313	Diesel RollOn RollOff	C3 Port Emissions: Main Engine
2280211313	Diesel Tanker	C3 Port Emissions: Main Engine
2280212313	Diesel Tour Boat	C3 Port Emissions: Main Engine
2280213313	Diesel Tug	C3 Port Emissions: Main Engine
2280214313	Diesel Refrigerated	C3 Port Emissions: Main Engine
2280215313	Diesel Cruise	C3 Port Emissions: Main Engine
2280216313	Diesel Passenger Other	C3 Port Emissions: Main Engine
2280201314	Diesel Barge	C3 Port Emissions: Auxiliary Engine
2280202314	Diesel Offshore support	C3 Port Emissions: Auxiliary Engine
2280203314	Diesel Bulk Carrier	C3 Port Emissions: Auxiliary Engine
2280204314	Diesel Commercial Fishing	C3 Port Emissions: Auxiliary Engine
2280205314	Diesel Container Ship	C3 Port Emissions: Auxiliary Engine
2280206314	Diesel Ferry	C3 Port Emissions: Auxiliary Engine
2280207314	Diesel General Cargo	C3 Port Emissions: Auxiliary Engine
2280208314	Diesel Government	C3 Port Emissions: Auxiliary Engine
2280209314	Diesel Miscellaneous	C3 Port Emissions: Auxiliary Engine
2280210314	Diesel RollOn RollOff	C3 Port Emissions: Auxiliary Engine
2280211314	Diesel Tanker	C3 Port Emissions: Auxiliary Engine
2280212314	Diesel Tour Boat	C3 Port Emissions: Auxiliary Engine
2280213314	Diesel Tug	C3 Port Emissions: Auxiliary Engine
2280214314	Diesel Refrigerated	C3 Port Emissions: Auxiliary Engine
2280215314	Diesel Cruise	C3 Port Emissions: Auxiliary Engine
2280216314	Diesel Passenger Other	C3 Port Emissions: Auxiliary Engine
2280201323	Diesel Barge	C3 Underway emissions: Main Engine
2280202323	Diesel Offshore support	C3 Underway emissions: Main Engine
2280203323	Diesel Bulk Carrier	C3 Underway emissions: Main Engine
2280204323	Diesel Commercial Fishing	C3 Underway emissions: Main Engine
2280205323	Diesel Container Ship	C3 Underway emissions: Main Engine
2280206323	Diesel Ferry	C3 Underway emissions: Main Engine
2280207323	Diesel General Cargo	C3 Underway emissions: Main Engine
2280208323	Diesel Government	C3 Underway emissions: Main Engine
2280209323	Diesel Miscellaneous	C3 Underway emissions: Main Engine
2280210323	Diesel RollOn RollOff	C3 Underway emissions: Main Engine
2280211323	Diesel Tanker	C3 Underway emissions: Main Engine

SCC	Level 3 Description	Level 4 Description
2280212323	Diesel Tour Boat	C3 Underway emissions: Main Engine
2280213323	Diesel Tug	C3 Underway emissions: Main Engine
2280214323	Diesel Refrigerated	C3 Underway emissions: Main Engine
2280215323	Diesel Cruise	C3 Underway emissions: Main Engine
2280216323	Diesel Passenger Other	C3 Underway emissions: Main Engine
2280201324	Diesel Barge	C3 Underway emissions: Auxiliary Engine
2280202324	Diesel Offshore support	C3 Underway emissions: Auxiliary Engine
2280203324	Diesel Bulk Carrier	C3 Underway emissions: Auxiliary Engine
2280204324	Diesel Commercial Fishing	C3 Underway emissions: Auxiliary Engine
2280205324	Diesel Container Ship	C3 Underway emissions: Auxiliary Engine
2280206324	Diesel Ferry	C3 Underway emissions: Auxiliary Engine
2280207324	Diesel General Cargo	C3 Underway emissions: Auxiliary Engine
2280208324	Diesel Government	C3 Underway emissions: Auxiliary Engine
2280209324	Diesel Miscellaneous	C3 Underway emissions: Auxiliary Engine
2280210324	Diesel RollOn RollOff	C3 Underway emissions: Auxiliary Engine
2280211324	Diesel Tanker	C3 Underway emissions: Auxiliary Engine
2280212324	Diesel Tour Boat	C3 Underway emissions: Auxiliary Engine
2280213324	Diesel Tug	C3 Underway emissions: Auxiliary Engine
2280214324	Diesel Refrigerated	C3 Underway emissions: Auxiliary Engine
2280215324	Diesel Cruise	C3 Underway emissions: Auxiliary Engine
2280216324	Diesel Passenger Other	C3 Underway emissions: Auxiliary Engine
2280301313	Residual Barge	C3 Port Emissions: Main Engine
2280302313	Residual Offshore support	C3 Port Emissions: Main Engine
2280303313	Residual Bulk Carrier	C3 Port Emissions: Main Engine
2280304313	Residual Commercial Fishing	C3 Port Emissions: Main Engine
2280305313	Residual Container Ship	C3 Port Emissions: Main Engine
2280306313	Residual Ferry	C3 Port Emissions: Main Engine
2280307313	Residual General Cargo	C3 Port Emissions: Main Engine
2280308313	Residual Government	C3 Port Emissions: Main Engine
2280309313	Residual Miscellaneous	C3 Port Emissions: Main Engine
2280310313	Residual RollOn RollOff	C3 Port Emissions: Main Engine
2280311313	Residual Tanker	C3 Port Emissions: Main Engine
2280312313	Residual Tour Boat	C3 Port Emissions: Main Engine
2280313313	Residual Tug	C3 Port Emissions: Main Engine
2280314313	Residual Refrigerated	C3 Port Emissions: Main Engine
2280315313	Residual Cruise	C3 Port Emissions: Main Engine
2280316313	Residual Passenger Other	C3 Port Emissions: Main Engine
2280301314	Residual Barge	C3 Port Emissions: Auxiliary Engine
2280302314	Residual Offshore support	C3 Port Emissions: Auxiliary Engine

SCC	Level 3 Description	Level 4 Description
2280303314	Residual Bulk Carrier	C3 Port Emissions: Auxiliary Engine
2280304314	Residual Commercial Fishing	C3 Port Emissions: Auxiliary Engine
2280305314	Residual Container Ship	C3 Port Emissions: Auxiliary Engine
2280306314	Residual Ferry	C3 Port Emissions: Auxiliary Engine
2280307314	Residual General Cargo	C3 Port Emissions: Auxiliary Engine
2280308314	Residual Government	C3 Port Emissions: Auxiliary Engine
2280309314	Residual Miscellaneous	C3 Port Emissions: Auxiliary Engine
2280310314	Residual RollOn RollOff	C3 Port Emissions: Auxiliary Engine
2280311314	Residual Tanker	C3 Port Emissions: Auxiliary Engine
2280312314	Residual Tour Boat	C3 Port Emissions: Auxiliary Engine
2280313314	Residual Tug	C3 Port Emissions: Auxiliary Engine
2280314314	Residual Refrigerated	C3 Port Emissions: Auxiliary Engine
2280315314	Residual Cruise	C3 Port Emissions: Auxiliary Engine
2280316314	Residual Passenger Other	C3 Port Emissions: Auxiliary Engine
2280301323	Residual Barge	C3 Underway emissions: Main Engine
2280302323	Residual Offshore support	C3 Underway emissions: Main Engine
2280303323	Residual Bulk Carrier	C3 Underway emissions: Main Engine
2280304323	Residual Commercial Fishing	C3 Underway emissions: Main Engine
2280305323	Residual Container Ship	C3 Underway emissions: Main Engine
2280306323	Residual Ferry	C3 Underway emissions: Main Engine
2280307323	Residual General Cargo	C3 Underway emissions: Main Engine
2280308323	Residual Government	C3 Underway emissions: Main Engine
2280309323	Residual Miscellaneous	C3 Underway emissions: Main Engine
2280310323	Residual RollOn RollOff	C3 Underway emissions: Main Engine
2280311323	Residual Tanker	C3 Underway emissions: Main Engine
2280312323	Residual Tour Boat	C3 Underway emissions: Main Engine
2280313323	Residual Tug	C3 Underway emissions: Main Engine
2280314323	Residual Refrigerated	C3 Underway emissions: Main Engine
2280315323	Residual Cruise	C3 Underway emissions: Main Engine
2280316323	Residual Passenger Other	C3 Underway emissions: Main Engine
2280301324	Residual Barge	C3 Underway emissions: Auxiliary Engine
2280302324	Residual Offshore support	C3 Underway emissions: Auxiliary Engine
2280303324	Residual Bulk Carrier	C3 Underway emissions: Auxiliary Engine
2280304324	Residual Commercial Fishing	C3 Underway emissions: Auxiliary Engine
2280305324	Residual Container Ship	C3 Underway emissions: Auxiliary Engine
2280306324	Residual Ferry	C3 Underway emissions: Auxiliary Engine
2280307324	Residual General Cargo	C3 Underway emissions: Auxiliary Engine
2280308324	Residual Government	C3 Underway emissions: Auxiliary Engine
2280309324	Residual Miscellaneous	C3 Underway emissions: Auxiliary Engine

SCC	Level 3 Description	Level 4 Description
2280310324	Residual RollOn RollOff	C3 Underway emissions: Auxiliary Engine
2280311324	Residual Tanker	C3 Underway emissions: Auxiliary Engine
2280312324	Residual Tour Boat	C3 Underway emissions: Auxiliary Engine
2280313324	Residual Tug	C3 Underway emissions: Auxiliary Engine
2280314324	Residual Refrigerated	C3 Underway emissions: Auxiliary Engine
2280315324	Residual Cruise	C3 Underway emissions: Auxiliary Engine
2280316324	Residual Passenger Other	C3 Underway emissions: Auxiliary Engine

The EPA received Automated Identification System (AIS) data from United States Coast Guard (USCG) to quantify all ship activity which occurred between January 1 and December 31, 2021. The International Maritime Organization's (IMO's) International Convention for the Safety of Life at Sea (SOLAS) requires AIS to be fitted aboard all international voyaging ships with gross tonnage of 300 or more, and all passenger ships regardless of size.⁸ In addition, the USCG has mandated that all commercial marine vessels continuously transmit AIS signals while transiting U.S. navigable waters. As the vast majority of C3 vessels meet these requirements, any omitted from the inventory due to lack of AIS adoption are deemed to have a negligible impact on national C3 emissions estimates. The activity data incorporated into this inventory reflect ship operations within 200 nautical miles of the official U.S. baseline and beyond. Activity data within the border of the U.S Exclusive Economic Zone and the North American ECA are included as well as some activity data outside of the ECA.

The 2021 CMV emissions modeling platform data were computed based on the AIS data from the USGS for the year of 2021. The AIS data were coupled with ship registry data that contained engine parameters, vessel power parameters, and other factors such as tonnage and year of manufacture which helped to separate the C3 vessels from the C1C2 vessels. Where specific ship parameters were not available, they were gap-filled. The types of vessels that remain in the C3 data set include bulk carrier, chemical tanker, liquified gas tanker, oil tanker, other tanker, container ship, cruise, ferry, general cargo, fishing, refrigerated vessel, roll-on/roll-off, tug, and yacht.

Prior to use, the AIS data were reviewed - data deemed to be erroneous were removed, and data found to be at intervals greater than 5 minutes were interpolated to ensure that each ship had data every five minutes. The five-minute average data provide a reasonably refined assessment of a vessel's movement. For example, using a five-minute average, a vessel traveling at 25 knots would be captured every two nautical miles that the vessel travels. For slower moving vessels, the distance between transmissions would be less.

The emissions were calculated for each C3 vessel in the dataset for each 5-minute time range and allocated to the location of the message following to the interval. Emissions were calculated according to Equation 2-2.

⁸ International Maritime Organization (IMO) Resolution MSC.99(73) adopted December 12th, 2000 and entered into force July 1st, 2002; as amended by SOLAS Resolution CONF.5/32 adopted December 13th, 2002.

$$Emissions_{interval} = Time (hr)_{interval} \times Power(kW) \times EF(\frac{g}{kWh}) \times LLAFF \quad \text{Equation 2-2}$$

Power is calculated for the propulsive (main), auxiliary, and auxiliary boiler engines for each interval and emission factor (EF) reflects the assigned emission factors for each engine, as described below. LLAFF represents the low load adjustment factor, a unitless factor which reflects increasing propulsive emissions during low load operations. Time indicates the activity duration time between consecutive intervals.

Emissions were computed according to a computed power need (kW) multiplied by the time (hr) and by an engine-specific emission factor (g/kWh) and finally by a low load adjustment factor that reflects increasing propulsive emissions during low load operations.

The resulting emissions were available at 5-minute intervals. Code was developed to aggregate these emissions to modeling grid cells and up to hourly levels so that the emissions data could be input to SMOKE for emissions modeling with SMOKE. Within SMOKE, the data were speciated into the pollutants needed by the air quality model,⁹ but since the data were already in the form of point sources at the center of each grid cell, and they were already hourly, no other processing was needed within SMOKE. SMOKE requires an annual inventory file to go along with the hourly data, so this annual file was generated for 2021.

On January 1st, 2015, the ECA initiated a fuel sulfur standard which regulated large marine vessels to use fuel with 1,000 ppm sulfur or less. These standards are reflected in the cmv_c3 inventories.

The resulting point emissions centered on each grid cell were converted to an annual point 2010 flat file format (FF10). A set of standard stack parameters were assigned to each release point in the cmv_c3 inventory. The assigned stack height was 65.62 ft, the stack diameter was 2.625 ft, the stack temperature was 539.6 °F, and the velocity was 82.02 ft/s. Emissions were computed for each grid cell needed for modeling.

2.4.3 Railway Locomotives (rail)

The rail sector includes all locomotives in the NEI nonpoint data category including line haul locomotives on Class 1, 2, and 3 railroads along with emissions from commuter rail lines and Amtrak. The rail sector excludes railway maintenance locomotives and point source yard locomotives. Railway maintenance emissions are included in the nonroad sector. The point source yard locomotives are included in the ptnonipm sector.

The rail emissions for the 2021 emissions modeling platform are based on the 2020 NEI. A projection factor of 1.05 was applied across all rail SCCs for the 2021 modeling platform (including railyards). This was based on Annual Energy Outlook (AEO). The 2021 AEO shows a slight increase in fuel use from 2020 to 2021 (2021/2020=1.05). The 2020 NEI is based on methods developed during the development of the 2017 NEI rail inventory by the Lake Michigan Air Directors Consortium (LADCO) and the State of Illinois

International Maritime Organization (IMO) Resolution MSC.99(73).

with support from various other states. Class I railroad emissions are based on confidential link-level line-haul activity GIS data layer maintained by the Federal Railroad Administration (FRA). In addition, the Association of American Railroads (AAR) provided national emission tier fleet mix information. Class II and III railroad emissions are based on a comprehensive nationwide GIS database of locations where short line and regional railroads operate. Passenger rail (Amtrak) emissions follow a similar procedure as Class II and III, except using a database of Amtrak rail lines. Yard locomotive emissions are based on a combination of yard data provided by individual rail companies, and by using Google Earth and other tools to identify rail yard locations for rail companies which did not provide yard data. Information on specific yards were combined with fuel use data and emission factors to create an emissions inventory for rail yards. Pollutant-specific factors were applied on top of the activity-based changes for the Class I rail. The inventory SCCs are shown in Table 2-19. More detailed information on the development of the 2021 emission modeling platform rail inventory is available in the 2020 NEI TSD and in the [Rail 2020 National Emissions Inventory Supplementary Document](#) on the 2020 NEI supporting data FTP site.

Table 2-19. SCCs for the Rail Sector

SCC	Sector	Description: Mobile Sources prefix for all
2285002006	Rail	Railroad Equipment; Diesel; Line Haul Locomotives: Class I Operations
2285002007	Rail	Railroad Equipment; Diesel; Line Haul Locomotives: Class II / III Operations
2285002008	Rail	Railroad Equipment; Diesel; Line Haul Locomotives: Passenger Trains (Amtrak)
2285002009	Rail	Railroad Equipment; Diesel; Line Haul Locomotives: Commuter Lines
2285002010	Rail	Railroad Equipment; Diesel; Yard Locomotives (nonpoint)
228500201	Rail	Railroad Equipment; Diesel; Yard Locomotives (point)

Class I Line-haul Methodology

For the 2020 inventory, the Class I railroads granted EPA permission to use the confidential link-level line haul activity geographic information system (GIS) data layer maintained and updated annually by the Federal Railroad Administration (FRA). At the time of inventory development, 2019 million gross ton (MGT) data was the most recent and complete data available. A map of the Class I railroad lines is shown in Figure 2-5. The dataset contains three columns indicating railroad ownership and nine columns indicating trackage rights for each rail segment. While most rail links have a single owner, some links have up to six different Class 1 railroad companies operating on it. To prepare the FRA data for use in the Class I line haul calculations, all segments associated with a railroad company were extracted to identify the full network for each company. This involved iterating through each of those twelve columns to identify all segments within each railroad company's network. This process was conducted seven times, one for each Class I railroad company. This resulted in a complete inventory of rail links trafficked by each Class I railroads with a record for each link/railroad company combination.

Figure 2-5. 2019 Class I Railroad Line Haul Activity



EPA collected 2020 Class I line haul fuel use data from the most recent R-1 submittals from the Surface Transportation Board.¹⁰ Consistent with previous inventory efforts, EPA summed line haul and work train fuel usage, Table 2-20.

Table 2-20. 2020 R-1 Reported Locomotive Fuel Use for Class I Railroads

Class I Railroad	Line Haul Fuel Use (gal)*
BNSF	1,137,598,007
Canadian National (CN)	96,337,392
Canadian Pacific (CPRS)	57,664,407
CSX Transportation (CSXT)	327,917,859
Kansas City Southern (KCS)	55,763,748
Norfolk Southern (NS)	342,470,779
Union Pacific (UP)	773,476,896

* Includes work train fuel usage

¹⁰ Surface Transportation Board. Available at <https://www.stb.gov/reports-data/economic-data/annual-report-financial-data/>. Retrieved 22 June 2021.

The Association of American Railroads (AAR) provided national Class I locomotive tier fleet mix information that reflects engine turnover in the nation. Given the impact of the pandemic in 2020, AAR provided a fleet mix that reflected active locomotives and excluded those that were held in storage. A locomotive's Tier level determines its allowable emission rates based on the year when it was built and/or re-manufactured. More accurate emission factors for each pollutant were calculated based on the percentage of the operating Class I line haul locomotives for each USEPA Tier-level category.

Class II and III Methodology

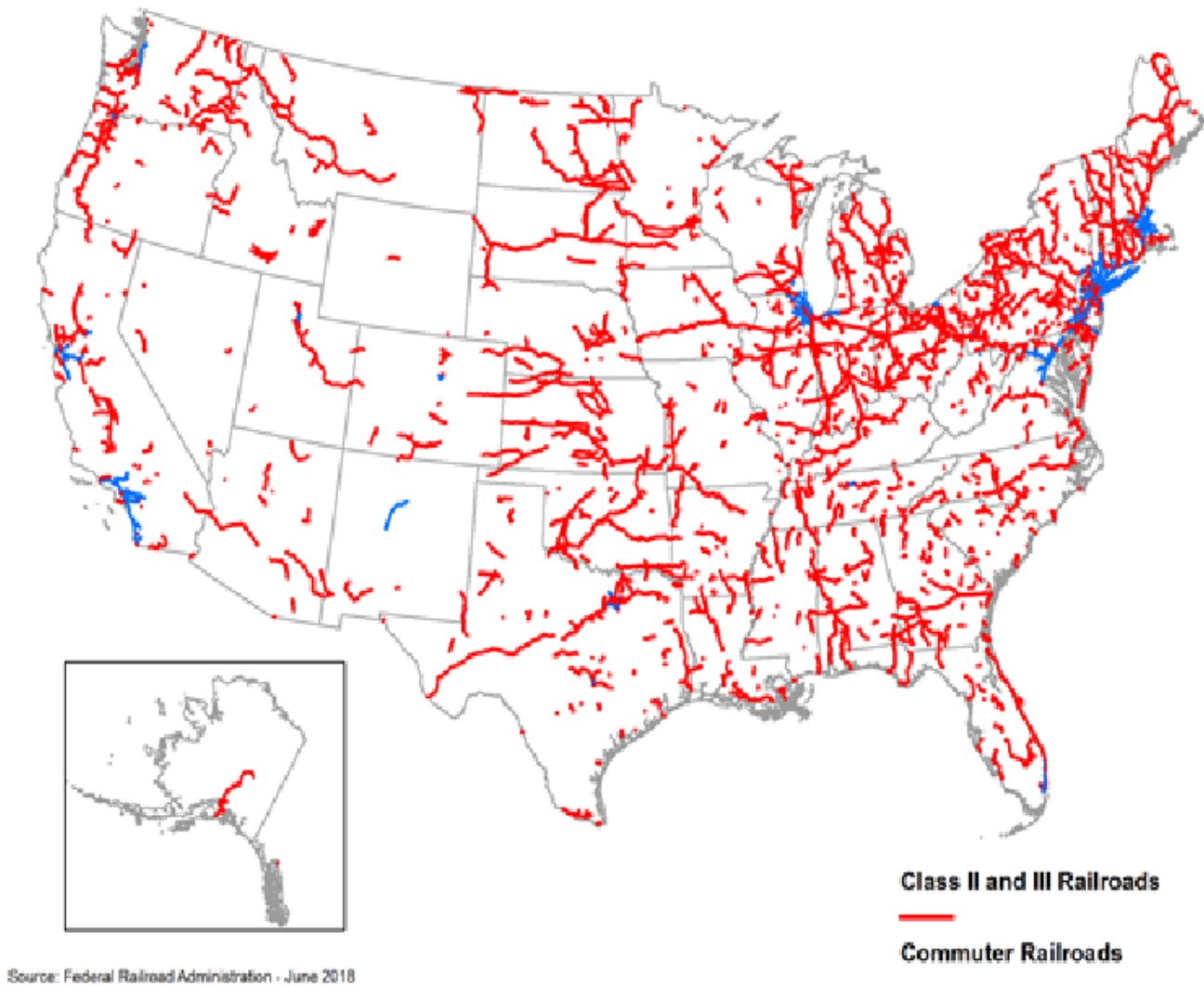
There are approximately 630 Class II and III Railroads operating in the United States, most of which are members of the American Short Line and Regional Railroad Association (ASLRRA). Data on Class II and III locomotive operations is publicly available from Bureau of Transportation Statistics' National Transportation Atlas Database (NTAD), along with related data including reporting mark, railroad name, route miles owned or operated, and total route miles of links.

Class II and III railroads are widely dispersed across the country (see Figure 2-6), often utilizing older, higher emitting locomotives than their Class I counterparts. AAR provided a national line-haul tier fleet mix profile for 2020 which reflects the trend toward older engines in this sector as shown in Table 2-21. The national fleet mix data was then used to calculate weighted average in-use emissions factors for the locomotives operated by the Class II and III railroads. Note that to be consistent with the 2017 inventory, the unweighted emission factors were the same as the Class I line haul due to the conservative use of the EPA's large locomotive conversion factor of 20.8 bhp-hr/gal. Emission factors for PM_{2.5}, SO₂, NH₃, VOC, and GHGs were calculated in the same manner as those used for Class I line-haul inventory described above.

Table 2-21. 2020 Class II/III Line Haul Fleet by Tier Level

Tier	2020 Class II/III Locomotive Count	Percent of Total Fleet
0	1,664	48%
1	31	1%
2	169	5%
3	160	5%
4	64	2%
Not Classified	1,359	39%
Total	3,447	100%

Figure 2-6. Class II and III Railroads in the United States



Commuter Rail Methodology

Commuter rail emissions were calculated in the same way as the Class II and III railroads. The primary difference is that the fuel use estimates were based on data collected by the Federal Transit Administration (FTA) for the National Transit Database. These fuel use estimates were replaced with reported fuel use statistics for MBTA (Massachusetts) and Metra (Illinois). The commuter railroads were separated from the Class II and III railroads so that the appropriate SCC codes could be entered into the emissions calculation sheet.

Intercity Passenger Methodology (Amtrak)

The calculation methodology mimics that used for the Class II and III and commuter railroads with a few modifications. Since link-level activity data for Amtrak was unavailable, the default assumption was made to evenly distribute Amtrak's 2020 reported fuel use across all of its diesel-powered route-miles shown in Figure 2-7.

Figure 2-7. Amtrak National Rail Network



Other Data Sources

The 2020 NEI locomotives sector includes data from SLT agency-provided emissions data, and an EPA dataset of locomotive emissions. The following agencies also submitted emissions to locomotive SCCs: Alaska Department of Environmental Conservation; California; Connecticut; District of Columbia; Maricopa County, AZ; Minnesota; North Carolina; Texas; Virginia; Washington; and Washoe County, NV.

2.4.4 Nonroad Mobile Equipment (nonroad)

The mobile nonroad equipment sector includes all mobile source emissions that do not operate on roads, excluding commercial marine vehicles, railways, and aircraft. Types of nonroad equipment include recreational vehicles, pleasure craft, and construction, agricultural, mining, and lawn and garden equipment. Nonroad equipment emissions were computed by running MOVES4 which incorporates the NONROAD model. MOVES3 incorporated updated nonroad engine population growth rates, nonroad Tier 4 engine emission rates, and sulfur levels of nonroad diesel fuels. The only change made for nonroad from MOVES3 to MOVES4 was a change to fuel properties. MOVES provides a complete set of

HAPs and incorporates updated nonroad emission factors for HAPs. MOVES4 was used for all states other than California, which uses their own model. California nonroad emissions were provided by the California Air Resources Board (CARB) for the 2020 NEI, as well as 2023. For the 2021 emissions modeling platform CARB nonroad emissions were interpolated between 2020 and 2023. CARB emissions were used in California for all pollutants except PAHs and CO₂, which were taken from MOVES.

MOVES creates a monthly emissions inventory for criteria air pollutants (CAPs) and a full set of HAPs, plus additional pollutants such as NONHAPTOG and ETHANOL, which are not included in the NEI but are used for speciation. MOVES provides estimates of NONHAPTOG along with the speciation profile code for the NONHAPTOG emission source. This was accomplished by using NHTOG#### as the pollutant code in the Flat File 2010 (FF10) inventory file that can be read into SMOKE, where #### is a speciation profile code. For California, NHTOG####-VOC and HAP-VOC ratios from MOVES-based emissions were applied to VOC emissions so that VOC emissions can be speciated consistently with other states.

MOVES also provides estimates of PM_{2.5} by speciation profile code for the PM_{2.5} emission source, using PM25_#### as the pollutant code in the FF10 inventory file, where #### is a speciation profile code. To facilitate calculation of PMC within SMOKE, and to help create emissions summaries, an additional pollutant representing total PM_{2.5} called PM25TOTAL was added to the inventory. As with VOC, PM25_####-PM25TOTAL ratios were calculated and applied to PM_{2.5} emissions in California so that PM_{2.5} emissions in California can be speciated consistently with other states.

MOVES4 outputs emissions data in county-specific databases, and a post-processing script converts the data into FF10 format. Additional post-processing steps were performed as follows:

- County-specific FF10s were combined into a single FF10 file.
- Emissions were aggregated from the more detailed SCCs modeled in MOVES to the SCCs modeled in SMOKE. A list of the aggregated SMOKE SCCs is in Appendix A of the 2016v1 platform nonroad specification sheet (NEIC, 2019).
- To reduce the size of the inventory, HAPs not needed for air quality modeling, such as dioxins and furans, were removed from the inventory.
- To reduce the size of the inventory further, all emissions for sources (identified by county/SCC) for which CAP emissions totaling less than 1×10^{-10} were removed from the inventory. The MOVES model attributes a very tiny amount of emissions to sources that are actually zero, for example, snowmobile emissions in Florida. Removing these sources from the inventory reduces the total size of the inventory by about 7%.
- Gas and particulate components of HAPs that come out of MOVES separately, such as naphthalene, were combined.
- VOC was renamed VOC_INV so that SMOKE does not speciate both VOC and NONHAPTOG, which would result in a double count.
- PM25TOTAL, referenced above, was also created at this stage of the process.

- Emissions for airport ground support vehicles (SCCs ending in -8005), and oil field equipment (SCCs ending in -10010), were removed from the inventory at this stage, to prevent a double count with the airports and np_oilgas sectors, respectively.
- California emissions from MOVES were deleted and replaced with the CARB-supplied emissions.

National Updates: Agricultural and Construction Equipment Allocation

The modified MOVES default database for MOVES3 containing refinements to construction and agricultural sectors, (movesdb20220105_nupdates) and state-submitted inputs in CDBs were used to run MOVES-Nonroad to produce emissions for all states other than California. California-submitted emissions were used. Updated *nrsurrogate*, *nrstatesurrogate*, and *nrbaseyearequippopulation* tables, along with instructions for utilizing these tables in MOVES runs, are available for download from EPA's ftp site: <https://gaftp.epa.gov/air/emismod/2016/v1/reports/nonroad/>).

Emissions Inside California

California nonroad emissions were provided by CARB for the 2020 NEI and 2023. The 2021 emissions were interpolated between 2020 and 2023 where pollutants were available in both data sets. All California nonroad inventories are annual, with monthly temporalization applied in SMOKE. Emissions for oil field equipment (SCCs ending in -10010) were removed from the California inventory in order to prevent a double count with the np_oilgas sector. VOC HAPs from California were incorporated into speciation similarly to VOC HAPs from MOVES elsewhere, e.g. model species BENZ is equal to HAP emissions for benzene as submitted by CARB. VOC and PM_{2.5} emissions were allocated to speciation profiles. Ratios of VOC (PM_{2.5}) by speciation profile to total VOC (PM_{2.5}) were calculated by county and SCC from the MOVES run in California, and then applied CARB-provided VOC (PM_{2.5}) in the inventory so that California nonroad emissions could be speciated consistently with the rest of the country.

State Submitted Data

The CDBs used to run MOVES-Nonroad to produce emissions for all states other than California were consistent with those used to develop the 2020 NEI. The following states submitted CDBs for the 2020 NEI: Arizona - Maricopa Co.; Connecticut; Georgia; Illinois; Indiana; Michigan; Minnesota; Ohio; Oregon; Texas; Utah; Washington; and Wisconsin..

Following the completion of the MOVES runs, railway maintenance emissions were removed from specific counties / census areas in Alaska because Alaska DEC specified that this type of activity does not happen in those areas. Specifically, emissions from SCCs 2285002015, 2285004015, and 2285006015 were removed from the following counties / census areas: 02013, 02016, 02050, 02060, 02063, 02066, 02070, 02100, 02105, 02110, 02130, 02150, 02158, 02164, 02180, 02185, 02188, 02195, 02198, 02220, 02240, 02275, and 02282. Alaska DEC also specified some counties / census areas in which logging and agricultural emissions do not happen, but the emissions for the specified SCCs were already zero in the specified areas.

For more information on the development of the nonroad emissions in the 2020 NEI see [Section 4 of the 2020 NEI TSD](#).

2.5 Fires (ptfire-rx, ptfire-wild, ptagfire)

Multiple types of fires are represented in the modeling platform. These include wild and prescribed fires that are grouped into the ptfire-wild and ptfire-rx sectors, respectively, and agricultural fires that comprise the ptagfire sector. All ptfire and ptagfire fires are in the United States. Fires outside of the United States are described in the ptfire_othna sector later in this document.

2.5.1 Wild and Prescribed Fires (ptfire-rx, ptfire-wild)

Wildfire and prescribed burning emissions are contained in the ptfire-wild and ptfire-rx sectors, respectively. The ptfire sector has emissions provided at geographic coordinates (point locations) and has daily emissions values. The ptfire-rx sector excludes agricultural burning and other open burning sources that are included in the ptagfire and nonpt sectors. Emissions are day-specific and include satellite-derived latitude/longitude of the fire's origin and other parameters associated with the emissions such as acres burned and fuel load, which allow estimation of plume rise.

The SCCs used for the ptfire-rx and ptfire-wild sources are shown in Table 2-22. The ptfire-rx and ptfire-wild inventories include separate SCCs for the flaming and smoldering combustion phases for wildfire and prescribed burns. Note that prescribed grassland fires or Flint Hills, Kansas has their own SCC (2801500171) in the inventory. Any wild grassland fires were assigned the standard wildfire SCCs shown in Table 2-22.

Table 2-22. SCCs included in the ptfire sector for the 2021 platform

SCC	Description
2801500171	Agricultural Field Burning - whole field set on fire;Fallow
2810001001	Forest Wildfires; Smoldering; Residual smoldering only (includes grassland wildfires)
2810001002	Forest Wildfires; Flaming (includes grassland wildfires)
2811015001	Prescribed Forest Burning; Smoldering; Residual smoldering only
2811015002	Prescribed Forest Burning; Flaming
2811020002	Prescribed Rangeland Burning

Fire Information Data

Inputs to SMARTFIRE2 for 2021 include:

- The National Oceanic and Atmospheric Administration's (NOAA's) Hazard Mapping System (HMS) fire location information
- National Incident Feature Services (NIFS) (formerly GeoMAC) wildland fire perimeter polygons
- The Incident Status Summary, also known as the "ICS-209", used for reporting specific information on fire incidents of significance.

- Hazardous fuel treatment reduction polygons for prescribed burns from the Forest Service Activity Tracking System (FACTS)
- Fire activity on federal lands from the United States Fish and Wildlife Service (USFWS) and other Department of Interior agencies

The Hazard Mapping System (HMS) was developed in 2001 by the National Oceanic and Atmospheric Administration's (NOAA) National Environmental Satellite and Data Information Service (NESDIS) as a tool to identify fires over North America in an operational environment. The system utilizes geostationary and polar orbiting environmental satellites. Automated fire detection algorithms are employed for each of the sensors. When possible, HMS data analysts apply quality control procedures for the automated fire detections by eliminating those that are deemed to be false and adding hotspots that the algorithms have not detected via a thorough examination of the satellite imagery.

The HMS product used for the 2021 inventory consisted of daily comma-delimited files containing fire detect information including latitude-longitude, satellite used, time detected, and other information. These detects were processed through Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation version 2 SmartFire2/BlueSky Pipeline (SF2/BSP).

National Incident Feature Services (NIFS) is an online wildfire mapping application designed for fire managers to access maps of current U.S. fire locations and perimeters. The wildfire perimeter data are based upon input data from incident intelligence sources from multiple agencies, GPS data, and infrared (IR) imagery from fixed wing and satellite platforms.

The Incident Status Summary, also known as the "ICS-209" is used for reporting specific information on significant fire incidents. The ICS-209 report is a critical interagency incident reporting tool giving daily 'snapshots' of the wildland fire management situation and individual incident information which include fire behavior, size, location, cost, and other information. Data from two tables in the ICS-209 database were merged and used for the 2021 ptfire inventory: the SIT209_HISTORY_INCIDENT_209_REPORTS table contained daily 209 data records for large fires, and the SIT209_HISTORY_INCIDENTS table contained summary data for additional smaller fires.

The US Forest Service (USFS) compiles a variety of fire information every year. Year 2021 data from the USFS Natural Resource Manager (NRM) Forest Activity Tracking System (FACTS) were acquired and used for emissions inventory development. This database includes information about activities related to fire/fuels, silviculture, and invasive species. The FACTS database consists of shapefiles for prescribed burns that provide acres burned and start and ending time information.

The US Fish and Wildlife Service (USFWS) also compiles wildfire and prescribed burn activity on their federal lands every year. Year 2021 data were acquired from USFWS through direct communication with USFWS staff and were used for 2021 platform development. The USFWS fire information provided fire type, acres burned, latitude-longitude, and start and ending times. The Department of Interior also provided National Fire Plan Operations and Reporting System (NFPORS) activity data that covers all other DOI agencies.

Fire Emissions Estimation Methodology

The national and S/L/T data mentioned earlier were used to estimate daily wildfire and prescribed burn emissions from flaming combustion and smoldering combustion phases for the 2021 inventory. Flaming combustion is more complete combustion than smoldering and is more prevalent with fuels that have a high surface-to-volume ratio, a low bulk density, and low moisture content. Smoldering combustion occurs without a flame, is a less complete burn, and produces some pollutants, such as PM_{2.5}, VOCs, and CO, at higher rates than flaming combustion. Smoldering combustion is more prevalent with fuels that have low surface-to-volume ratios, high bulk density, and high moisture content. Models sometimes differentiate between smoldering emissions that are lofted with a smoke plume and those that remain near the ground (residual emissions), but for the purposes of the inventory the residual smoldering emissions were allocated to the smoldering SCCs listed in Table 2-22. The lofted smoldering emissions were assigned to the flaming emissions SCCs in Table 2-22.

Figure 2-8 is a schematic of the data processing stream for the inventory of wildfire and prescribed burn sources. The ptfire-rx and ptfire-wild inventory sources were estimated using Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation version 2 (SMARTFIRE2) and Blue Sky Pipeline. SMARTFIRE2 is an algorithm and database system that is within a geographic information system (GIS). SMARTFIRE2 combines multiple sources of fire information and reconciles them into a unified GIS database. It reconciles fire data from space-borne sensors and ground-based reports, thus drawing on the strengths of both data types while avoiding double-counting of fire events. At its core, SMARTFIRE2 is an association engine that links reports covering the same fire in any number of multiple databases. In this process, all input information is preserved, and no attempt is made to reconcile conflicting or potentially contradictory information (for example, the existence of a fire in one database but not another).

For the 2021 platform, the national and S/L/T fire information was input into SMARTFIRE2 and then merged and associated based on user-defined weights for each fire information dataset. The output from SMARTFIRE2 was daily acres burned by fire type, and latitude-longitude coordinates for each fire. The fire type assignments were made using the fire information datasets. If the only information for a fire was a satellite detect for fire activity, then the flow described in Figure 2-8 was used to make fire type assignment by state and by month in conjunction with the default fire type assignments shown in Figure 2-9.

Figure 2-8. Processing flow for fire emission estimates in the 2021 inventory

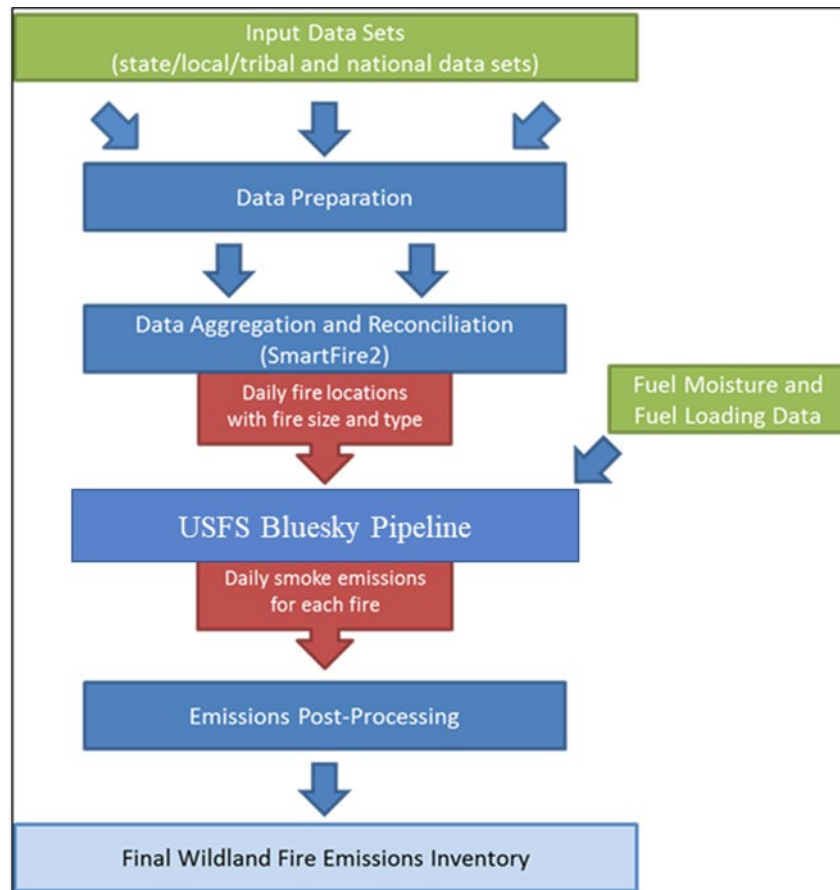
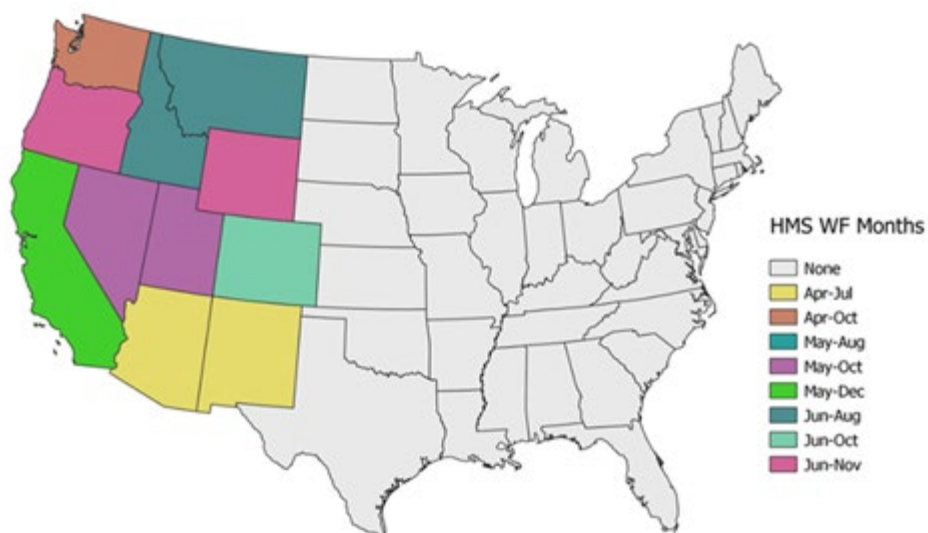


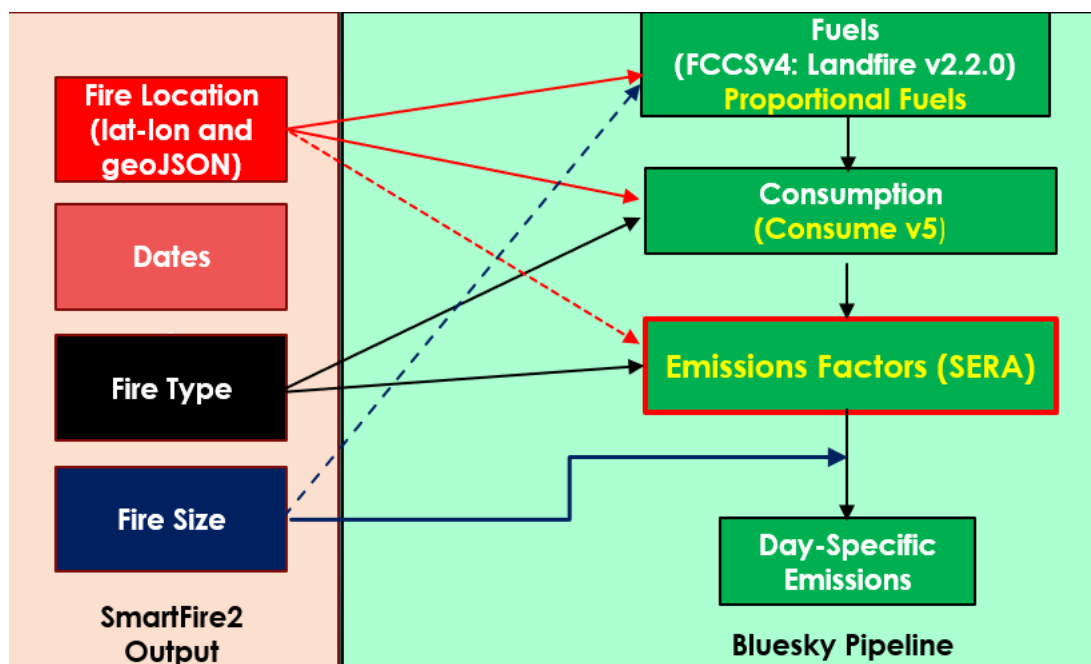
Figure 2-9. Default fire type assignment by state and month where data are only from satellites

2020 NEI HMS Default Wildfire Type Months



The second system used to estimate emissions is the BlueSky Modeling Pipeline (BSP). The framework supports the calculation of fuel loading and consumption, and emissions using various models depending on the available inputs as well as the desired results. The contiguous United States and Alaska, where Fuel Characteristic Classification System (FCCS) fuel loading data are available, were processed using the modeling chain described in Figure 2-10. The Smoke Emissions Reference Application (SERA) in the BSP generates all the CAP emission factors for wildland fires used in the 2021 study. SERA factors can vary by phase, fire type, region, fuel type and more pollutants. SERA emission factors are available here: <https://depts.washington.edu/nwfire/sera/index.php>. SERA consists of existing peer-reviewed emission factors (EFs) of 276 known air pollutants. The SERA database enables the analysis and summaries of existing EFs, and creation of average EFs to be used in decision support tools for smoke management, including BSP. HAP emission factors were obtained from Urbanski's (2014) work and applied by region and by fire type.

Figure 2-10. Blue Sky Modeling Pipeline



The FCCSv3 cross-reference was implemented along with the LANDFIREv1 (at 200 meter resolution) to provide better fuel bed information for the BlueSky Pipeline (BSP). The LANDFIREv2 was aggregated from the native resolution and projection to 200 meter using a nearest-neighbor methodology. Aggregation and reprojection were required for the proper function on BSP.

The final products from this process were annual and daily FF10-formatted emissions inventories. These SMOKE-ready inventory files contain both CAPs and HAPs. The BAFM HAP emissions from the inventory were used directly in modeling and were not overwritten with VOC speciation profiles (i.e., an “integrate HAP” use case).

2.5.2 Point source Agriculture Fires (ptagfire)

In the NEI, agricultural fires are stored as county-annual emissions and are part of the nonpoint data category. For this study agricultural fires are modeled as day specific fires derived from satellite data for the year 2021 in a similar way to the emissions in ptfire.

Daily year-specific agricultural burning emissions are derived from HMS fire activity data, which contains the date and location of remote-sensed anomalies. The activity is filtered using the 2021 USDA cropland data layer (CDL). Satellite fire detects over agricultural lands are assumed to be agricultural burns and assigned a crop type. Detects that are not over agricultural lands are output to a separate file for use in the ptfire sector. Each detect is assigned an average size of between 40 and 120 acres varying by state. Grassland/pasture fires were moved to the ptfire sectors for this 2021 modeling platform. Depending on their origin, grassland fires are in both ptfire-rx and ptfire-wild sectors because both fire types do involve grassy fuels.

The point source agricultural fire (ptagfire) inventory sector contains daily agricultural burning emissions. Daily fire activity was derived from the NOAA Hazard Mapping System (HMS) fire activity data. The agricultural fires sector includes SCCs starting with '28015'. The first three levels of descriptions for these SCCs are: 1) Fires - Agricultural Field Burning; Miscellaneous Area Sources; 2) Agriculture Production - Crops - as nonpoint; and 3) Agricultural Field Burning - whole field set on fire. The SCC 2801500000 does not specify the crop type or burn method, while the more specific SCCs specify field or orchard crops and, in some cases, the specific crop being grown. The SCCs for this sector listed are in Table 2-23.

Table 2-23. SCCs included in the ptagfire sector

SCC	Description
2801500000	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Unspecified crop type and Burn Method
2801500141	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Bean (red): Headfire Burning
2801500150	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Corn: Burning Techniques Not Important
2801500160	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Cotton: Burning Techniques Not Important
2801500220	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Rice: Burning Techniques Not Significant
2801500250	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Sugar Cane: Burning Techniques Not Significant
2801500262	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; Field Crop is Wheat: Backfire Burning
2801500264	Miscellaneous Area Sources; Agriculture Production - Crops - as nonpoint; Agricultural Field Burning - whole field set on fire; DoubleCrop Winter Wheat and Soybeans
2811020002	Miscellaneous Area Sources;Other Combustion - as Event;Prescribed Rangeland Burning;Flaming

Another feature of the ptagfire database is that the satellite detections for 2021 were filtered out to exclude areas covered by snow during the winter months. To do this, the daily snow cover fraction per grid cell was extracted from a 2021 meteorological Weather Research Forecast (WRF) model simulation. The locations of fire detections were then compared with this daily snow cover file. For any day in which a grid cell had snow cover, the fire detections in that grid cell on that day were excluded from the inventory. Due to the inconsistent reporting of fire detections from the Visible Infrared Imaging Radiometer Suite (VIIRS) platform, any fire detections in the HMS dataset that were flagged as VIIRS or Suomi National Polar-orbiting Partnership satellite were excluded. In addition, certain crop types (corn and soybeans) have been excluded from these specific midwestern states: Iowa, Kansas, Indiana, Illinois, Michigan, Missouri, Minnesota, Wisconsin, and Ohio. The reason for these crop types being excluded is because states have indicated that these crop types are not burned.

Heat flux for plume rise was calculated using the size and assumed fuel loading of each daily agricultural fire. This information is needed for a plume rise calculation within a chemical transport modeling system.

The daily agricultural emissions were converted from a tabular format into the SMOKE-ready daily point flat file format. The daily emissions were also aggregated into annual values by location and converted into the annual point flat file format.

For this modeling platform, a SMOKE update allows the use of HAP integration for speciation for PTDAY inventories. The 2021 agricultural fire inventories include emissions for HAPs, so HAP integration was used for this study.

2.6 Biogenic Sources (beis)

Biogenic emissions were computed based on the 2021 meteorology data used for the 2021 platform and were developed using the Biogenic Emission Inventory System version 4 (BEIS4) within CMAQ. BEIS4 creates gridded, hourly, model-species emissions from vegetation and soils. It estimates CO, VOC (most notably isoprene, terpene, and sesquiterpene), and NO emissions for the contiguous U.S. and for portions of Mexico and Canada. In the BEIS4 two-layer canopy model, the layer structure varies with light intensity and solar zenith angle (Pouliot and Bash, 2015). Both layers include estimates of sunlit and shaded leaf area based on solar zenith angle and light intensity, direct and diffuse solar radiation, and leaf temperature (Bash et al., 2015). BEIS4 computes the seasonality of emissions using the 1-meter soil temperature (SOIT2) instead of the BIOSEASON file, and canopy temperature and radiation environments are now modeled using the driving meteorological model's (WRF) representation of leaf-area index (LAI) rather than the estimated LAI values from BELD data alone. See [these CMAQ Release Notes](https://github.com/USEPA/CMAQ/wiki/CMAQ-Release-Notes) for technical information on BEIS4: <https://github.com/USEPA/CMAQ/wiki/CMAQ-Release-Notes:-Emissions-Updates:-BEIS-Biogenic-Emissions>. The variables output from the Meteorology-Chemistry Interface Processor (MCIP) that are used to convert WRF outputs to CMAQ inputs are shown in Table 2-24.

Table 2-24. Meteorological variables required by BEIS4

Variable	Description
LAI	leaf-area index
PRSFC	surface pressure
Q2	mixing ratio at 2 m
RC	convective precipitation per met TSTEP
RGRND	solar rad reaching surface
RN	nonconvective precipitation per met TSTEP
RSTOMI	inverse of bulk stomatal resistance
SLYTP	soil texture type by USDA category
SOIM1	volumetric soil moisture in top cm
SOIT1	soil temperature in top cm
TEMPG	skin temperature at ground
USTAR	cell averaged friction velocity
RADYNI	inverse of aerodynamic resistance
TEMP2	temperature at 2 m
WSAT_PX	soil saturation from (Pleim-Xiu Land Surface Model) PX-LSM

The Biogenic Emissions Landcover Database version 6 (BELD6) was used as the input gridded land use information in generating the biogenic emissions estimates. BELD version 5 (BELD5) was used to generate 2017 NEI estimates. There are now two different BELD6 datasets that are input into BEIS4. The gridded landuse and the other is the gridded dry leaf biomass (grams/m²) values for various vegetation types. The BELD6 includes the following datasets:

- High resolution tree species and biomass data from Wilson et al. 2013a, and Wilson et al. 2013b for which species names were changed from non-specific common names to scientific names
- Tree species biogenic volatile organic carbon (BVOC) emission factors for tree species were taken from the NCAR Enclosure database (Wiedinmyer, 2001)
 - <https://www.sciencedirect.com/science/article/pii/S1352231001004290>
- Agricultural land use from [US Department of Agriculture \(USDA\) crop data layer](#)
- Global Moderate Resolution Imaging Spectroradiometer (MODIS) 20 category data with enhanced lakes and Fraction of Photosynthetically Active Radiation (FPAR) for vegetation coverage from [National Center for Atmospheric Research \(NCAR\)](#)
- Canadian BELD land use, updates to Version 4 of the Biogenic Emissions Landuse Database (BELD4) for Canada and Impacts on Biogenic VOC Emissions (https://www.epa.gov/sites/default/files/2019-08/documents/800am_zhang_2_0.pdf).

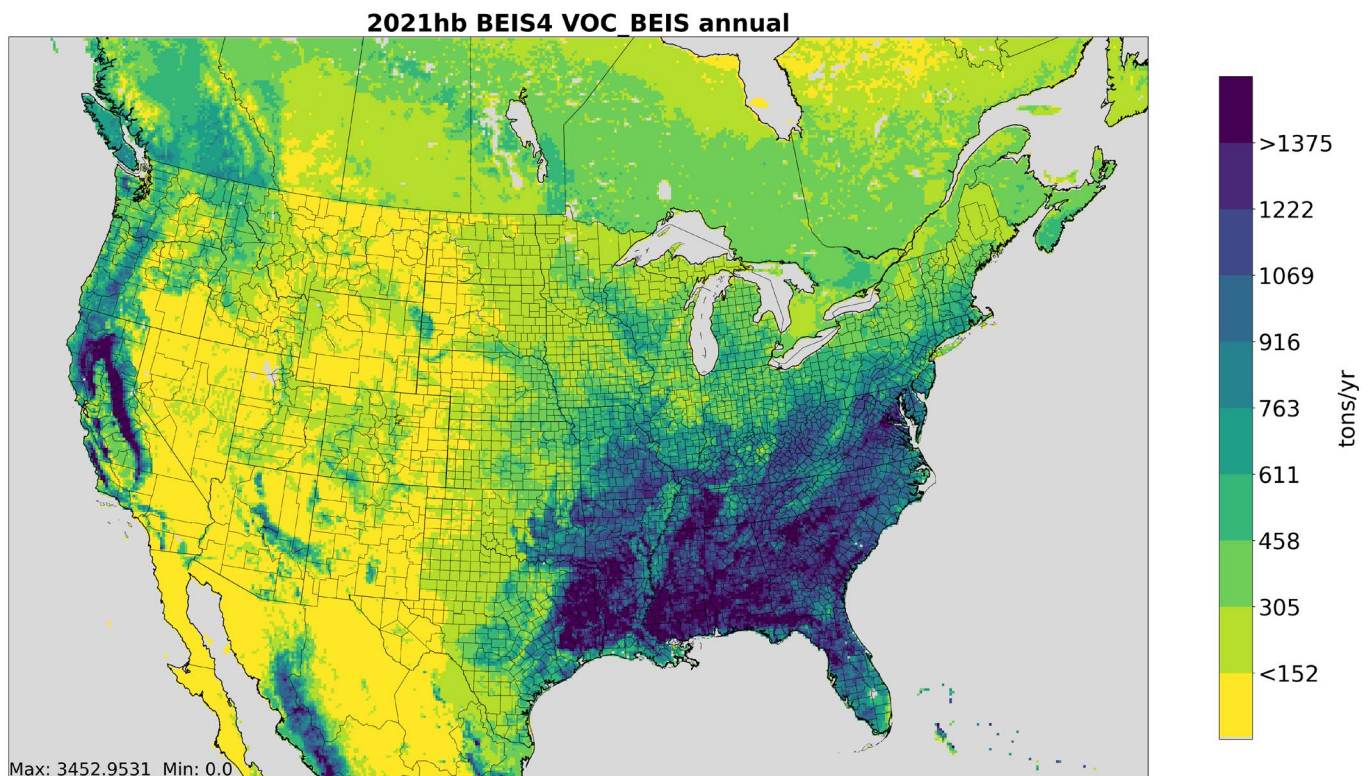
Bug fixes included in BEIS4 included the following:

- Solar radiation attenuation in the shaded portion of the canopy was using the direct beam photosynthetically active radiation (PAR) when the diffuse beam PAR attenuation coefficient should have been used.

- This update had little impact on the total emissions but did result in slightly higher emissions in the morning and evening transition periods for isoprene, methanol and Methylbutenol (MBO).
- The fraction of solar radiation in the sunlit and shaded canopy layers, SOLSUN and SOLSHADE respectively were estimated using a planar surface. These should have been estimated based on the PAR intercepted by a hemispheric surface rather than a plane.
 - This update can result in an earlier peak in leaf temperature, approximately up to an hour.
- The quantum yield for isoprene emissions (ALPHA) was updated to the mean value in [Niinemets et al. 2010a](#) and the integration coefficient (CL) was updated to yield 1 when PAR = 1000 following [Niinemets et al 2010b](#).
 - This updated resulted in a slight reduction in isoprene, methanol, and MBO emissions.

Biogenic emissions computed with BEIS were used to review and prepare summaries, but were left out of the CMAQ-ready merged emissions in favor of inline biogenics produced during the CMAQ model run itself using the same algorithm described above but with finer time steps within the air quality model. Figure 2-11 provides an annual estimate of the biogenic VOC emissions in year 2021 from BEIS4.

Figure 2-11. Annual biogenic VOC BEIS4 emissions for the 12US1 domain



2.7 Sources Outside of the United States

The emissions from Canada and Mexico are included as part of the emissions modeling sectors: canmex_point, canmex_area, canada_afdust, canada_ptdust, canada_onroad, mexico_onroad,

canmex_ag, and canada_og2D. The canmex_ag sector is processed as a separate sector for reporting and tracking purposes, and unlike in other recent emissions platforms, the Canada ag sources are area sources in this platform rather than pre-gridded point sources. As in prior platforms, Fugitive dust emissions in Canada are represented as both area sources (canada_afdust sector, formerly “othafdust”) and point sources (canada_ptdust sector, formerly “othptdust”). Due to the large number of individual points, low-level oil and gas emissions in Canada are processed separately from the canmex_point sector to reduce the number of individual points to track within CMAQ, and also to reduce the size of the model-ready emissions files.

Emissions in these sectors were taken from 2020 and 2023 inventories provided by Environment and Climate Change Canada (ECCC), interpolated to 2021. ECCC provided the following inventories; the sectors in which they were incorporated are listed and the inventories are described in more detail below:

- Agricultural livestock and fertilizer, area source format (canmex_ag sector)
- Surface-level oil and gas emissions in Canada (canada_og2D sector)
- Agricultural fugitive dust, point source format (canada_ptdust sector)
- Other area source dust (canada_afdust sector)
- Onroad (canada_onroad sector)
- Nonroad and rail (canmex_area sector)
- Airports (canmex_point sector)
- Other area sources (canmex_area sector)
- Other point sources (canmex_point sector)

The 2021 CMV data included coastal waters of Canada and Mexico with emissions derived from AIS data. These NEI emissions were used for all areas of Canada and Mexico and are included in the cmv_c1c2 and cmv_c3 sectors. Both the C1C2 and C3 emissions were developed in a point source format with point locations at the center of the 12km grid cells.

Other than the CB6 species of NBAFM present in the speciated point source data, there are no explicit HAP emissions in these Canadian inventories. In addition to emissions inventories, the ECCC 2020 dataset also included shapefiles for creating spatial surrogates. These surrogates were used for this study.

While emissions in the 2020 platform were adjusted at the monthly level to reflect COVID pandemic effects, no such adjustments were made for 2021 modeling.

2.7.1 Point Sources in Canada and Mexico (canmex_point)

Canadian point source inventories provided by ECCC include emissions for airports and other point sources. The Canadian point source inventory is pre-speciated for the CB6 chemical mechanism.

Point sources in Mexico were compiled in two parts. New emissions inventories representing 2018 developed through a collaboration between EPA and SEMARNAT were used for the six Mexico border states: Baja California, Sonora, Chihuahua, Coahuila, Nuevo Leon, and Tamaulipas. Mexico inventories for other states were based on inventories projected from the Inventario Nacional de Emisiones de

Mexico, 2016 (Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT)), projected to 2019 as part of the 2019 emissions modeling platform. For the emissions carried forward from the 2019 platform, the point source emissions were converted to English units and into the FF10 format that could be read by SMOKE, missing stack parameters were gapfilled using SCC-based defaults, latitude and longitude coordinates were verified and adjusted if they were not consistent with the reported municipality. Only CAPs are covered in the Mexico point source inventory.

2.7.2 Fugitive Dust Sources in Canada (canada_afdust, canada_ptdust)

Fugitive dust sources of particulate matter emissions excluding land tilling from agricultural activities, were provided by Environment and Climate Change Canada (ECCC) for 2020 and 2023, and were interpolated to 2021 for this study. This inventory no longer contains agricultural dust. Different source categories were provided as gridded point sources and area (nonpoint) source inventories. Gridded point source emissions resulting from land tilling due to agricultural activities were provided as part of the ECCC emission inventory. The provided wind erosion emissions were removed. A transport fraction adjustment that reduces dust emissions based on land cover types was applied to both point and nonpoint dust emissions, along with a meteorology-based (precipitation and snow/ice cover) zero-out of emissions when the ground is snow covered or wet.

2.7.3 Agricultural Sources in Canada and Mexico (canmex_ag)

Agricultural emissions from Canada and Mexico, excluding fugitive dust, are included in the canmex_ag sector. Canadian agricultural emissions were provided by Environment and Climate Change Canada (ECCC) as part of their 2020 and 2023 emission inventories (interpolated to 2021). Unlike in recent platforms, Canadian agricultural were not represented as point sources, instead they were represented as area sources and gridded using spatial surrogates. In Mexico, agricultural sources were based on new emissions inventories representing 2018 for the six Mexico border states (Baja California, Sonora, Chihuahua, Coahuila, Nuevo Leon, and Tamaulipas), and emissions from the 2019 emissions platform (SEMARNAT-provided 2016, projected to 2019) were carried forward for all other states.

2.7.4 Surface-level Oil and Gas Sources in Canada (canada_og2D)

Canadian point source inventories provided by ECCC included oil and gas emissions, and were interpolated from 2020 and 2023 to 2021. A very large number of these oil and gas point sources are surface level emissions, appropriate to be modeled in layer 1. Reducing the size of the canmex_point sector improves air quality model run time because plume rise calculations are needed for fewer sources, so these surface level oil and gas sources were placed into the canada_og2D sector for layer 1 modeling.

2.7.5 Nonpoint and Nonroad Sources in Canada and Mexico (canmex_area)

ECCC provided year 2020 and 2023 at the Canada province, and in some cases sub-province, resolution emissions from for nonpoint and nonroad sources (canmex_area). 2021 was interpolated from the 2020 and 2023 emissions. The nonroad sources were monthly while the nonpoint and rail emissions were annual.

In Mexico, nonroad and nonpoint sources were based on new emissions inventories representing 2018 for the six Mexico border states (Baja California, Sonora, Chihuahua, Coahuila, Nuevo Leon, and

Tamaulipas), and emissions from the 2019 emissions platform (SEMARNAT-provided 2016, projected to 2019) were carried forward for all other states.

2.7.6 Onroad Sources in Canada and Mexico (canada_onroad, mexico_onroad)

The onroad emissions for Canada and Mexico are in the canada_onroad and mexico_onroad sectors, respectively. ECCC provided year 2020 and 2023 at the Canada province, and in some cases sub-province resolution. 2021 was interpolated from the 2020 and 2023 emissions.

For Mexico onroad emissions, a version of the MOVES model for Mexico was run for 2020 and 2023. 2021 was then interpolated. This provided the same VOC HAPs and speciated VOCs as for the U.S. MOVES model (ERG, 2016a). This includes NBAFM plus several other VOC HAPs such as toluene, xylene, ethylbenzene and others. Except for VOC HAPs that are part of the speciation, no other HAPs are included in the Mexico onroad inventory (such as particulate HAPs nor diesel particulate matter).

2.7.7 Fires in Canada and Mexico (ptfire_othna)

Annual 2021 wildland fire emissions for Mexico, Canada, Central America, and Caribbean nations are included in the ptfire_othna sector. [Annual](#) Canadian fire activity was developed by processing the Canadian Wildland Fire Information System's National Burned Area Composite (NBAC) and NOAA's Hazard Mapping System (HMS) through SMARTFIRE 2¹¹. Emissions were estimated from the wildland fire activity using BlueSky pipeline with Canadian Fire Behavior Prediction (FBP) fuel beds mapped to Fuel Characteristic Classification System (FCCS) fuel beds. Fires in Mexico, Central America, and the Caribbean, were developed from the Fire Inventory from NCAR (FINN) v2.5 daily fire emissions for 2021 (Wiedenmyer, 2023). For FINN fires, listed vegetation type codes of 1 and 9 are defined as agricultural burning, all other fire detections and assumed to be wildfires. All wildland fires that are not defined as agricultural are assumed to be wildfires rather than prescribed. FINN fire detects of less than 50 square meters (0.012 acres) are removed from the inventory. The locations of FINN fires are geocoded from latitude and longitude to FIPS code.

2.7.8 Ocean Chlorine, Ocean Sea Salt, and Volcanic Mercury

The ocean chlorine gas emission estimates are based on the build-up of molecular chlorine (Cl₂) concentrations in oceanic air masses (Bullock and Brehme, 2002). Data at 36 km and 12 km resolution were available and were not modified other than the model-species name "CHLORINE" was changed to "CL2" to support CMAQ modeling.

For mercury, the volcanic mercury emissions that were used in the recent modeling platforms were not included in this study. The emissions were originally developed for a 2002 multipollutant modeling platform with coordination and data from Christian Seigneur and Jerry Lin for 2001 (Seigneur et. al, 2004 and Seigneur et. al, 2001).). The volcanic emissions from the most recent eruption were not included in the because they have diminished by the year 2019. Thus no volcanic emissions were included.

Because of mercury bidirectional flux within the latest version of CMAQ, no other natural mercury emissions are included in the emissions merge step.

¹¹ See <https://www.cmascenter.org/conference/2023/slides/2023-10-18-1350-2021-Canada-WF-Updates-CMAS.pptx>

3 Emissions Modeling

The CMAQ and CAMx air quality models require hourly emissions of specific gas and particle species for the horizontal and vertical grid cells contained within the modeled region (i.e., modeling domain). To provide emissions in the form and format required by the model, it is necessary to “pre-process” the “raw” emissions (i.e., emissions input to SMOKE) for the sectors described above in Section 2. In brief, the process of emissions modeling transforms the emissions inventories from their original temporal resolution, pollutant resolution, and spatial resolution into the hourly, speciated, gridded and vertical resolution required by the air quality model. Emissions modeling includes temporal allocation, spatial allocation, and pollutant speciation. Emissions modeling sometimes includes the vertical allocation (i.e., plume rise) of point sources, but many air quality models also perform this task because it greatly reduces the size of the input emissions files if the vertical layers of the sources are not included.

As seen in Section 2, the temporal resolutions of the emissions inventories input to SMOKE vary across sectors and may be hourly, daily, monthly, or annual total emissions. The spatial resolution may be individual point sources; totals by county (U.S.), province (Canada), or municipio (Mexico); or gridded emissions. This section provides some basic information about the tools and data files used for emissions modeling as part of the modeling platform.

3.1 Emissions Modeling Overview

SMOKE version 5.0 was used to process the raw emissions inventories into emissions inputs for each modeling sector into a format compatible with CMAQ. SMOKE executables and source code are available from the Community Multiscale Analysis System (CMAS) Center at <http://www.cmascenter.org>. Additional information about SMOKE is available from <http://www.smoke-model.org>. For sectors that have plume rise, the in-line plume rise capability allows for the use of emissions files that are much smaller than full three-dimensional gridded emissions files. For quality assurance of the emissions modeling steps, emissions totals by specie for the entire model domain are output as reports that are then compared to reports generated by SMOKE on the input inventories to ensure that mass is not lost or gained during the emissions modeling process.

When preparing emissions for the air quality model, emissions for each sector are processed separately through SMOKE, and then the final merge program (Mrggrid) is run to combine the model-ready, sector-specific 2-D gridded emissions across sectors. The SMOKE settings in the run scripts and the data in the SMOKE ancillary files control the approaches used by the individual SMOKE programs for each sector. Table 3-1 summarizes the major processing steps of each platform sector with the columns as follows.

The “Spatial” column shows the spatial approach used: “point” indicates that SMOKE maps the source from a point location (i.e., latitude and longitude) to a grid cell; “surrogates” indicates that some or all of the sources use spatial surrogates to allocate county emissions to grid cells; and “area-to-point” indicates that some of the sources use the SMOKE area-to-point feature to grid the emissions (further described in Section 3.4.2). The “Speciation” column indicates that all sectors use the SMOKE speciation step, though biogenics speciation is done within the Tmpbeis3 program and not as a separate SMOKE step. The “Inventory resolution” column shows the inventory temporal resolution from which SMOKE needs to calculate hourly emissions. Note that for some sectors (e.g., onroad, beis), there is no input

inventory; instead, activity data and emission factors are used in combination with meteorological data to compute hourly emissions.

Table 3-1. Key emissions modeling steps by sector.

Platform sector	Spatial	Speciation	Inventory resolution	Plume rise
afdust_adj	Surrogates	Yes	Annual	
Airports	Point	Yes	Annual	None
Beis	Pre-gridded land use	in BEIS4	computed hourly in CMAQ	
Fertilizer	EPIC	No	computed hourly in CMAQ	
livestock	Surrogates	Yes	Daily	
cmv_c1c2	Point	Yes	hourly	in-line
cmv_c3	Point	Yes	hourly	in-line
nonpt	Surrogates & area-to-point	Yes	Annual	
nonroad	Surrogates	Yes	monthly	
np_oilgas	Surrogates	Yes	Annual	
onroad	Surrogates	Yes	monthly activity, computed hourly	
onroad_ca_adj	Surrogates	Yes	monthly activity, computed hourly	
canada_onroad	Surrogates	Yes	monthly	
mexico_onroad	Surrogates	Yes	monthly	
canada_afdust	Surrogates	Yes	annual & monthly	
canmex_area	Surrogates	Yes	monthly	
canmex_point	Point	Yes	monthly	in-line
canada_ptdust	Point	Yes	annual	None
canada_og2D	Point	Yes	monthly	None
canmex_ag	Surrogates	Yes	annual	
ptagfire	Point	Yes	daily	in-line
pt_oilgas	Point	Yes	annual	in-line
ptegu	Point	Yes	daily & hourly	in-line
ptfire-rx	Point	Yes	daily	in-line
ptfire-wild	Point	Yes	daily	in-line
ptfire_othna	Point	Yes	daily	in-line
ptnonipm	Point	Yes	annual	in-line
rail	Surrogates	Yes	annual	
rwc	Surrogates	Yes	annual	
np_solvents	Surrogates	Yes	annual	

The “plume rise” column in indicates the sectors for which the “in-line” approach is used. These sectors are the only ones with emissions in aloft layers based on plume rise. The term “in-line” means that the plume rise calculations are done inside of the air quality model instead of being computed by SMOKE. In all of the “in-line” sectors, all sources are output by SMOKE into point source files which are subject to plume rise calculations in the air quality model. In other words, no emissions are output to layer 1 gridded emissions files from those sectors as has been done in past platforms. The air quality model computes the plume rise using stack parameters, the Briggs algorithm, and the hourly emissions in the SMOKE output files for each emissions sector. The height of the plume rise determines the model layers into which the emissions are placed. The plume top and bottom are computed, along with the plumes’ distributions into the vertical layers that the plumes intersect. The pressure difference across each layer divided by the pressure difference across the entire plume is used as a weighting factor to assign the emissions to layers. This approach gives plume fractions by layer and source. Day-specific point fire emissions are treated differently in CMAQ. After plume rise is applied, there are emissions in every layer from the ground up to the top of the plume.

Note that SMOKE has the option of grouping sources so that they are treated as a single stack when computing plume rise. For the modeling cases discussed in this document, no grouping was performed because grouping combined with “in-line” processing will not give identical results as “offline” processing (i.e., when SMOKE creates 3-dimensional files). This occurs when stacks with different stack parameters or latitude and longitudes are grouped, thereby changing the parameters of one or more sources. The most straightforward way to get the same results between in-line and offline is to avoid the use of stack grouping.

Biogenic emissions can be modeled two different ways in the CMAQ model. The BEIS model in SMOKE can produce gridded biogenic emissions that are then included in the gridded CMAQ-ready emissions inputs, or alternatively, CMAQ can be configured to create “in-line” biogenic emissions within CMAQ itself. For this study, the in-line biogenic emissions option was used, and so biogenic emissions from BEIS were not included in the gridded CMAQ-ready emissions.

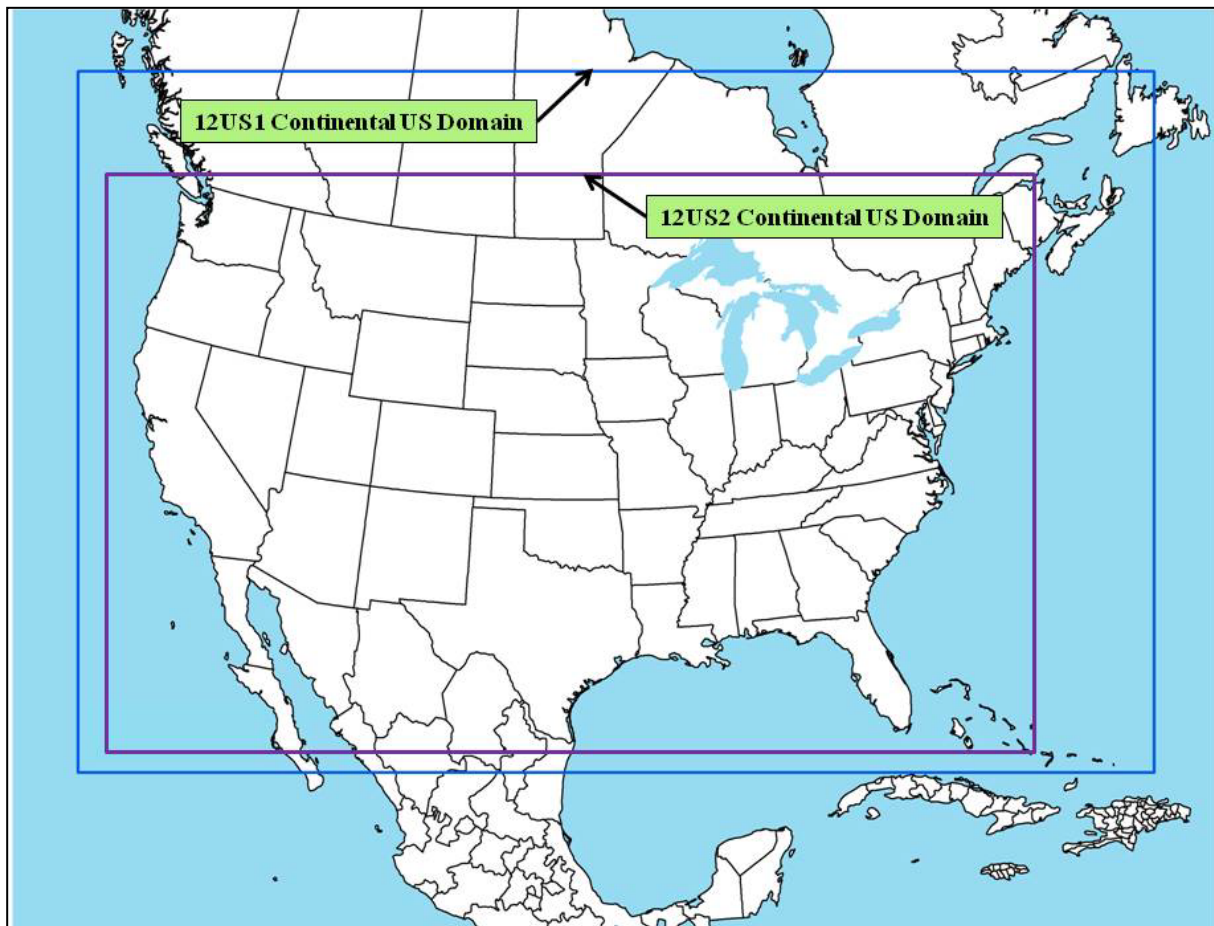
For this study, SMOKE was run for the larger 12-km CONTinental United States “CONUS” modeling domain (12US1) shown in Figure 3-1, but the air quality model was run on the smaller 12-km domain (12US2). More specifically, SMOKE was run on the 12US1 domain and emissions were extracted from 12US1 data files to create 12US2 emissions. The grids used a Lambert-Conformal projection, with Alpha = 33, Beta = 45 and Gamma = -97, with a center of X = -97 and Y = 40. In addition, SMOKE was run for grids over Alaska, Hawaii, and Puerto Rico plus the Virgin Islands. Later sections provide details on the spatial surrogates and area-to-point data used to accomplish spatial allocation with SMOKE. Table 3-2 describes the grids. WRF, SMOKE, and CMAQ all presume the Earth is a sphere with a radius of 6370000 m.

Table 3-2. Descriptions of the platform grids

Common Name	Grid Cell Size	Description	Grid name	Parameters listed in SMOKE grid description (GRIDDESC) file: projection name, xorig, yorig, xcell, ycell, ncols, nrows, nthik
Continental 12km grid	12 km	Entire conterminous US plus some of Mexico/Canada	12US1_459X299	'LAM_40N97W', -2556000, -1728000, 12.D3, 12.D3, 459, 299, 1
US 12 km or "smaller" CONUS-12	12 km	Smaller 12km CONUS plus some of Mexico/Canada	12US2	'LAM_40N97W', -2412000, -1620000, 12.D3, 12.D3, 396, 246, 1
Alaska 9km	9 km	Small 9 km Alaska with parts of Canada	9AK1	LAM_36N_155W', -1107000, -1134000, 9000, 9000, 312, 252, 1
Hawaii 3km	3 km	Small 3 km Hawaii	3HI1	LAM_21N_157W', -391500, -346500, 3000, 3000, 225, 201, 1
Puerto Rico & Virgin Islands 3km	3 km	Small 3 km covering Puerto Rico and the Virgin Islands	3PR1	LAM_18N_66W', -274500, -202500, 3000, 3000, 150, 150, 1

Figure 3-1. Air quality modeling domains

a) 12US1 and 12US2



3.2 Chemical Speciation

Chemical speciation involves the process of translating emissions from the inventory into the chemical mechanism-specific “model species” needed by an air quality model. Using the CB6R5_AE7 chemical mechanism as an example, these model species either represent explicit chemical compounds (e.g., acetone, benzene, ethanol) or groups of species (i.e., “lumped species;” e.g., PAR, OLE, KET). Table 3-3 lists the model species generated by SMOKE for this mechanism. Table 3-4 and Table 3-5 list additional model species that are generated when performing toxics modeling, and Table 3-6 lists the mapping between individual polycyclic aromatic hydrocarbons (PAHs) to the PAH groups used in toxics modeling.

Table 3-3. Emission model species produced for CB6R5_AE7 for CMAQ

Inventory Pollutant	Model Species	Model species description
Cl ₂	CL2	Atomic gas-phase chlorine
HCl	HCL	Hydrogen Chloride (hydrochloric acid) gas
CO	CO	Carbon monoxide
NO _x	NO	Nitrogen oxide
NO _x	NO2	Nitrogen dioxide
NO _x	HONO	Nitrous acid
SO ₂	SO2	Sulfur dioxide
SO ₂	SULF	Sulfuric acid vapor
NH ₃	NH3	Ammonia
NH ₃	NH3_FERT	Ammonia from fertilizer
VOC	AACD	Acetic acid
VOC	ACET	Acetone
VOC	ALD2	Acetaldehyde
VOC	ALDX	Propionaldehyde and higher aldehydes
VOC	APIN	Alpha pinene
VOC	BENZ	Benzene
VOC	CAT1	Methyl-catechols
VOC	CH4	Methane
VOC	CRES	Cresols
VOC	CRON	Nitro-cresols
VOC	ETH	Ethene
VOC	ETHA	Ethane
VOC	ETHY	Ethyne
VOC	ETOH	Ethanol
VOC	FACD	Formic acid
VOC	FORM	Formaldehyde
VOC	GLY	Glyoxal
VOC	GLYD	Glycolaldehyde
VOC	IOLE	Internal olefin carbon bond (R-C=C-R)
VOC	ISOP	Isoprene
VOC	ISPD	Isoprene Product
VOC	IVOC	Intermediate volatility organic compounds
VOC	KET	Ketone Groups
VOC	MEOH	Methanol

Inventory Pollutant	Model Species	Model species description
VOC	MGLY	Methylglyoxal
VOC	NAPH	Naphthalene
VOC	NVOL	Non-volatile compounds
VOC	OLE	Terminal olefin carbon bond (R-C=C)
VOC	PACD	Peroxyacetic and higher peroxy-carboxylic acids
VOC	PAR	Paraffin carbon bond
VOC	PRPA	Propane
VOC	SESQ	Sesquiterpenes (from biogenics only)
VOC	SOAALK	Secondary Organic Aerosol (SOA) tracer
VOC	TERP	Terpenes (from biogenics only)
VOC	TOL	Toluene and other monoalkyl aromatics
VOC	UNR	Unreactive
VOC	XYLMN	Xylene and other polyalkyl aromatics, minus naphthalene
Naphthalene	NAPH	Naphthalene from inventory
Benzene	BENZ	Benzene from the inventory
Acetaldehyde	ALD2	Acetaldehyde from inventory
Formaldehyde	FORM	Formaldehyde from inventory
Methanol	MEOH	Methanol from inventory
PM ₁₀	PMC	Coarse PM > 2.5 microns and ≤ 10 microns
PM _{2.5}	PEC	Particulate elemental carbon ≤ 2.5 microns
PM _{2.5}	PNO3	Particulate nitrate ≤ 2.5 microns
PM _{2.5}	POC	Particulate organic carbon (carbon only) ≤ 2.5 microns
PM _{2.5}	PSO4	Particulate Sulfate ≤ 2.5 microns
PM _{2.5}	PAL	Aluminum
PM _{2.5}	PCA	Calcium
PM _{2.5}	PCL	Chloride
PM _{2.5}	PFE	Iron
PM _{2.5}	PK	Potassium
PM _{2.5}	PH2O	Water
PM _{2.5}	PMG	Magnesium
PM _{2.5}	PMN	Manganese
PM _{2.5}	PMOTHR	PM _{2.5} not in other AE6 species
PM _{2.5}	PNA	Sodium
PM _{2.5}	PNCOM	Non-carbon organic matter
PM _{2.5}	PNH4	Ammonium
PM _{2.5}	PSI	Silica
PM _{2.5}	PTI	Titanium

Table 3-4. Additional HAP gaseous model species generated for toxics modeling

Inventory Pollutant	Model Species
Acetaldehyde	ALD2_PRIMARY
Formaldehyde	FORM_PRIMARY
Acetonitrile	ACETONITRILE
Acrolein	ACROLEIN
Acrylic acid	ACRYLICACID
Acrylonitrile	ACRYLONITRILE
Benzo[a]Pyrene	BENZOAPYRNE
1,3-Butadiene	BUTADIENE13
Carbon tetrachloride	CARBONTET
Carbonyl Sulfide	CARBSULFIDE
Chloroform	CHCL3
Chloroprene	CHLOROPRENE
1,4-Dichlorobenzene(p)	DICHLOROBENZENE
1,3-Dichloropropene	DICHLOROPROPENE
Ethylbenzene	ETHYLBENZ
Ethylene dibromide (Dibromoethane)	BR2_C2_12
Ethylene dichloride (1,2-Dichloroethane)	CL2_C2_12
Ethylene oxide	ETOX
Hexamethylene-1,6-diisocyanate	HEXAMETH_DIIS
Hexane	HEXANE
Hydrazine	HYDRAZINE
Maleic Anyhydride	MAL_ANYHYDRIDE
Methyl Chloride	METHCHLORIDE
Methylene chloride (Dichloromethane)	CL2_ME
Specific PAHs assigned with URE =0	PAH_000E0
Specific PAHs assigned with URE =9.6E-06 (previously 1.76E-5)	PAH_176E5
Specific PAHs assigned with URE =4.8E-05 (previously 8.8E-5)	PAH_880E5
Specific PAHs assigned with URE =9.6E-05 (previously 1.76E-4)	PAH_176E4
Specific PAHs assigned with URE =9.6E-04 (previously 1.76E-3)	PAH_176E3
Specific PAHs assigned with URE =9.6E-03 (previously 1.76E-2)	PAH_176E2
Specific PAHs assigned with URE =0.01 (previously 1.01E-2)	PAH_101E2
Specific PAHs assigned with URE =1.14E-1	PAH_114E1
Specific PAHs assigned with URE =9.9E-04 (previously 1.92E-3)	PAH_192E3
Propylene dichloride (1,2-Dichloropropane)	PROPDICHLORIDE
Quinoline	QUINOLINE
Styrene	STYRENE
1,1,2,2-Tetrachloroethane	CL4_ETHANE1122
Tetrachloroethylene (Perchloroethylene)	CL4_ETHE
Toluene	TOLU
2,4-Toluene diisocyanate	TOL_DIIS
Trichloroethylene	CL3_ETHE
Triethylamine	TRIETHYLAMINE

Inventory Pollutant	Model Species
m-xylene, o-xylene, p-xylene, xylenes (mixed isomers)	XYLENES
Vinyl chloride	CL_ETHE

Table 3-5. Additional HAP particulate model species generated for toxics modeling

Inventory Pollutant	Model Species
Arsenic	ARSENIC_C, ARSENIC_F
Beryllium	BERYLLIUM_C, BERYLLIUM_F
Cadmium	CADMIUM_C, CADMIUM_F
Chromium VI, Chromic Acid (VI), Chromium Trioxide	CHROMHEX_C, CHROMHEX_F
Chromium III	CHROMTRI_C, CHROMTRI_F
Lead	LEAD_C, LEAD_F
Manganese	MANGANESE_C, MANGANESE_F
Mercury ¹	HGIIGAS, HGNRVA, PHGI
Nickel, Nickel Oxide, Nickel Refinery Dust	NICKEL_C, NICKEL_F
Diesel-PM10, Diesel-PM25	DIESEL_PMC, DIESEL_PMFINE, DIESEL_PMEC, DIESEL_PMOC, DIESEL_PMNO3, DIESEL_PMSO4

¹Mercury is multi-phase

Table 3-6. PAH/POM pollutant groups

PAH Group	NEI Pollutant Code	NEI Pollutant Description	URE 1/(µg/m ³)
PAH_000E0	120127	Anthracene	0
PAH_000E0	129000	Pyrene	0
PAH_000E0	85018	Phenanthrene	0
PAH_101E2	56495	3-Methylcholanthrene	0.01
PAH_114E1	57976	7,12-Dimethylbenz[a]Anthracene	0.114
PAH_176E2	189559	Dibenzo[a,i]Pyrene	9.6E-03
PAH_176E2	189640	Dibenzo[a,h]Pyrene	9.6E-03
PAH_176E2	191300	Dibenzo[a,l]Pyrene	9.6E-03
PAH_176E2	7496028	6-Nitrochrysene	9.6E-03
PAH_176E3	192654	Dibenzo[a,e]Pyrene	9.6E-04
PAH_176E3	194592	7H-Dibenzo[c,g]carbazole	9.6E-04
PAH_176E3	3697243	5-Methylchrysene	9.6E-04
PAH_176E3	41637905	Methylchrysene	9.6E-04
PAH_176E3	53703	Dibenzo[a,h]Anthracene	9.6E-04
PAH_176E4	193395	Indeno[1,2,3-c,d]Pyrene	9.6E-05
PAH_176E4	205823	Benzo[j]Fluoranthene	9.6E-05
PAH_176E4	205992	Benzo[b]Fluoranthene	9.6E-05
PAH_176E4	224420	Dibenzo[a,j]Acridine	9.6E-05
PAH_176E4	226368	Dibenz[a,h]acridine	9.6E-05
PAH_176E4	5522430	1-Nitropyrene	9.6E-05
PAH_176E4	56553	Benz[a]Anthracene	9.6E-05
PAH_176E5	207089	Benzo[k]Fluoranthene	9.6E-06
PAH_176E5	218019	Chrysene	9.6E-06

PAH Group	NEI Pollutant Code	NEI Pollutant Description	URE 1/(µg/m ³)
PAH_176E5	86748	Carbazole	9.6E-06
PAH_192E3	8007452	Coal Tar	9.9E-04
PAH_880E5	130498292	PAH, total	4.8E-05
PAH_880E5	191242	Benzo[g,h,i,j]Perylene	4.8E-05
PAH_880E5	192972	Benzo[e]Pyrene	4.8E-05
PAH_880E5	195197	Benzo(c)phenanthrene	4.8E-05
PAH_880E5	198550	Perylene	4.8E-05
PAH_880E5	203123	Benzo(g,h,i)Fluoranthene	4.8E-05
PAH_880E5	203338	Benzo(a)fluoranthene	4.8E-05
PAH_880E5	206440	Fluoranthene	4.8E-05
PAH_880E5	208968	Acenaphthylene	4.8E-05
PAH_880E5	2381217	1-Methylpyrene	4.8E-05
PAH_880E5	2422799	12-Methylbenz(a)Anthracene	4.8E-05
PAH_880E5	250	PAH/POM - Unspecified	4.8E-05
PAH_880E5	2531842	2-Methylphenanthrene	4.8E-05
PAH_880E5	26914181	Methylantracene	4.8E-05
PAH_880E5	284	Extractable Organic Matter (EOM)	4.8E-05
PAH_880E5	56832736	Benzofluoranthenes	4.8E-05
PAH_880E5	65357699	Methylbenzopyrene	4.8E-05
PAH_880E5	779022	9-Methyl Anthracene	4.8E-05
PAH_880E5	832699	1-Methylphenanthrene	4.8E-05
PAH_880E5	83329	Acenaphthene	4.8E-05
PAH_880E5	86737	Fluorene	4.8E-05
PAH_880E5	90120	1-Methylnaphthalene	4.8E-05
PAH_880E5	91576	2-Methylnaphthalene	4.8E-05
PAH_880E5	91587	2-Chloronaphthalene	4.8E-05
PAH_880E5	N590	Polycyclic aromatic compounds (includes PAH/POM)	4.8E-05

The TOG and PM_{2.5} profiles used to speciate emissions are part of the SPECIATE v5.2 database (<https://www.epa.gov/air-emissions-modeling/speciate>). The SPECIATE database is developed and maintained by the EPA's Office of Research and Development (ORD), Office of Transportation and Air Quality (OTAQ), and the Office of Air Quality Planning and Standards (OAQPS), in cooperation with Environment Canada (EPA, 2016). These profiles are processed using the EPA's S2S-Tool (<https://github.com/USEPA/S2S-Tool>) to generate the GSPRO and GSCNV files needed by SMOKE. As with previous platforms, some Canadian point source inventories are provided from Environment Canada as pre-specified emissions.

Speciation profiles (GSPRO files) and cross-references (GSREF files) for this platform are available in the SMOKE input files for the platform. Emissions of VOC and PM_{2.5} emissions by county, sector, and profile for all sectors other than onroad mobile can be found in the sector summaries. Total emissions for each model species by state and sector can be found in the state-sector totals workbook.

The following updates to profile assignments were made to this modeling platform and vary from prior years:

- For PM2.5:
 - All GSPRO files were generated by the S2S-Tool, dated 09-11-2023, and utilized SPECIATE v5.3.
 - Update of the CMV speciation cross-reference files to utilize the SCC updates for this sector and use the new CROC profiles introduced in SPECIATE v5.3.
 - Update onroad and nonroad mobile cross-reference files to utilize the CROC profiles introduced in SPECIATE v5.3.
- For VOC:
 - All GSPRO and GSCNV files were generated by the S2S-Tool, dated 09-11-2023, and utilized SPECIATE v5.3.
 - All oil and gas well completion and abandoned wells emissions were updated (or added in the case of abandoned wells) from 1101 and 8949, respectively, to 95404 and 95403, respectively. However, this update was not performed for basin-specific profiles that were output by the O&G Tool.
 - Update of the CMV speciation cross-reference files to utilize the SCC updates for this sector and use the new GROC profiles introduced in SPECIATE v5.3.
 - Update usage of 95120a to 95120c.
 - Update onroad and nonroad mobile cross-reference files to utilize the GROC profiles introduced in SPECIATE v5.3.

3.2.1 VOC speciation

The base emissions inventory for this modeling platform includes total VOC and individual HAP emissions. Often, individual HAPs are components of VOC (HAP-VOC), and these HAP-VOCs are included (“integrated”) in the speciation process. This HAP integration is performed in a way to ensure double counting of emitted mass does not occur and requires specific data processing by the S2S-Tool and user input in SMOKE.

To incorporate HAP emissions from the base inventory into the modeling platform, one of two methods are performed. (1) Integrate, HAP-use is a method where the mass of integrated HAP-VOCs is summed and subtracted from VOC, and the residual mass (NONHAPVOC) is speciated using a renormalized speciation profile that does not include the integrated HAP-VOCs (they are subtracted from the profile and then the profile is renormalized to 100%). (2) No-Integrate, HAP-use is a method where the mass of VOC is speciated using a speciation profile that does not include the integrated HAP-VOCs (they are subtracted from the profile and the profile is not renormalized to 100%). In this scenario, the HAP-VOC and VOC portions of the inventory are difficult to harmonize, and it is assumed that the proportions of HAPs from these sources are adequately captured in the speciation profile used to speciate the VOC emissions (which is why there is no renormalization). In addition, HAPs can be introduced into a modeling platform using speciation profiles. In this scenario, HAP-VOC emissions are “generated” through VOC speciation and are not incorporated from the base inventory. This method is called “Criteria” speciation. An illustration of these methods is shown in Figure 3-2 and the integration methods used for this platform for each sector are shown in Table 3-7.

Figure 3-2. Process of integrating HAPs and speciating VOC in a modeling platform

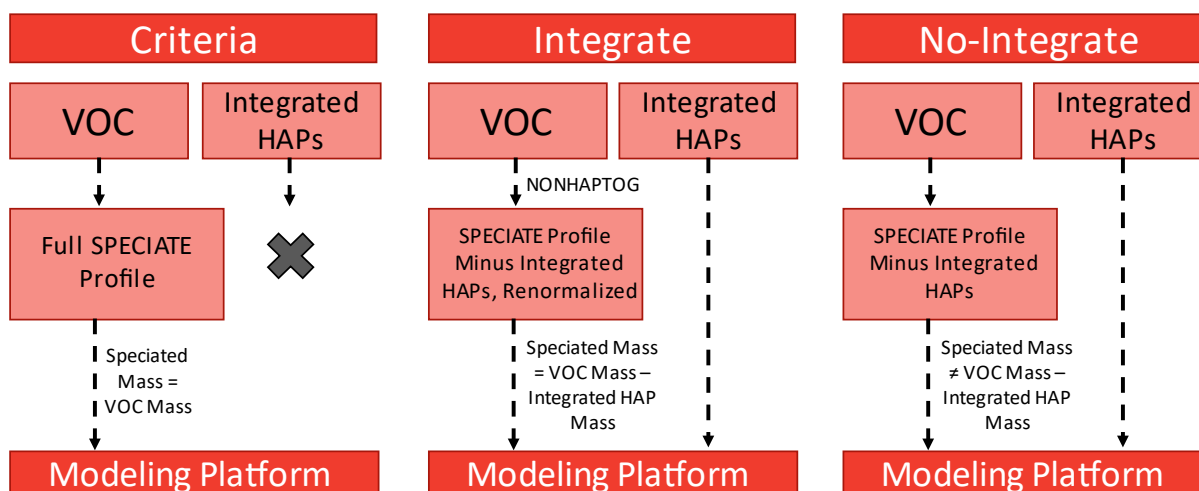


Table 3-7. Integration status for each platform sector

Platform Sector	Approach for Integrating NEI emissions of Naphthalene (N), Benzene (B), Acetaldehyde (A), Formaldehyde (F) and Methanol (M)
afdust	N/A – sector contains no VOC
airports	No integration, use NBAFM in inventory
beis	N/A – sector contains no inventory pollutant “VOC”; but rather specific VOC species
cmv_c1c2	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
cmv_c3	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
fertilizer	N/A – sector contains no VOC
livestock	Full integration (NBAFM)
nonpt	Partial integration (NBAFM)
nonroad	Full integration (internal to MOVES)
np_oilgas	Partial integration (NBAFM)
onroad	Full integration (internal to MOVES)
canada_onroad	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
mexico_onroad	Full integration (internal to MOVES-Mexico); however, MOVES-MEXICO speciation was older CB6, so post-SMOKE emissions were converted to CB6R3AE6
canada_afdust	N/A – sector contains no VOC
canmex_area	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
canmex_point	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
canada_ptdust	N/A – sector contains no VOC
canada_og2D	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
canmex_ag	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
pt_oilgas	No integration, use NBAFM in inventory
ptagfire	Full integration (NBAFM)
ptegu	No integration, use NBAFM in inventory
ptfire-rx	Full integration (NBAFM)
ptfire-wild	Partial integration (NBAFM)
ptfire_othna	No integration, no NBAFM in inventory, create NBAFM from VOC speciation
ptnonipm	No integration, use NBAFM in inventory

Platform Sector	Approach for Integrating NEI emissions of Naphthalene (N), Benzene (B), Acetaldehyde (A), Formaldehyde (F) and Methanol (M)
rail	Full integration (NBAFM)
rwf	Full integration (NBAFM)
np_solvents	Partial integration (NBAFM)

The HAPs integrated from the base inventory into the modeling platform are sector and chemical mechanism specific. In recent years, CB6R3_AE7 has been the primary chemical mechanism used at the EPA. Within that mechanism, naphthalene (NAPH), benzene (BENZ), acetaldehyde (ALD2), formaldehyde (FORM), and methanol (MEOH) are explicit HAP-VOCs, and these compounds are collectively referred to as NBAFM. Since NBAFM are explicitly modeled in CB6R3_AE7, these species have become the default collection of integrated HAP species at the EPA. MOVES, the EPA's mobile emissions model, features additional species that are explicitly modeled (e.g., ethanol). These species (Table 3-8) are also incorporated directly into modeling platforms. To incorporate these species, additional files from the S2S-Tool are required. For California, speciation of NONHAPTOG is performed on CARB's VOC submissions using the county-specific speciation profile assignments generated by MOVES in California.

Table 3-8. Integrated species from MOVES sources

MOVES ID	Pollutant Name
5	Methane (CH ₄)
20	Benzene
21	Ethanol
22	MTBE
24	1,3-Butadiene
25	Formaldehyde
26	Acetaldehyde
27	Acrolein
40	2,2,4-Trimethylpentane
41	Ethyl Benzene
42	Hexane
43	Propionaldehyde
44	Styrene
45	Toluene
46	Xylene
185	Naphthalene gas

Several sectors require VOC speciation to occur at the county-level and consistent speciation profiles cannot be applied across the nation. To accomplish this, the GSREF is setup to provide profiles that are "blended" at the county/SCC-level using proportions included in the input file. These variable VOC speciation methods are year-specific and applied in the oil and gas sector and for various mobile emissions sources. In both the np_oilgas and pt_oilgas sector, VOC speciation profiles are weighted to reflect region-specific application of controls, differences in gas composition, and variable sources of emissions (e.g., varying proportions of emissions from associated gas, condensate tanks, crude oil tanks,

dehydrators, liquids unloading and well completions). The Nonpoint Oil and Gas Emissions Estimation Tool generates an intermediate file that provides SCC and county-specific emissions proportions, which are subsequently incorporated into the modeling platform.

For onroad and nonroad mobile sources, the speciation of total organic gas and particulate matter emissions has historically been performed within MOVES. However, this is now performed outside of MOVES as a post-processing step. This has the advantages of making MOVES simpler, faster to run, and makes it easier to change or update chemical mechanisms and speciation profiles used in the emissions modeling process. Some speciation is still performed inside MOVES (i.e., for “integrated species”). In many cases, these integrated species have effects like temperature or fuel effects which are not always well captured by external speciation profiles. For total organic gases, MOVES calculates 15 integrated species, such as methane and benzene, and the remainder is called NONHAPTOG and speciated outside MOVES.

In MOVES, speciation profiles are assigned by emission process, fuel subtype, regulatory class, and model year. Each of these dimensions are available in MOVES output except for fuel subtype, which is aggregated as part of each fuel type. To apply speciation outside of MOVES and make it compatible with the needs of SMOKE, we need to determine the speciation profile mapping by SMOKE process (aggregation of MOVES emission processes) and SMOKE Source Classification Code (SCC), which are defined by fuel type, source type, and road type. To support use of new ROC-based speciation profiles for mobile sources, during nonroad inventory post-MOVES processing, speciation profile assignments were updated to both NONHAPTOG and PM2.5 in a one-to-one manner. As well, to support use of these new profiles, PM2.5 was split into four parts: PEC and PSO4 (based on the new speciation profiles); total organic matter, or TOM (PNCOM plus PEC); and residual_PM, is RESID_PM (all other PM species). These profile updates are included in the tables below.

Table 3-9. Mobile Speciation Profile Updates

Pollutant	Old profile	New profile
PM2.5	8992	100CROC
PM2.5	8993	101CROC
PM2.5	8994	103CROC (starts) 102CROC (other)
PM2.5	8995	103CROC
PM2.5	8996	104CROC
PM2.5	95219a	105CROC
PM2.5	95220a	106CROC
NONHAPTOG	1001	107GROC
NONHAPTOG	8757	101GROC (starts) 103GROC (other)
NONHAPTOG	8774	104GROC
NONHAPTOG	8775	105GROC
NONHAPTOG	8855	108GROC (starts) 109GROC (other)

NONHAPTOG	8751a	100GROC (starts) 102GROC (other)
NONHAPTOG	95335a	106GROC
NONHAPTOG	95335a	106GROC
NONHAPTOG	95327	110GROC
NONHAPTOG	95328	111GROC
NONHAPTOG	95329	112GROC
NONHAPTOG	95330	113GROC
NONHAPTOG	95331	114GROC
NONHAPTOG	95332	115GROC
NONHAPTOG	95333	116GROC
NONHAPTOG	8775	105GROC
NONHAPTOG	1001	107GROC
NONHAPTOG	8860	117GROC
PM2_5	8996	109CROC
PM2_5	91106	108CROC
PM2_5	91113	107CROC
PM2_5	95219	105CROC

For this platform, MOVES runs were performed in inventory mode¹² for each representative county and season (i.e., winter and summer) to compute NONHAPTOG output by emission process, fuel type, regulatory class, and model year. Emissions were then disaggregated by fuel subtype using the market share of each fuel blend in each county. Then, emissions were normalized and aggregated to calculate the percentage of total NONHAPTOG emissions that should be speciated by each profile for each SMOKE SCC and process. Finally, these percentages were applied in SMOKE-MOVES to all counties based on their representative county. A MOVES post-processing tool was used to generate the speciation cross-references (GSREFs) for SMOKE from the outputs of the inventory mode runs.

To generate onroad emissions and to perform the subsequent speciation, SMOKE-MOVES was first run to estimate emissions and both the MEPROC and INVTABLE files were used to control which pollutants are processed and eventually integrated. From there, the NONHAPTOG emission factor tables produced by MOVES were speciated within SMOKE using the GSREF files and the NONHAPTOG GSPRO files generated by the S2S-Tool. Further details on speciation methods involving MOVES can be found in Table 3-10 and in the [associated technical reports](#) (EPA-420-R-22-017, EPA-420-R-23-006).

¹² Inventory mode was run rather than rates mode because: 1) MOVES inventory mode is faster than rates mode, 2) there are several dimensions of rates mode output which are not relevant to the assigning of speciation profiles, such as speed bin and temperature profile and 3) weighting speciation profiles by their emissions inventory is both easier and more accurate than by MOVES output activity or emission rates.

Table 3-10. Mobile NOx and HONO fractions

Fuel	Model Years	Process	NO	NOx	HONO
Gasoline	1960-1980	Running Exh.	0.975	0.017	0.008
Gasoline	1981-1990	Running Exh.	0.932	0.06	0.008
Gasoline	1991-1995	Running Exh.	0.954	0.038	0.008
Gasoline	1996-2050	Running Exh.	0.836	0.156	0.008
Gasoline	1960-1980	Start Exh.	0.975	0.017	0.008
Gasoline	1981-1990	Start Exh.	0.961	0.031	0.008
Gasoline	1991-1995	Start Exh.	0.987	0.005	0.008
Gasoline	1996-2050	Start Exh.	0.951	0.041	0.008
Diesel	1960-2003	Exhaust	0.9622	0.0298	0.008
Diesel	2004-2006	Exhaust	0.9325	0.0595	0.008
Diesel	2007-2009	Exhaust	0.7529	0.2381	0.008
Diesel	2010-2060	Exhaust	0.8035	0.1885	0.008

In Canada, a GSPRO_COMBO file is used to generate speciated gasoline emissions that account for various ethanol mixes. In Mexico, onroad emissions are pre-speciated from the MOVES-Mexico model, thus eliminating the need for a GSPRO_COMBO file. For both Canada and Mexico, nonroad VOC emissions are not defined by mode (e.g., exhaust versus evaporative), which necessitates the need for a GSPRO_COMBO file that splits total VOC into exhaust and evaporative components. In addition, MOVES-Mexico uses an older version of MOVES that is hardcoded for an older version of the CB6 chemical mechanism ("CB6-CAMx"). This version does not generate the model species XYLMN or SOAALK, so additional post-processing is performed to generate those emissions:

- $XYLMN = XYL[1] - 0.966 * NAPHTHALENE[1]$
- $PAR = PAR[1] - 0.00001 * NAPHTHALENE[1]$
- $SOAALK = 0.108 * PAR[1]$

3.2.2 PM speciation

Like VOC speciation, PM_{2.5} speciation does feature integrated species from the base inventory, though there are far fewer (only BC and SO₄). The remaining mass is either TOM (total organic matter) or RESID_PM (residual PM = PM_{2.5} – BC – SO₄ – TOM), which is speciated using SPECIATE profiles that were post-processed using the S2S-Tool. Small adjustments to the methods were needed to accommodate the reporting by California. Since California does not provide speciated PM_{2.5} emissions, total PM_{2.5} emissions for onroad and nonroad sources in California were speciated using the profile proportions estimated by MOVES in California. Finally, onroad brake and tire wear PM_{2.5} emissions were speciated in the *moves2smk* postprocessor using the SPECIATE profiles 95462 and 95460, respectively.

3.2.2.1 Diesel PM

Diesel PM emissions are explicitly included in the NEI using the pollutant names DIESEL-PM10 and DIESEL-PM25 for select mobile sources whose engines burn diesel or residual oil fuels. This includes sources in onroad, nonroad, point airport ground support equipment, point locomotives, nonpoint locomotives, and all PM from diesel or residual oil fueled nonpoint CMV. These emissions are equal to

their primary PM₁₀-PRI and PM₂₅-PRI counterparts, are exclusively from exhaust (i.e., do not include brake/tire wear), and are exclusively used in toxics modeling. Diesel PM is then speciated in SMOKE using the same speciation profiles and methods as primary PM, except that diesel PM is mapped to model species that feature “DIESEL_PM” in their species name.

3.2.3 NO_x speciation

In the NEI, NO_x emissions are inventoried on a NO₂ weighted basis, but must be speciated into NO, NO₂, and HONO. Table 3-11 provides the NO_x speciation profiles used in EPA’s modeling platforms. The only difference between the two profiles is the allocation of some NO₂ mass to HONO in the “HONO” profile. HONO emissions from mobile sources have been identified in tunnel studies and its inclusion in emissions inventories is important for urban chemistry. Here, a HONO to NO_x ratio of 0.008 was selected (Sarwar, 2008). In this modeling platform, all non-mobile sources use the “NHONO” profile, all non-onroad mobile sources (including nonroad, cmv, and rail) use the “HONO” profile, and all onroad NO_x speciation occurs within MOVES. For further details on NO_x speciation within MOVES, please see the [associated technical report](#).

Table 3-11. NO_x speciation profiles

Profile	Pollutant	Species	Mass Split Factor
HONO	NOX	NO2	0.092
HONO	NOX	NO	0.9
HONO	NOX	HONO	0.008
NHONO	NOX	NO2	0.1
NHONO	NOX	NO	0.9

3.2.4 Sulfuric Acid Vapor (SULF)

Sulfuric acid vapor (SULF) is added for coal and distillate oil fuel combustion sources to the emissions files using SO₂ emissions from the base inventory. This process utilizes profiles assignments in the GSREF file and the profiles were derived using data from AP-42 (EPA, 1998). The weight fraction of added sulfuric acid vapor is fuel specific, assumes that gaseous sulfate is primarily H₂SO₄, and is calculated as follows:

$$SULF\ emissions = SO_2\ emissions \times \frac{fraction\ of\ S\ emitted\ as\ sulfate}{fraction\ of\ S\ emitted\ as\ SO_2} \times \frac{MW\ H_2SO_4}{MW\ SO_2}$$

In the above, the molecular weight (*MW*) of sulfate and sulfur dioxide are 98 g/mol and 64 g/mol, respectively. The fractions of sulfur emissions emitted as sulfate and sulfur dioxide, as well as the resulting sulfuric acid vapor split factors, by fuel, are summarized in Table 3-12 and Table 3-13 below.

Table 3-12. Sulfate Split Factor Computation

Fuel	SCCs	Profile Code	Fraction as SO ₂	Fraction as Sulfate	Split Factor (Mass Fraction)
Bituminous	1-0X-002-YY X is 1, 2, or 3 YY is 01-19 21-0Z-002-000 Z is 2, 3, or 4	95014	0.95	0.014	$.014/.95 * 98/64 = 0.0226$
Subbituminous	1-0X-002-YY X is 1, 2, or 3 YY is 21-38	87514	0.875	0.014	$.014/.875 * 98/64 = 0.0245$
Lignite	1-0X-003-YY X is 1, 2, or 3 YY is 01-18	75014	0.75	0.014	$.014/.75 * 98/64 = 0.0286$
Residual oil	1-0X-004-YY X is 1, 2, or 3 YY is 01-06 21-0Z-005-000 Z is 2, 3, or 4	99010	0.99	0.01	$.01/.99 * 98/64 = 0.0155$
Distillate oil	1-0X-005-YY X is 1, 2, or 3 YY is 01-06 21-0Z-004-000 Z is 2, 3, or 4	99010	0.99	0.01	Same as residual oil

Table 3-13. SO₂ speciation profiles

Profile	pollutant	species	split factor
95014	SO ₂	SULF	0.0226
95014	SO ₂	SO ₂	1
87514	SO ₂	SULF	0.0245
87514	SO ₂	SO ₂	1
75014	SO ₂	SULF	0.0286
75014	SO ₂	SO ₂	1
99010	SO ₂	SULF	0.0155
99010	SO ₂	SO ₂	1

3.2.5 Speciation of Metals and Mercury

Metals and mercury emissions from the base inventory require speciation for use in modeling. Non-mercury metals must be speciated into coarse and fine size ranges for use in CMAQ, and Table 3-14, summarizes the particle size profiles used for each data category.

Table 3-14. Particle Size Speciation of Metals

Source Type	Profile	Pollutant	Fine	Coarse
Onroad	OARS	Arsenic	0.95	0.05
Onroad	ONMN	Manganese	0.4375	0.5625
Onroad	ONNI	Nickel	0.83	0.17
Onroad	CRON	Chromhex	0.86	0.14
Nonroad	NOARS	Arsenic	0.83	0.17
Nonroad	NONMN	Manganese	0.67	0.33
Nonroad	NONNI	Nickel	0.49	0.51
Nonroad	CRNR	Chromhex	0.80	0.20
Stationary	STANI	Nickel	0.59	0.41
Stationary	STACD	Cadmium	0.76	0.24
Stationary	STAMN	Manganese	0.67	0.33
Stationary	STAPB	Lead	0.74	0.26
Stationary	STABE	Beryllium	0.68	0.32
Stationary	CRSTA	Chromhex	0.71	0.29
Stationary	STARS	Arsenic	0.59	0.41

Mercury is speciated into one of the three forms used by CMAQ; elemental, divalent gaseous, and divalent particulate. Table 3-15 provides the mercury speciation profiles used in the modeling platform. All relevant SCCs were mapped to these profiles within the GSREF. A caveat is the onroad and nonroad sectors, where mercury emissions are pre-speciated in MOVES, nonroad emissions from California, which use the appropriate profiles below, and onroad emissions from California, where MOVES-based speciation is applied.

Table 3-15. Mercury Speciation Profiles

Profile Code	Description	Elemental	Divalent Gas	Particulate
HGCEM	Cement kiln exhaust	0.66	0.34	0
HGCLI	Cement clinker cooler	0	0	1
HBCMB	Fuel combustion	0.5	0.4	0.1
HGCRE	Human cremation	0.8	0.15	0.05
HGELE	Elemental only (used?)	1	0	0
HGGEO	Geothermal power plants	0.87	0.13	0
HGGLD	Gold mining	0.8	0.15	0.05
HGHCL	Chlor-Alkali plants	0.972	0.028	0
HGINC	Waste incineration	0.2	0.6	0.2
HGIND	Industrial average	0.73	0.22	0.05
HGMD	Mobile diesel	0.56	0.29	0.15
HGMG	Mobile gas	0.915	0.082	0.003
HGMET	Metal production	0.8	0.15	0.005
HGMWI	Medical waste incineration	0.2	0.6	0.2
HGPETCOKE	Petroleum coke	0.6	0.3	0.1

3.3 Temporal Allocation

Temporal allocation is the process of distributing aggregated emissions to a finer temporal resolution, thereby converting annual emissions to hourly emissions as is required by CMAQ. While the total emissions are important, the timing of the occurrence of emissions is also essential for accurately simulating ozone, PM, and other pollutant concentrations in the atmosphere. Many emissions inventories are annual or monthly in nature. Temporal allocation takes these aggregated emissions and distributes the emissions to the hours of each day. This process is typically done by applying temporal profiles to the inventories in this order: monthly, day of the week, and diurnal, with monthly and day-of-week profiles applied only if the inventory is not already at that level of detail.

The temporal factors applied to the inventory were selected using some combination of country, state, county, SCC, and pollutant. Table 3-16 summarizes the temporal aspects of emissions modeling by comparing the key approaches used for temporal processing across the sectors. In the table, “Daily temporal approach” refers to the temporal approach for getting daily emissions from the inventory using the SMOKE Temporal program. The values given are the values of the SMOKE L_TYPE setting. The “Merge processing approach” refers to the days used to represent other days in the month for the merge step. If this is not “all,” then the SMOKE merge step runs only for representative days, which could include holidays as indicated by the right-most column. The values given are those used for the SMOKE M_TYPE setting (see below for more information).

Table 3-16. Temporal settings used for the platform sectors in SMOKE

Platform sector short name	Inventory resolutions	Monthly profiles used?	Daily temporal approach	Merge processing approach	Process holidays as separate days
afdust_adj	Annual	Yes	week	all	Yes
airports	Annual	Yes	all	all	No
beis	Hourly		n/a	all	No
cmv_c1c2	Annual & hourly		All	all	No
cmv_c3	Annual & hourly		All	all	No
fertilizer	Monthly		met-based	All	Yes
livestock	Daily	No	met-based	All	No
nonpt	Annual	Yes	week	week	Yes
nonroad	Monthly		mwdss	mwdss	Yes
np_oilgas	Annual	Yes	aveday	aveday	No
onroad	Annual & monthly ¹		all	all	Yes
onroad_ca_adj	Annual & monthly ¹		all	all	Yes
canada_afdust	Annual & monthly	Yes	week	all	No
canmex_area	Monthly		week	week	No
canada_onroad	Monthly		week	week	No
mexico_onroad	Monthly		week	week	No
canmex_point	Monthly	Yes	mwdss	mwdss	No
canada_ptdust	Annual	Yes	week	all	No
canmex_ag	Annual	Yes	mwdss	mwdss	No
canada_og2D	Monthly		mwdss	mwdss	No
pt_oilgas	Annual	Yes	mwdss	mwdss	Yes
Ptegu	Annual & hourly	Yes ²	all	All	No
ptnonipm	Annual	Yes	mwdss	mwdss	Yes
ptagfire	Daily		all	all	No
ptfire-rx	Daily		all	all	No
ptfire-wild	Daily		all	all	No
ptfire_othna	Daily		all	all	No
rail	Annual	Yes	aveday	aveday	No
rlc	Annual	No ³	met-based ³	all	No ³
np_solvents	Annual	Yes	aveday	aveday	No

¹Note the annual and monthly “inventory” actually refers to the activity data (VMT, hoteling, and VPOP) for onroad. VMT and hoteling is monthly and VPOP is annual. The actual emissions are computed on an hourly basis.

²Only units that do not have matching hourly CEMS data use monthly temporal profiles.

³Except for 3 SCCs that do not use met-based speciation

The following values are used in the table. The value “all” means that hourly emissions were computed for every day of the year and that emissions potentially have day-of-year variation. The value “week”

means that hourly emissions were computed for all days in one “representative” week, representing all weeks for each month. This means emissions have day-of-week variation, but not week-to-week variation within the month. The value “mwdss” means hourly emissions for one representative Monday, representative weekday (Tuesday through Friday), representative Saturday, and representative Sunday for each month. This means emissions have variation between Mondays, other weekdays, Saturdays and Sundays within the month, but not week-to-week variation within the month. The value “aveday” means hourly emissions computed for one representative day of each month, meaning emissions for all days within a month are the same. Special situations with respect to temporal allocation are described in the following subsections.

In addition to the resolution, temporal processing includes a ramp-up period for several days prior to January 1, 2021, which is intended to mitigate the effects of initial condition concentrations. The ramp-up period was 10 days (December 22-31, 2020). For all anthropogenic sectors, emissions from December 2021 were used to fill in surrogate emissions for the end of December 2020. For biogenic emissions, December 2020 emissions were computed using year 2020 meteorology.

3.3.1 Use of FF10 format for finer than annual emissions

The FF10 inventory format for SMOKE provides a consolidated format for monthly, daily, and hourly emissions inventories. With the FF10 format, a single inventory file can contain emissions for all 12 months and the annual emissions in a single record. This helps simplify the management of numerous inventories. Similarly, daily and hourly FF10 inventories contain individual records with data for all days in a month and all hours in a day, respectively.

SMOKE prevents the application of temporal profiles on top of the “native” resolution of the inventory. For example, a monthly inventory should not have annual-to-month temporal allocation applied to it; rather, it should only have month-to-day and diurnal temporal allocation. This becomes particularly important when specific sectors have a mix of annual, monthly, daily, and/or hourly inventories. The flags that control temporal allocation for a mixed set of inventories are discussed in the SMOKE documentation. The modeling platform sectors that make use of monthly values in the FF10 files are nonroad, onroad (for activity data), and all Canada and Mexico inventories except for agriculture. Commercial marine vessels in cmv_c3 and cmv_c1c2 use hourly data in the FF10 files.

3.3.2 Temporal allocation for non-EGU sources (ptnonipm)

Most temporal profiles in ptnonipm result in primarily constant emissions for each day of the year, although some have lower emissions on Sundays. An update in the 2018 platform was an analysis of monthly temporal profiles for non-EGU point sources in the ptnonipm sector. A number of profiles were found to be not quite flat over the months but were so close to flat that the difference was not meaningful. These profiles were replaced in the cross reference to point instead to the flat monthly profile. The codes for the profiles that were replaced were: 202, 214, 220, 221, 222, 223, 227, 257, 263, 264, 265, 266, 267, 269, 271, 272, 279, 280, 295, 302, 303, 304, 305, 306, 309, 310, 327, 329, 332, and 333.

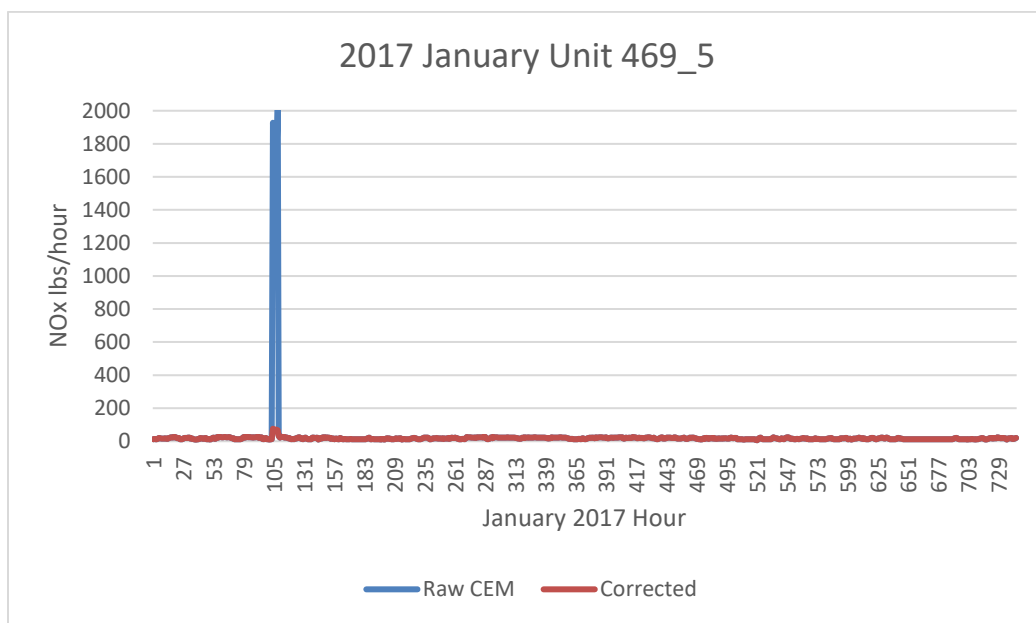
3.3.3 Electric Generating Utility temporal allocation (ptegu)

Electric generating unit (EGU) sources matched to ORIS units were temporally allocated to hourly emissions needed for modeling using the hourly CEMS data for units that could be matched to the CEMS

emissions. Those hourly data were processed through v2.1 of the CEMCorrect tool to mitigate the impact of unmeasured values in the data.

The temporal allocation procedure for EGUs in the base year is differentiated by whether or not the unit could be directly matched to a unit with CEMS data via its ORIS facility code and boiler ID. Note that for units matched to CEMS data, annual totals of their emissions input to CMAQ may be different than the values in the annual inventory because the CEMS data replace the NO_x and SO₂ annual inventory data for the seasons in which the CEMS are operating. If a CEMS-matched unit is determined to be a partial year reporter, as can happen for sources that run CEMS only in the summer, emissions totaling the difference between the annual emissions and the total CEMS emissions are allocated to the non-summer months. Prior to use of the CEMS data in SMOKE it is processed through the CEMCorrect tool. The CEMCorrect tool identifies hours for which the data were not measured as indicated by the data quality flags in the CEMS data files. Unmeasured data can be filled in with maximum values and thereby cause erroneously high values in the CEMS data. When data were flagged as unmeasured and the values were found to be more than three times the annual mean for that unit, the data for those hours were replaced with annual mean values (Adelman et al., 2012). These adjusted CEMS data were then used for the remainder of the temporal allocation process described below (see Figure 3-3 for an example).

Figure 3-3. Eliminating unmeasured spikes in CEMS data



The region, fuel, and type (peaking or non-peaking) must be identified for each input EGU with CEMS data so the data can be used to generate profiles. The identification of peaking units was done using hourly heat input data from 2021 and the two previous years (2019 and 2020). The heat input was summed for each year. Equation 1 shows how the annual heat input value is converted from heat units (BTU/year) to power units (MW) using the NEEDS v6 derived unit-level heat rate (BTU/kWh). In equation 2 a capacity factor is calculated by dividing the annual unit MW value by the NEEDS v6 unit capacity

value (MW) multiplied by the hours in the year. A peaking unit was defined as any unit that had a maximum capacity factor of less than 0.2 for every year (2019, 2020, and 2021) and a 3-year average capacity factor of less than 0.1.

Equation 1. Annual unit power output

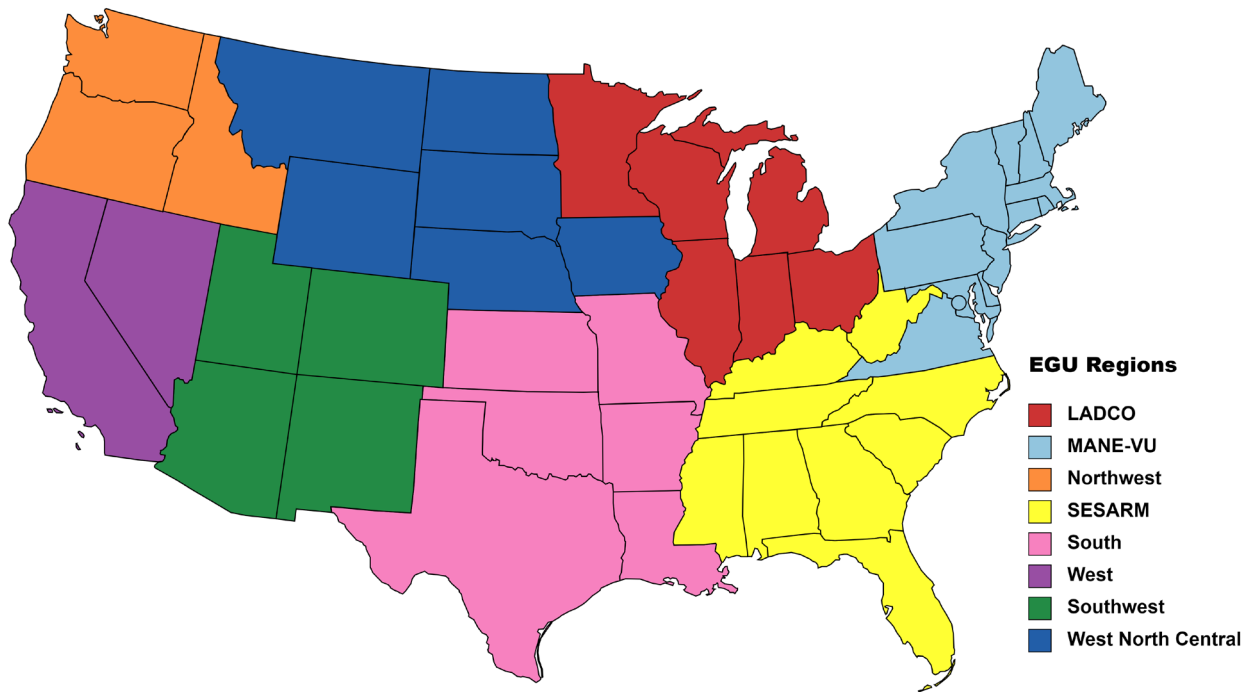
$$\text{Annual Unit Output (MW)} = \frac{\sum_{i=0}^{8760} \text{Hourly HI (BTU)} * 1000 \left(\frac{\text{MW}}{\text{kW}} \right)}{\text{NEEDS Heat Rate} \left(\frac{\text{BTU}}{\text{kWh}} \right)}$$

Equation 2. Unit capacity factor

$$\text{Capacity Factor} = \frac{\text{Annual Unit Output (MW)}}{\text{NEEDS Unit Capacity} \left(\frac{\text{MW}}{\text{h}} \right) * 8760 \text{ (h)}}$$

Input regions were determined from one of the eight EGU modeling regions based on MJO and climate regions. Regions were used to group units with similar climate-based load demands. Region assignment is made on a state level, where all units within a state were assigned to the appropriate region. Unit fuel assignments were made using the primary NEEDS v6 fuel. Units fueled by bituminous, subbituminous, or lignite were assigned to the coal fuel type. Natural gas units were assigned to the gas fuel type. Distillate and residual fuel oil were assigned to the oil fuel type. Units with any other primary fuel were assigned the “other” fuel type. Figure 3-4 shows the regions used to generate the profiles. Currently there are 64 unique profiles available based on 8 regions, 4 fuels, and 2 for peaking unit status (peaking and non-peaking).

Figure 3-4. Regions used to Compute Temporal non-CEMS EGU Temporal Profiles



The daily and diurnal profiles were calculated for each region, fuel, and peaking type group from the year 2021 CEMS heat input values. The heat input values were summed for each input group to the annual level at each level of temporal resolution: monthly, month-of-day, and diurnal. The sum by temporal resolution value was then divided by the sum of annual heat input in that group to get a set of temporalization factors. Diurnal factors were created for both the summer and winter seasons to account for the variation in hourly load demands between the seasons. For example, the sum of all hour 1 heat input values in the group was divided by the sum of all heat inputs over all hours to get the hour 1 factor. Each grouping contained 12 monthly factors, up to 31 daily factors per month, and two sets of 24 hourly factors. The profiles were weighted by unit size where the units with more heat input have more influence on the shape of the profile. Composite profiles were created for each region and type across all fuels as a way to provide profiles for a fuel type that does not have hourly CEMS data in that region. Figure 3-5 shows peaking and non-peaking daily temporal profiles for the gas fuel type in the LADCO region. Figure 3-6 shows the diurnal profiles for the coal fuel type in the Mid-Atlantic Northeast Visibility Union (MANE-VU) region.

Figure 3-5. Example Daily Temporal Profiles for the LADCO Region and the Gas Fuel Type

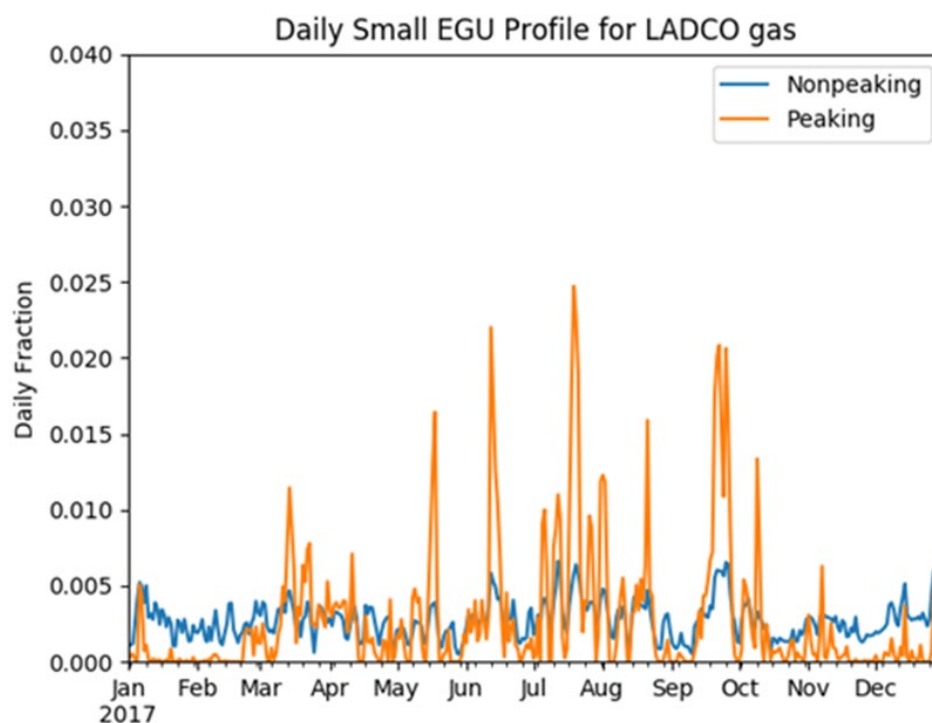
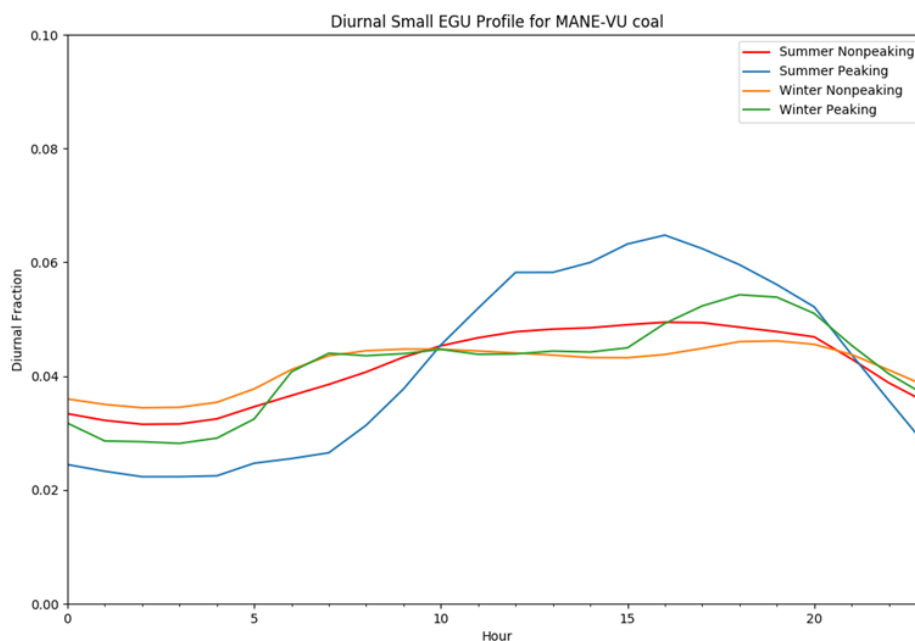


Figure 3-6. Example Diurnal Temporal Profiles for the MANE-VU Region and the Coal Fuel Type



SMOKE uses a cross-reference file to select a monthly, daily, and diurnal profile for each source. For the 2021 platform, the temporal profiles were assigned in the cross-reference at the unit level to EGU sources without hourly CEMS data. An inventory of all EGU sources without CEMS data was used to identify the region, fuel type, and type (peaking/non-peaking) of each source. The region used to select

the temporal profile is assigned based on the state from the unit FIPS. The fuel was assigned by SCC to one of the four fuel types: coal, gas, oil, and other. A fuel type unit assignment is made by summing the VOC, NOX, PM2.5, and SO2 for all SCCs in the unit. The SCC that contributed the highest total emissions to the unit for selected pollutants was used to assign the unit fuel type. Peaking units were identified as any unit with an oil, gas, or oil fuel type with a NAICS of 22111 or 221112. Some units may be assigned to a fuel type within a region that does not have an available input unit with a matching fuel type in that region. These units without an available profile for their group were assigned to use the regional composite profile. MWC and cogen units were identified using the NEEDS primary fuel type and cogeneration flag, respectively, from the NEEDS v6 database. Assignments for each unit needed a profile were made using the regions shown in Figure 3-4.

3.3.4 Airport Temporal allocation (airports)

Airport temporal profiles were updated to 2021-specific temporal profiles for all airports other than Alaska seaplanes. Hourly airport operations data were obtained from the Aviation System Performance Metrics (ASPM) Airport Analysis website (<https://aspm.faa.gov/apm/sys/AnalysisAP.asp>). A report of 2021 hourly Departures and Arrivals for Metric Computation by airport was generated. An overview of the ASPM metrics is at http://aspmhelp.faa.gov/index.php/Aviation_Performance_Metrics_%28APM%29. Figure 3-7 shows examples of diurnal airport profiles for the Phoenix airport (PHX) and the default profile for Texas.

Month-to-day and Annual-to month temporal profiles were developed based on a separate query of the 2021 Aviation System Performance Metrics (ASPM) Airport Analysis (<https://aspm.faa.gov/apm/sys/AnalysisAP.asp>). A report of all airport operations (takeoffs and landings) by day for 2021 was generated. Annual-to-month profiles were derived directly from the daily airport operations report and examples are shown for Wisconsin and Atlanta in Figure 3-8.

For 2020, all airport SCCs (i.e., 2275*, 2265008005, 2267008005, 2268008005 and 2270008005) were assigned to individual commercial airports where a match could be made between the inventory facility and the FAA identifier in the ASPM derived data. State average profiles were calculated as the average of the temporal fractions for all airports within a state. The state average profiles were assigned by state to all airports in the inventory that did not have an airport specific match in the ASPM data. Package processing hubs at the Memphis (MEM), Indianapolis (IND), Louisville (SDF), and Chicago Rockford (RFD) airports produced peaks in the average state profiles at times not typical for activity in smaller commercial airports. These packaging hubs were removed from the state averages. Airports that required state-defaults in states lacking ASPM data use national average profiles calculated from the average of the state temporal profiles.

Alaska seaplanes, which are outside the CONUS domain use the monthly profile in Figure 3-9. These were assigned based on the facility ID.

Figure 3-7. 2021 Airport Diurnal Profiles for PHX and state of Texas

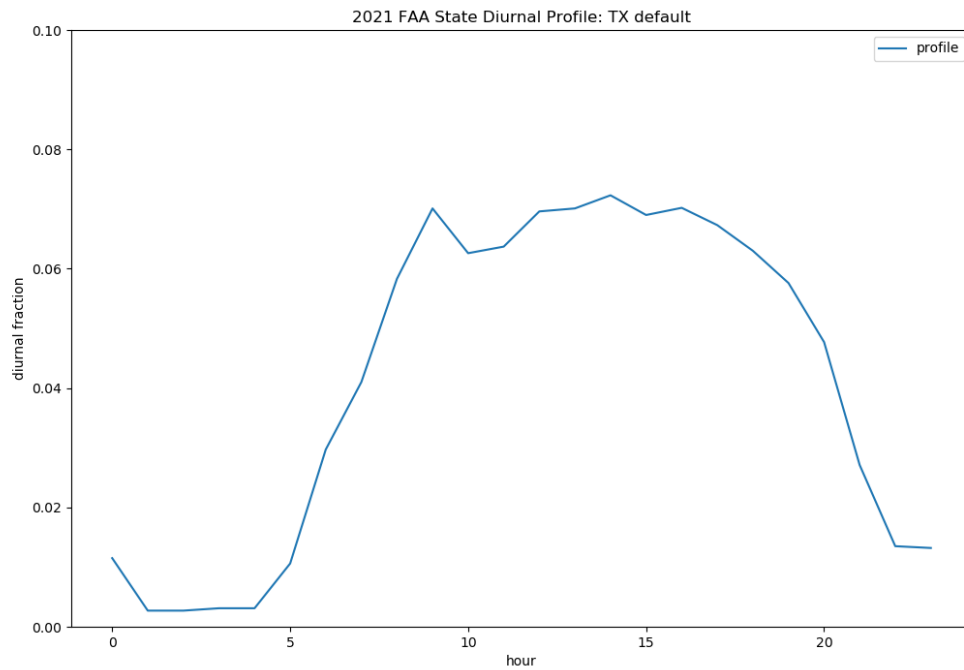
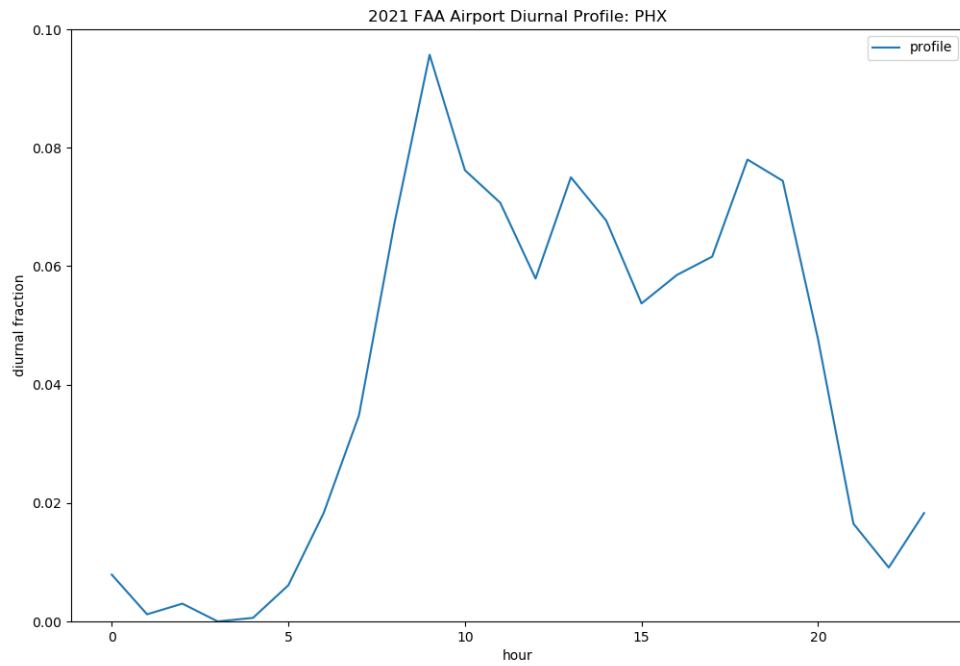


Figure 3-8. 2021 Wisconsin and Atlanta annual-to-month profile for airport emissions

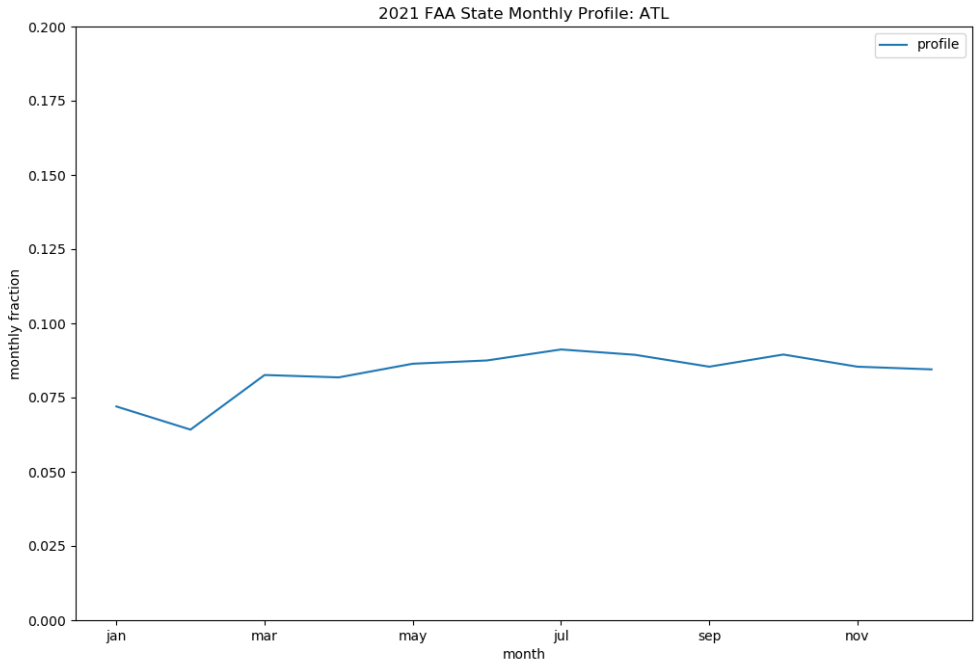
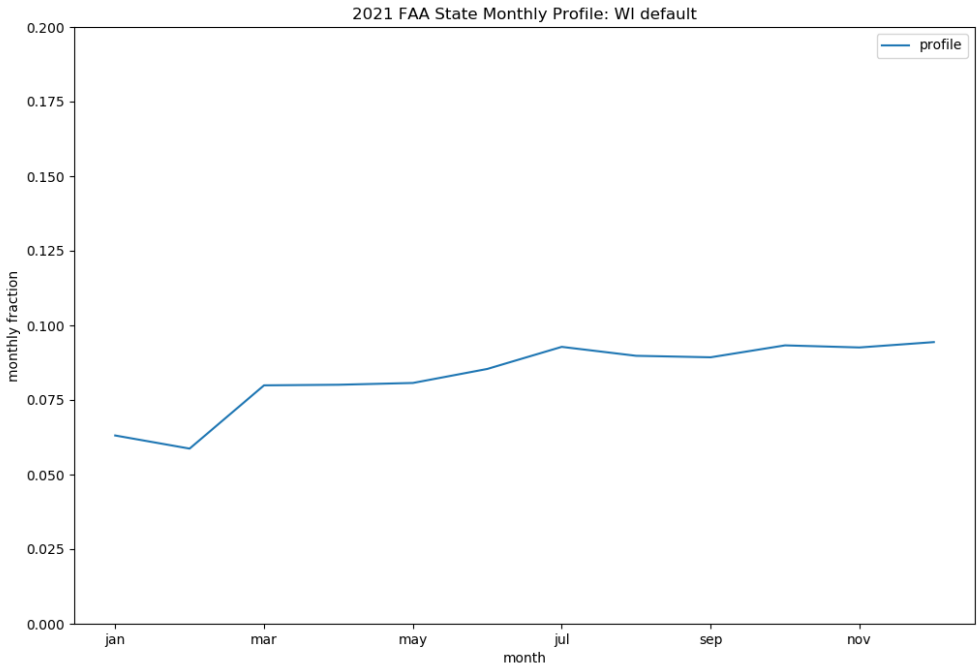
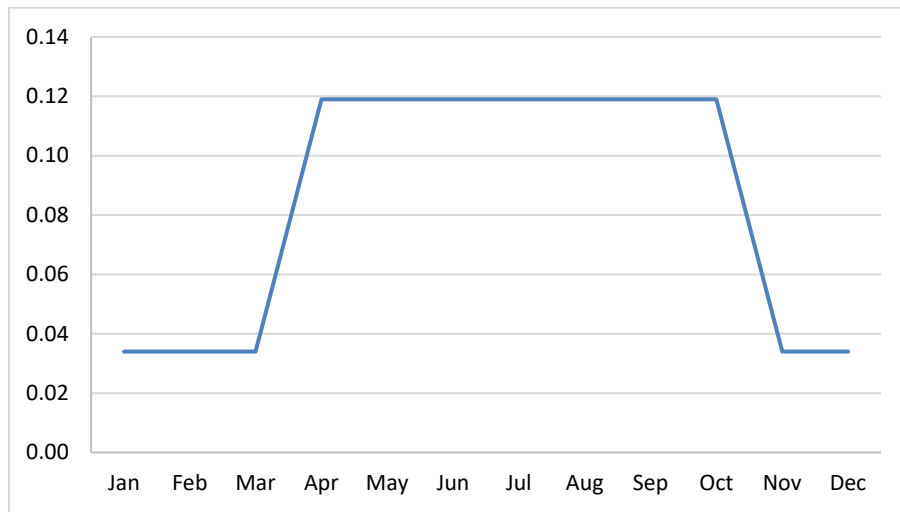


Figure 3-9. Alaska seaplane profile



3.3.5 Residential Wood Combustion Temporal allocation (rwc)

There are many factors that impact the timing of when emissions occur, and for some sectors this includes meteorology. The benefits of utilizing meteorology as a method for temporal allocation are: (1) a meteorological dataset consistent with that used by the AQ model is available (e.g., outputs from WRF); (2) the meteorological model data are highly resolved in terms of spatial resolution; and (3) the meteorological variables vary at hourly resolution and can, therefore, be translated into hour-specific temporal allocation.

The SMOKE program Gentpro provides a method for developing meteorology-based temporal allocation. Currently, the program can utilize three types of temporal algorithms: annual-to-day temporal allocation for residential wood combustion (RWC); month-to-hour temporal allocation for agricultural livestock NH_3 ; and a generic meteorology-based algorithm for other situations. Meteorological-based temporal allocation was used for portions of the rwc sector and for the entire ag sector.

Gentpro reads in gridded meteorological data (output from MCIP) along with spatial surrogates and uses the specified algorithm to produce a new temporal profile that can be input into SMOKE. The meteorological variables and the resolution of the generated temporal profile (hourly, daily, etc.) depend on the selected algorithm and the run parameters. For more details on the development of these algorithms and running Gentpro, see the Gentpro documentation and the SMOKE documentation at

http://www.cmascenter.org/smoke/documentation/3.1/GenTPRO_TechnicalSummary_Aug2012_Final.pdf and <https://www.cmascenter.org/smoke/documentation/4.5/html/ch05s03s05.html>, respectively.

For the RWC sector, two different algorithms for calculating temporal allocation are used. For most SCCs in the sector, in which wood burning is more prominent on colder days, Gentpro was used to compute annual to day-of-year temporal profiles based on the daily minimum temperature. These profiles distribute annual RWC emissions to the coldest days of the year. On days where the minimum temperature does not drop below a user-defined threshold, RWC emissions for most sources in the

sector are zero. Conversely, the program temporally allocates the largest percentage of emissions to the coldest days. Similar to other temporal allocation profiles, the total annual emissions do not change, only the distribution of the emissions within the year is affected. The temperature threshold for RWC emissions was 50 °F for most of the country, and 60 °F for the following states: Alabama, Arizona, California, Florida, Georgia, Louisiana, Mississippi, South Carolina, and Texas. The algorithm is as follows:

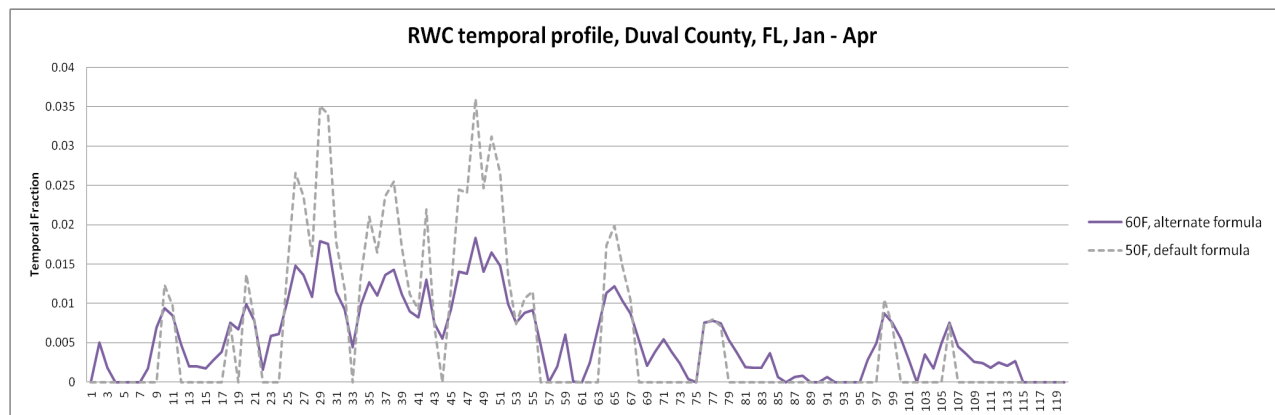
If $T_d \geq T_t$: no emissions that day
 If $T_d < T_t$: daily factor = $0.79 \cdot (T_t - T_d)$

where (T_d = minimum daily temperature; T_t = threshold temperature, which is 60 degrees F in southern states and 50 degrees F elsewhere).

Once computed, the factors were normalized to sum to 1 to ensure that the total annual emissions are unchanged (or minimally changed) during the temporal allocation process.

Figure 3-10 illustrates the impact of changing the temperature threshold for a warm climate county. The plot shows the temporal fraction by day for Duval County, Florida, for the first four months of 2007. The default 50 °F threshold creates large spikes on a few days, while the 60 °F threshold dampens these spikes and distributes a small amount of emissions to the days that have a minimum temperature between 50 and 60 °F.

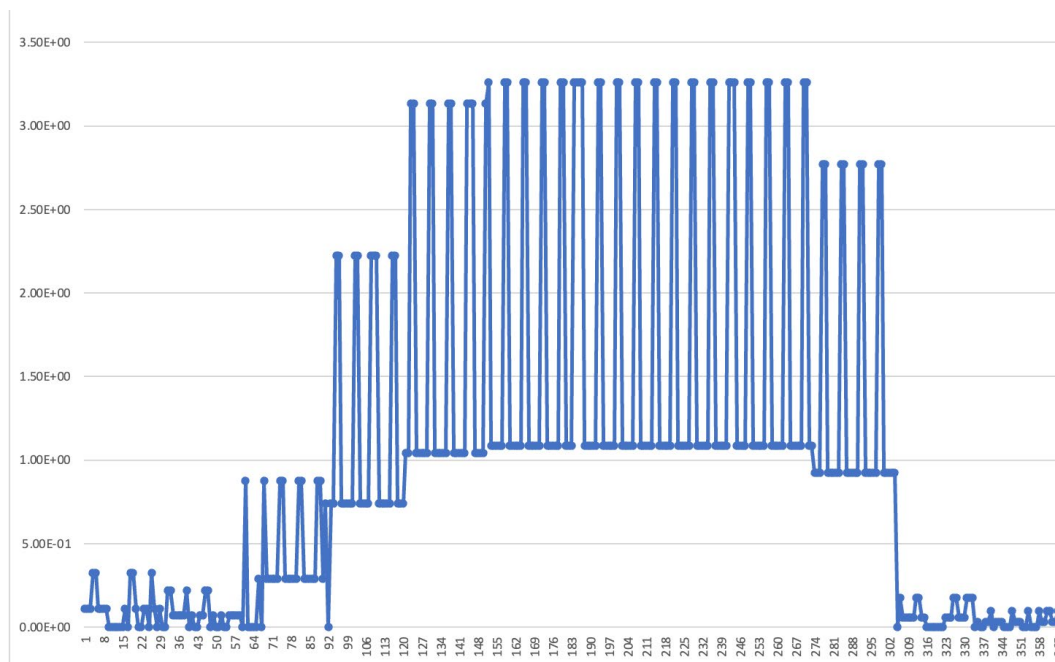
Figure 3-10. Example of RWC temporal allocation using a 50 versus 60 °F threshold



For the 2021 emissions modeling platform, a separate algorithm is used to determine temporal allocation of recreational wood burning, e.g. fire pits (SCC 2104008700) and is applied by Gentpro. Recreational wood burning depends on both minimum and maximum daily temperatures by county, and also uses a day-of-week temporal profile (61500) in which emissions are much higher on weekends than on weekdays. According to the recreational wood burning algorithm, only days in which the temperature falls within a range of 50°F and 80°F at some point during the day receive emissions. On days when the maximum temperature is less than 50°F or the minimum temperature is above 80°F, the daily temporal factor is zero. For all other days, the day-of-week profile 61500 is applied, which has 33% of the emissions on each weekend day and lower emissions on weekdays. An example is shown in Figure 3-11. As a result of applying this algorithm, northern states have more recreational wood burning in

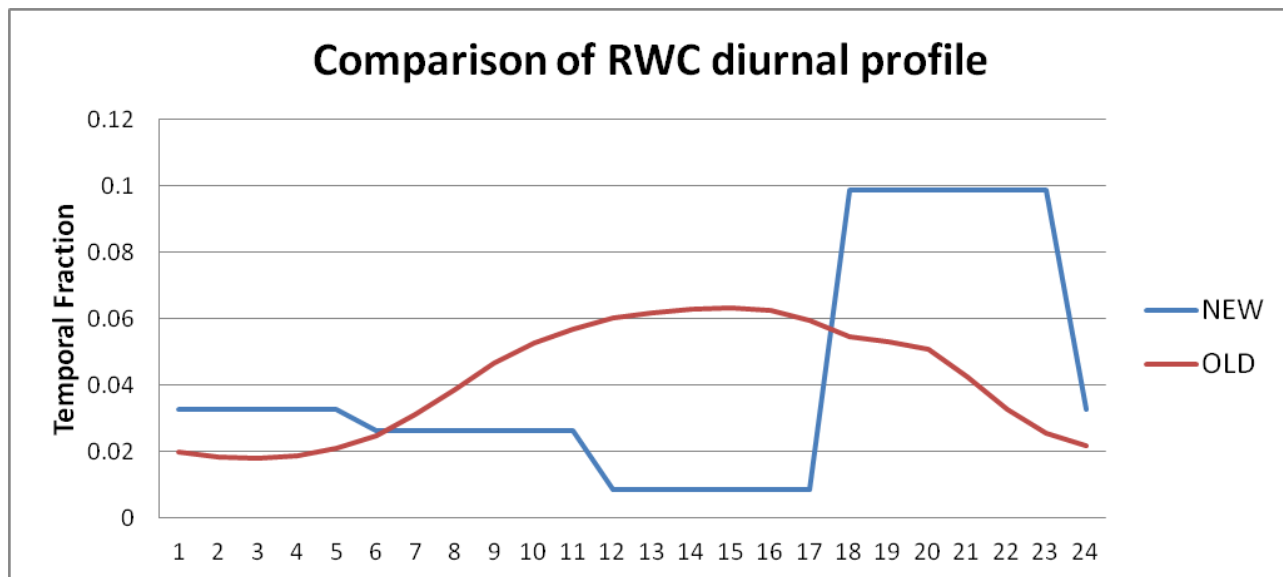
summer months while southern states show a flatter pattern with emissions distributed more evenly throughout the months.

Figure 3-11. Example of Annual-to-day temporal pattern of recreational wood burning emissions



The diurnal profile used for most RWC sources (see Figure 3-12) places more of the RWC emissions in the morning and the evening when people are typically using these sources. This profile is based on a 2004 MANE-VU survey based temporal profiles (https://s3.amazonaws.com/marama.org/wp-content/uploads/2019/11/04184303/Open_Burning_Residential_Areas_Emissions_Report-2004.pdf). This profile was created by averaging three indoor and three RWC outdoor temporal profiles from counties in Delaware and aggregating them into a single RWC diurnal profile. This new profile was compared to a concentration-based analysis of aethalometer measurements in Rochester, New York (Wang *et al.* 2011) for various seasons and days of the week and was found that the new RWC profile generally tracked the concentration based temporal patterns.

Figure 3-12. RWC diurnal temporal profile



The temporal profiles for hydronic heaters” (i.e., SCCs=2104008610 [outdoor], 2104008620 [indoor], and 2104008620 [pellet-fired]) are not based on temperature data, because the meteorologically based temporal allocation used for the rest of the rwc sector did not agree with observations for how these appliances are used.

For hydronic heaters, the annual-to-month, day-of-week and diurnal profiles were modified based on information in the New York State Energy Research and Development Authority’s (NYSERDA) “Environmental, Energy Market, and Health Characterization of Wood-Fired Hydronic Heater Technologies, Final Report” (NYSERDA, 2012), as well as a Northeast States for Coordinated Air Use Management (NESAUM) report “Assessment of Outdoor Wood-fired Boilers” (NESAUM, 2006). A Minnesota 2008 Residential Fuelwood Assessment Survey of individual household responses (MDNR, 2008) provided additional annual-to-month, day-of-week, and diurnal activity information for OHH as well as recreational RWC usage.

Data used to create the diurnal profile for hydronic heaters, shown in Figure 3-13, are based on a conventional single-stage heat load unit burning red oak in Syracuse, New York.

Annual-to-month temporal allocation for OHH was computed from the MDNR 2008 survey and is illustrated in Figure 3-14. The hydronic heater emissions still exhibit strong seasonal variability, but do not drop to zero because many units operate year-round for water and pool heating.

Figure 3-13. Data used to produce a diurnal profile for hydronic heaters

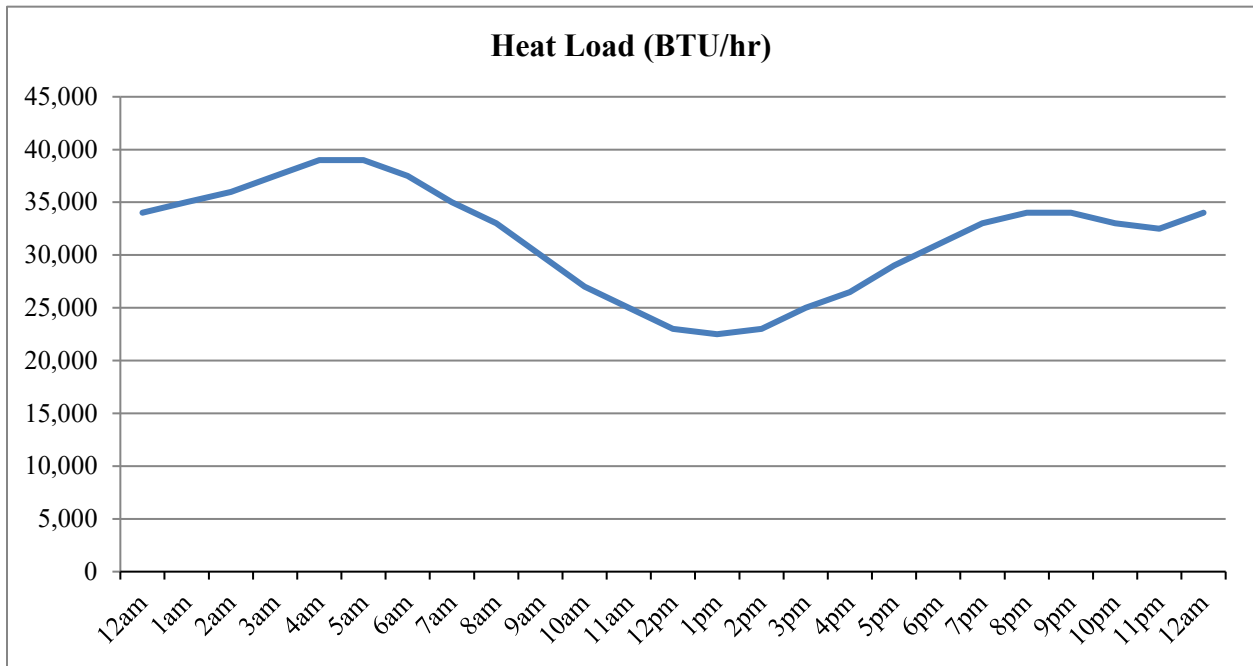
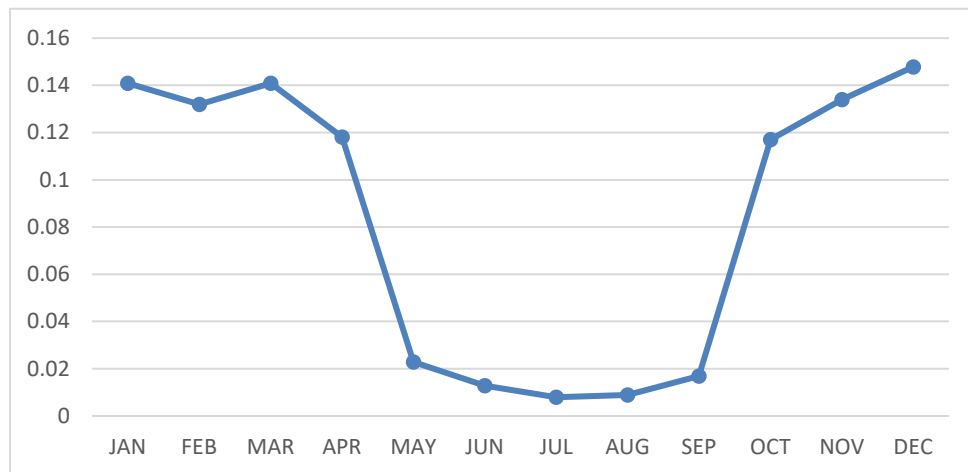


Figure 3-14. Monthly temporal profile for hydronic heaters



3.3.6 Agricultural Ammonia Temporal Profiles (livestock)

For the ag sector, agricultural GenTPRO temporal allocation was applied to livestock emissions and to all pollutants within the sector, not just NH₃. The GenTPRO algorithm is based on an equation derived by Jesse Bash of EPA ORD based on the Zhu, Henze, et al. (2014) empirical equation. This equation is based on observations from the TES satellite instrument with the GEOS-Chem model and its adjoint to estimate

diurnal NH₃ emission variations from livestock as a function of ambient temperature, aerodynamic resistance, and wind speed. The equations are:

$$E_{i,h} = [161500/T_{i,h} \times e^{(-1380/T_{i,h})}] \times AR_{i,h} \quad \text{Equation 3-1}$$

$$PE_{i,h} = E_{i,h} / \text{Sum}(E_{i,h}) \quad \text{Equation 3-2}$$

where

- $PE_{i,h}$ = Percentage of emissions in county i on hour h
- $E_{i,h}$ = Emission rate in county i on hour h
- $T_{i,h}$ = Ambient temperature (Kelvin) in county i on hour h
- $AR_{i,h}$ = Aerodynamic resistance in county i

Some examples plots of the profiles by animal type in different parts of the country are shown in Figure 3-15.

To develop month-to-hour temporal profiles of livestock emissions, GenTPRO was run using the “BASH_NH3” profile method to create for these sources. Because these profiles distribute to the hour based on monthly emissions, the monthly emissions were obtained from a monthly inventory, or from an annual inventory that has been temporalized to the month. Figure 3-16 compares the daily emissions for Minnesota from the “old” approach (uniform monthly profile) with the “new” approach (GenTPRO generated month-to-hour profiles) for 2014. Although the GenTPRO profiles show daily (and hourly) variability, the monthly total emissions are the same between the two approaches.

Figure 3-15. Examples of livestock temporal profiles in several parts of the country

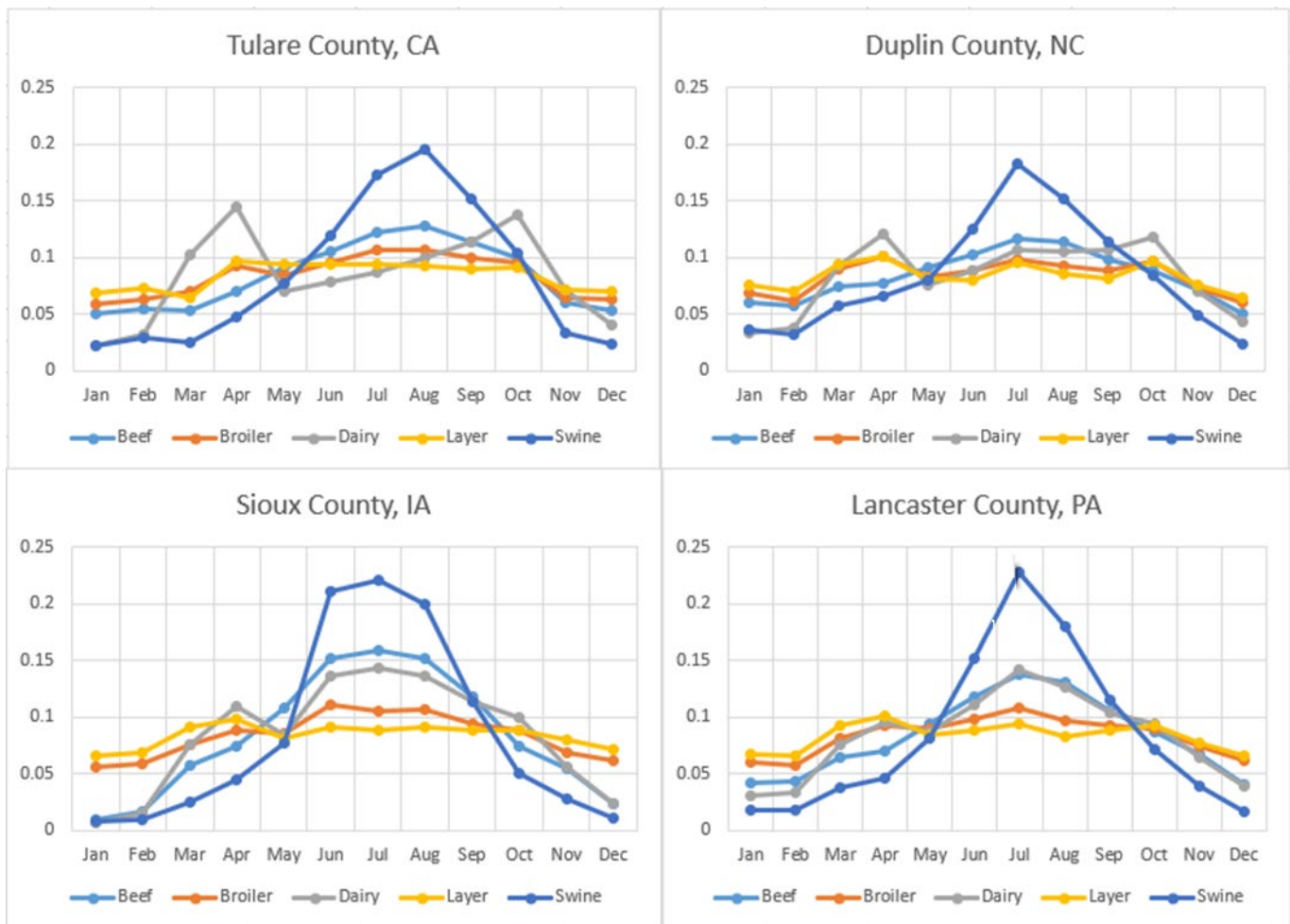
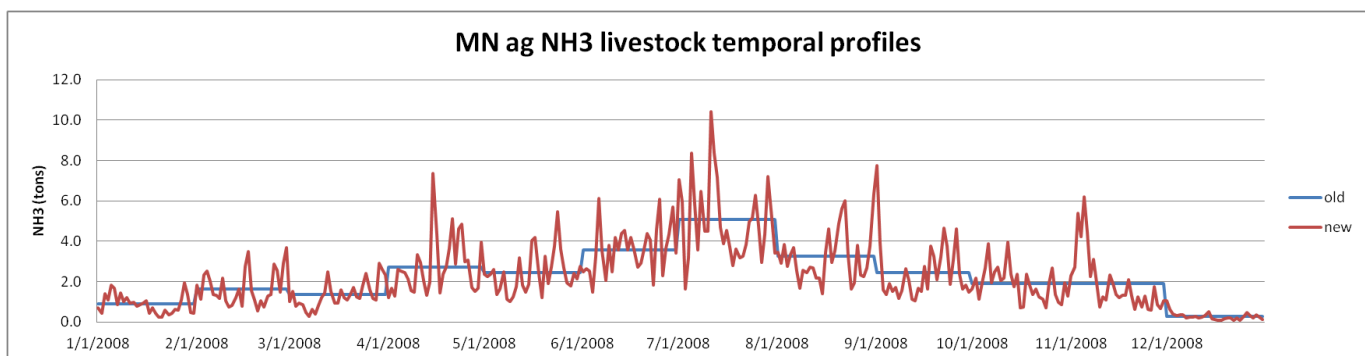


Figure 3-16. Example of animal NH₃ emissions temporal allocation approach (daily total emissions)



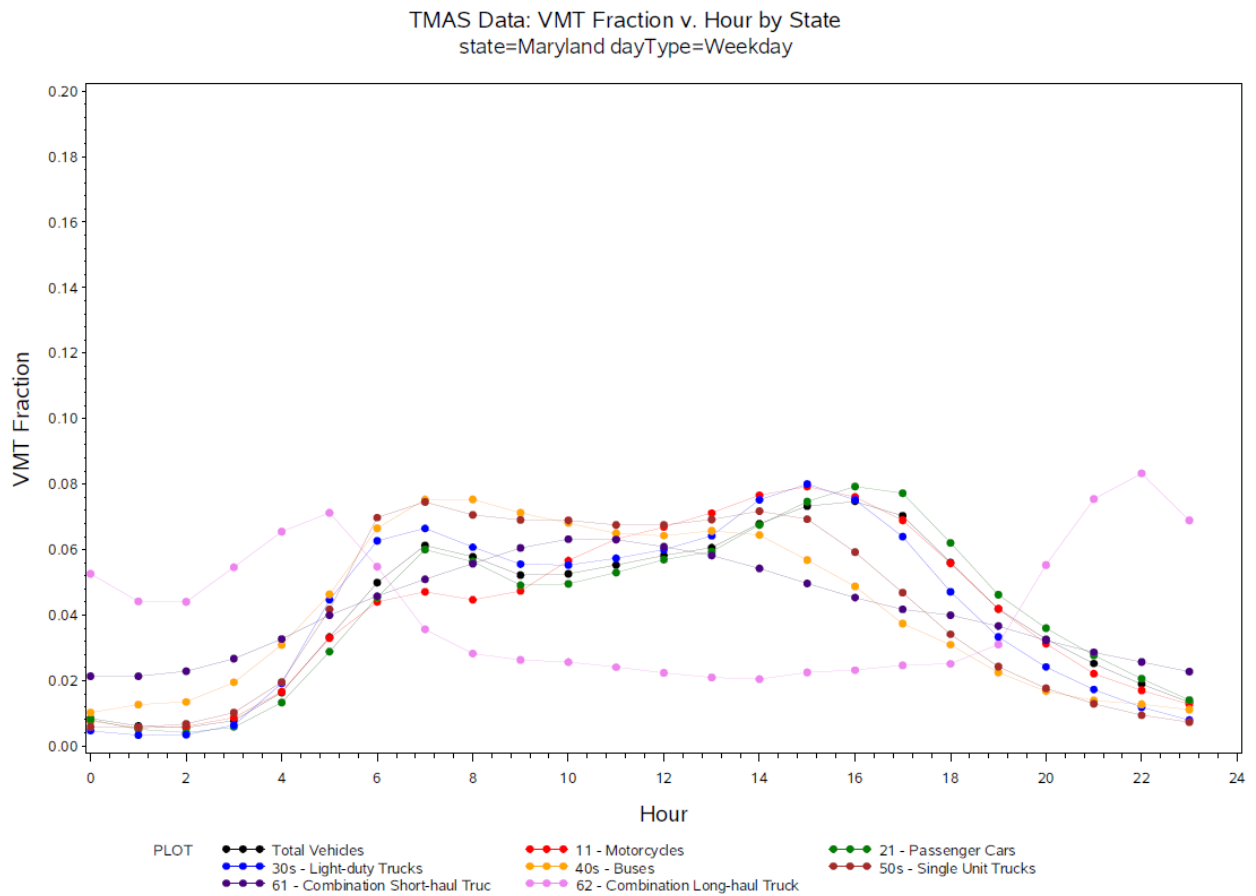
3.3.7 Oil and gas temporal allocation (np_oilgas)

Monthly temporalization of np_oilgas emissions is based primarily on year-specific monthly factors from the Oil and Gas Tool (OGT). Factors were specific to each county and SCC. For use in SMOKE, each unique set of factors was assigned a label (OG20M_0001 through OG20M_6306), and then a SMOKE-formatted ATPRO_MONTHLY and an ATREF were developed. This dataset of monthly temporal factors included profiles for all counties and SCCs in the Oil and Gas Tool inventory. Because we are using non-tool datasets in some states, this monthly temporalization dataset did not cover all counties and SCCs in the entire inventory used for this study. To fill in the gaps in those states, state average monthly profiles for oil, natural gas, and combination sources were calculated from Energy Information Administration (EIA) data and assigned to each county/SCC combination not already covered by the OGT monthly temporal profile dataset. Coal bed methane (CBM) and natural gas liquid sources were assigned flat monthly profiles where there was not already a profile assignment in the ERG dataset.

3.3.8 Onroad mobile temporal allocation (onroad)

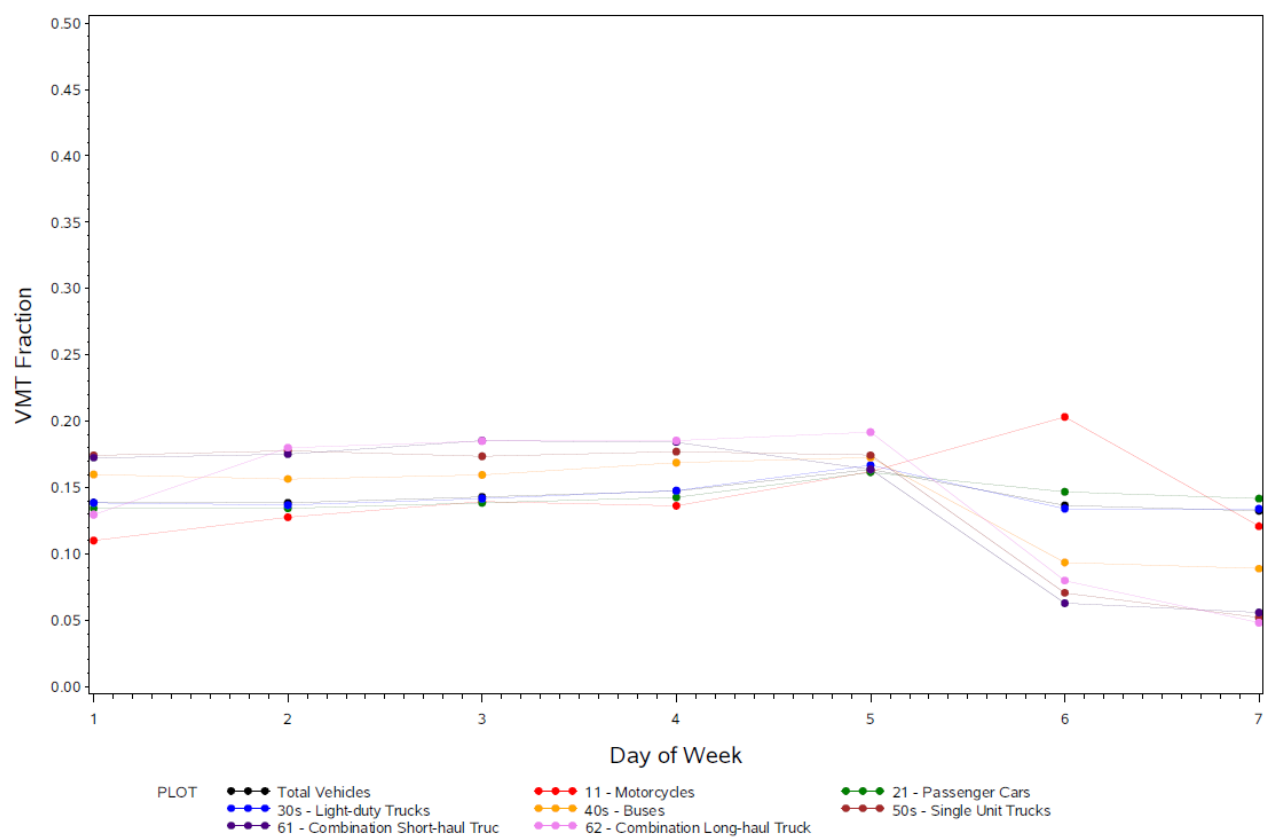
For the onroad sector, the temporal distribution of emissions is a combination of traditional temporal profiles and the influence of meteorology. For the 2021 platform EPA utilized the FHWA's Travel Monitoring and Analysis System (TMAS). This system measures monthly traffic volume, by class and weight. The primary purpose for using TMAS in 2021 platform was to replace the month VMT distribution from 2020, which were marked by the COVID pandemic shutdowns in mid-March/April. The 2021 TMAS month VMT distribution looks more like a typical nonpandemic year. We also used day/hour distributions from the same dataset because they were available and for the correct year (2021). TMAS data was processed for each state, for each month, and vehicle class. Figure 3-17 shows TMAS data. The first plot shows hour of the day for the state of Maryland. Note that you can see the the rush hour in the morning and the evening. The second plot shows the state of Minnesota for the month of June. Notice that motorcycles come out in the spring (winter months show less VMT for motorcycles) and are driven more on Saturday. The third plot shows an annual, by month plot of Montana. Note that there is an increase in passenger cars and light duty trucks during the month of July. This may be due to an increase in tourism during the warmer months.

Figure 3-17. TMAS Data: VMT Fraction by Hour of Day, Day of Week, and Month of Year

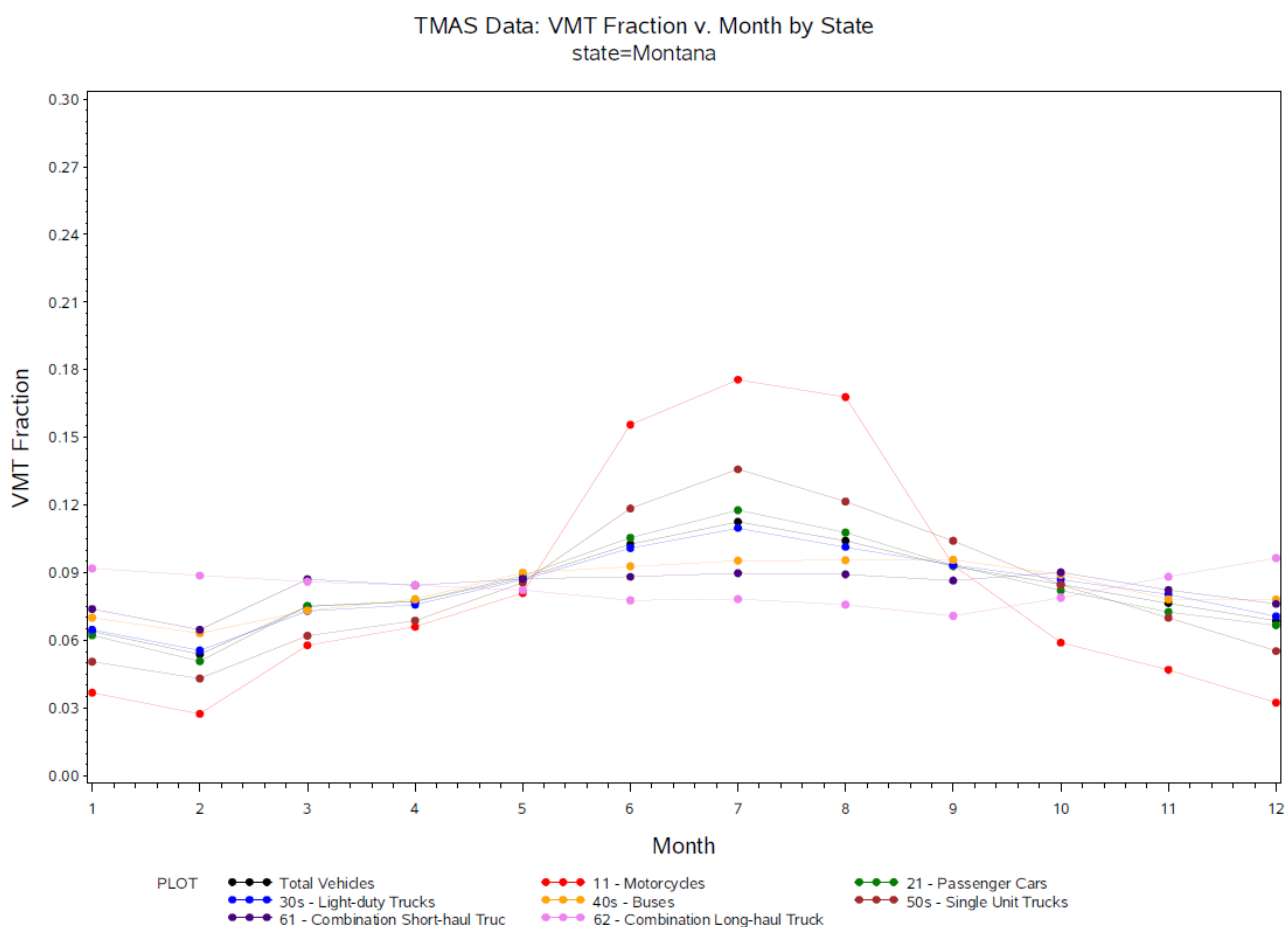


/proj1/EPA_2020_NEI/FHWA_TMAS_Class_Data_2021/plot_TMAS_hour.sas 13SEP23 13:59

TMAS Data: VMT Fraction v. Day of Week by State
state=Minnesota monthID=6



/proj1/EPA_2020_NEI/FHWA_TMAS_Class_Data_2021/plot_TMAS_day.sas 22SEP23 13:48



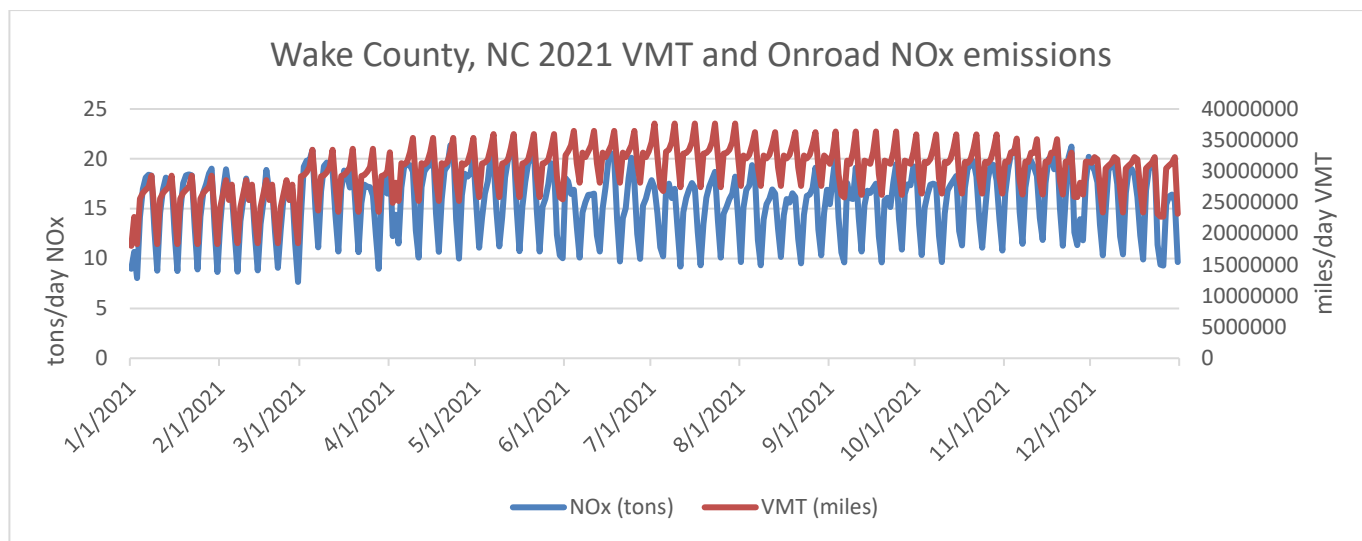
/proj1/EPA_2020_NEI/FHWA_TMAS_Class_Data_2021/plot_TMAS.sas 15SEP23 15:17

The “inventories” referred to in Table 3-16 consist of activity data for the onroad sector, not emissions. VMT is the activity data used for on-network rate-per-distance (RPD) processes. For the off-network emissions from the rate-per-profile (RPP) and rate-per-vehicle (RPV) processes, the VPOP activity data are annual and do not need temporal allocation. For rate-per-hour (RPH) processes that result from hoteling of combination trucks, the HOTELING inventory is annual and was temporalized to month, day of the week, and hour of the day through temporal profiles. Day-of-week and hour-of-day temporal profiles are also used to temporalize the starts activity used for rate-per-start (RPS) processes, and the off-network idling (ONI) hours activity used for rate-per-hour-ONI (RPHO) processes. The inventories for starts and ONI activity contain monthly activity so that monthly temporal profiles are not needed.

For on-roadway RPD processes, the VMT activity data are annual for some sources and monthly for other sources, depending on the source of the data. Sources without monthly VMT were temporalized from annual to month through temporal profiles. VMT was also temporalized from month to day of the week, and then to hourly through temporal profiles. The RPD processes also use hourly speed distributions (SPDIST) as discussed in Section 2.3. For onroad, the temporal profiles and SPDIST will impact not only the distribution of emissions through time but also the total emissions. SMOKE-MOVES calculates emissions for RPD processed based on the VMT, speed and meteorology. Thus, if the VMT or

speed data were shifted to different hours, it would align with different temperatures and hence different emission factors. In other words, two SMOKE-MOVES runs with identical annual VMT, meteorology, and MOVES emission factors, will have different total emissions if the temporal allocation of VMT changes. Figure 3-18 (from 2020) illustrates the temporal allocation of the onroad activity data (i.e., VMT) and the pattern of the emissions that result after running SMOKE-MOVES. In this figure, it can be seen that the meteorologically varying emission factors add variation on top of the temporal allocation of the activity data.

Figure 3-18. Example temporal variability of VMT compared to onroad NO_x emissions



Meteorology is not used in the development of the temporal profiles, but rather it impacts the calculation of the hourly emissions through the program Movesmrg. The result is that the emissions vary at the hourly level by grid cell. More specifically, the on-network (RPD) and the off-network parked and stationary vehicle (RPV, RPH, RPHO, RPS, and RPP) processes use the gridded meteorology (MCIP) either directly or indirectly. For RPD, RPV, RPH, RPHO, and RPS, Movesmrg determines the temperature for each hour and grid cell and uses that information to select the appropriate emission factor for the specified SCC/pollutant/mode combination. For RPP, instead of reading gridded hourly meteorology, Movesmrg reads gridded daily minimum and maximum temperatures. The total of the emissions from the combination of these six processes (RPD, RPV, RPH, RPHO, RPS, and RPP) comprise the onroad sector emissions. In summary, the temporal patterns of emissions in the onroad sector are influenced by meteorology.

Day-of-week, hour-of-day, and month-of-year temporal profiles for VMT were developed from TMAS data. Data were provided for motorcycles (11), passenger vehicles (21), lightduty trucks (30s), buses (40s), single unit trucks (50s), and combination short-haul trucks (61), and combination long-haul trucks (62). The dataset includes temporal profiles for individual states.

The StreetLight temporal profiles were used in areas of the contiguous United States that did not submit temporal profiles of sufficient detail for the 2020 NEI. For this platform, the data selection hierarchy favored local input data over EPA-developed information, with the exception of the three MOVES tables `hourVMTFraction`, `dayVMTFraction`, and `avgSpeedDistribution` where county-level, telematics-based

EPA Defaults were adopted for the NEI universally. For 2021, data from December 2020 were used for speed distributions. For hoteling, day-of-week profiles are the same as non-hoteling for combination trucks, while hour-of-day non-hoteling profiles for combination trucks were inverted to create new hoteling profiles that peak overnight instead of during the day.

Temporal profiles for RPHO are based on the same temporal profiles as the on-network processes in RPD, but since the on-network profiles are road-type-specific and ONI is not road-type-specific, the RPHO profiles were assigned to use rural unrestricted profiles for counties considered "rural" and urban unrestricted profiles for counties considered "urban". RPS uses the same day-of-week profiles as on-network processes in RPD, but uses a separate set of diurnal temporal profiles specifically for starts activity. For starts, there are two hour-of-day temporal profiles for each source type, one for weekdays and one for weekends. The starts diurnal temporal profiles are applied nationally and are based on the default starts-hour-fraction tables from MOVES.

3.3.9 Nonroad mobile temporal allocation (nonroad)

For nonroad mobile sources, temporal allocation is performed differently for different SCCs. Beginning with the final 2011 platform, improvements to temporal allocation of nonroad mobile sources were made to make the temporal profiles more realistically reflect real-world practices. The specific updates were made for agricultural sources (e.g., tractors), construction, and commercial residential lawn and garden sources.

Figure 3-19 shows two previously existing temporal profiles (9 and 18) and a newer temporal profile (19) which has lower emissions on weekends. In this platform, construction and commercial lawn and garden sources use the new profile 19 which has lower emissions on weekends. Residential lawn and garden sources continue to use profile 9 and agricultural sources continue to use profile 18.

Figure 3-19. Example Nonroad Day-of-week Temporal Profiles

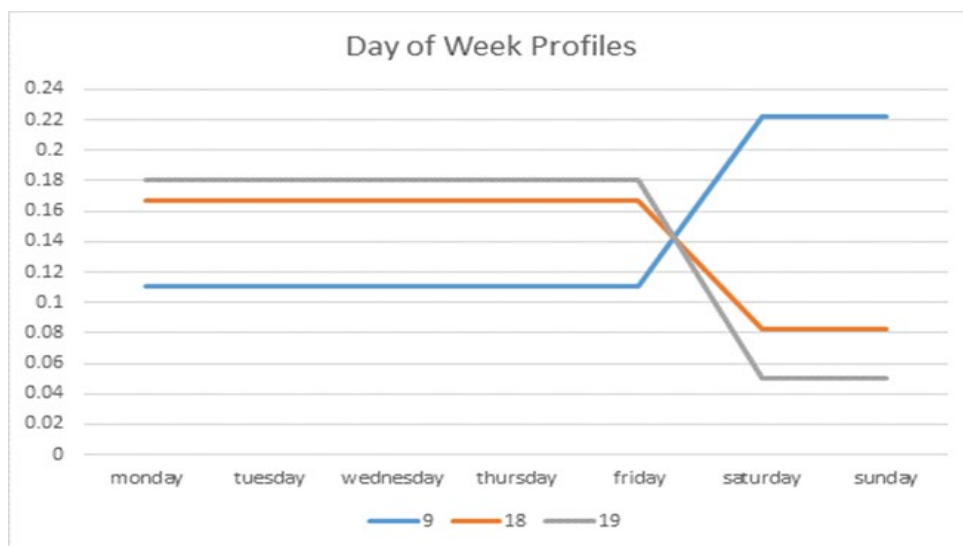
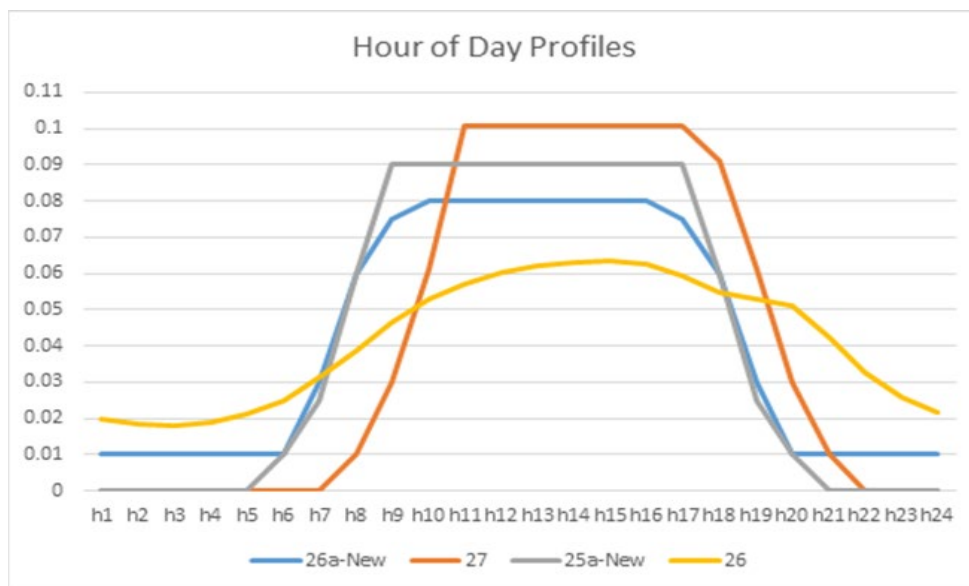


Figure 3-20 shows the previously existing temporal profiles 26 and 27 along with newer temporal profiles (25a and 26a) which have lower emissions overnight. In this platform, construction sources use profile 26a. Commercial lawn and garden and agriculture sources use the profiles 26a and 25a, respectively. Residential lawn and garden sources use profile 27.

Figure 3-20. Example Nonroad Diurnal Temporal Profiles



For the nonroad sector, while the NEI only stores the annual totals, the modeling platform uses monthly inventories from output from MOVES. For California, CARB’s annual inventory was temporalized to monthly using monthly temporal profiles applied in SMOKE by SCC.

3.3.10 Additional sector specific details (afdust, beis, cmv, rail, nonpt, ptfire-rx, ptfire-wild)

For the afdust sector, meteorology is not used in the development of the temporal profiles, but it is used to reduce the total emissions based on meteorological conditions. These adjustments are applied through sector-specific scripts, beginning with the application of land use-based gridded transport fractions and then subsequent zero-outs for hours during which precipitation occurs or there is snow cover on the ground. The land use data used to reduce the NEI emissions explain the amount of emissions that are subject to transport. This methodology is discussed in (Pouliot et al., 2010), and in “Fugitive Dust Modeling for the 2008 Emissions Modeling Platform” (Adelman, 2012). The precipitation adjustment is applied to remove all emissions for hours where measurable rain occurs, or where there is snow cover. Therefore, the afdust emissions vary day-to-day based on the precipitation and/or snow cover for each grid cell and hour. Both the transport fraction and meteorological adjustments are based on the gridded resolution of the platform; therefore, somewhat different emissions will result from different grid resolutions. Application of the transport fraction and meteorological adjustments prevents the overestimation of fugitive dust impacts in the grid modeling as compared to ambient samples.

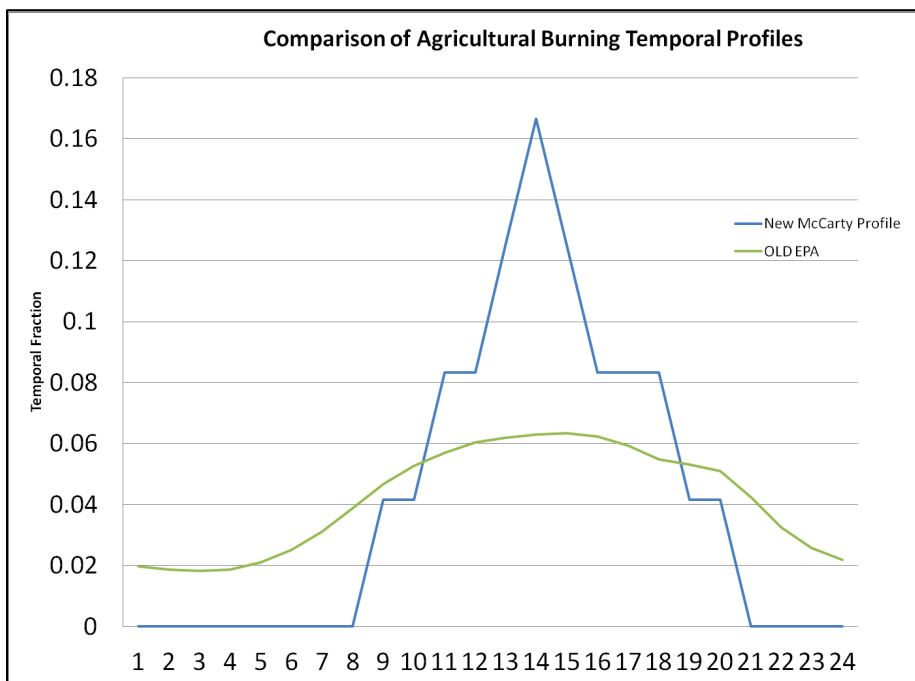
Biogenic emissions from the BEIS model vary each day of the year because they are developed using meteorological data including temperature, surface pressure, and radiation/cloud data. The emissions are computed using appropriate emission factors according to the vegetation in each model grid cell, while taking the meteorological data into account.

For the cmv sectors, most areas use hourly emission inventories derived from the 5-minute AIS data. In some areas where AIS data are not available, such as in Canada between the St. Lawrence Seaway and the Great Lakes and in the southern Caribbean, the flat temporal profiles are used for hourly and day-of-week values. Most regions without AIS data also use a flat monthly profile, with some offshore areas using an average monthly profile derived from the 2008 ECA inventory monthly values. These areas without AIS data also use flat day of week and hour of day profiles.

For the rail sector, monthly profiles from the 2016 platform were used. Monthly temporal allocation for rail freight emissions is based on AAR Rail Traffic Data, Total Carloads and Intermodal, for 2016. For passenger trains, monthly temporal allocation is flat for all months. Rail passenger miles data is available by month but it is not known how closely rail emissions track with passenger activity since passenger trains run on a fixed schedule regardless of how many passengers are aboard, and so a flat profile is chosen for passenger trains. Rail emissions are allocated with flat day of week profiles, and most emissions are allocated with flat hourly profiles.

For the ptagfire sector, the inventories are in the daily point fire format FF10 PTDAY. The diurnal temporal profile for ag fires reflects the fact that burning occurs during the daylight hours - see Figure 3-21 (McCarty et al., 2009). This puts most of the emissions during the work-day and suppresses the emissions during the middle of the night.

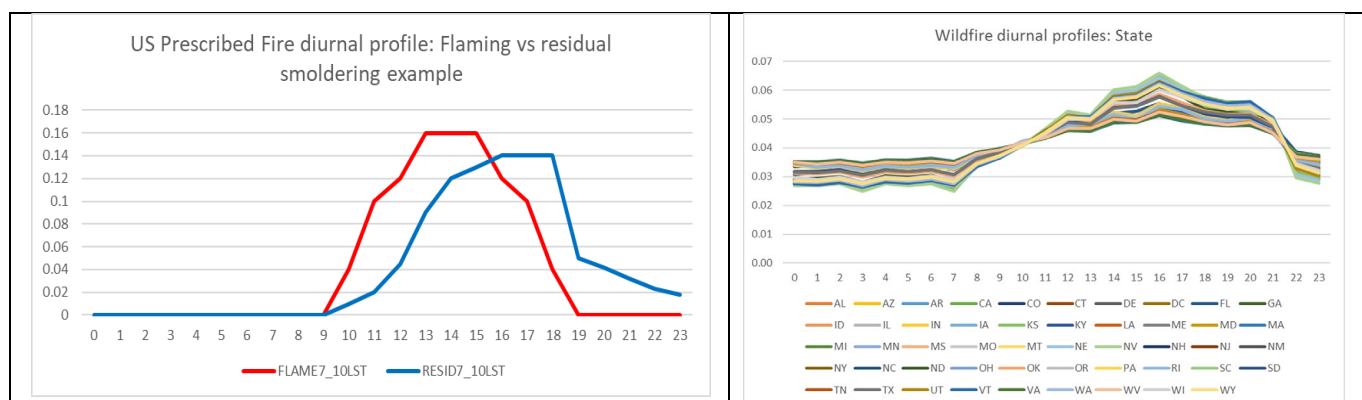
Figure 3-21. Agricultural burning diurnal temporal profile



Industrial processes that are not likely to shut down on Sundays, such as those at cement plants, use profiles that include emissions on Sundays, while those that would shut down on Sundays use profiles that reflect Sunday shutdowns.

For the ptfire sectors, the inventories are in the daily point fire format FF10 PTDAY, so temporal profiles are only used to go from day-specific to hourly emissions. Separate hourly profiles for prescribed and wildfires were used. For ptfire, state-specific hourly profiles were used, with distinct profiles for prescribed fires and wildfires. Figure 3-22 below shows the profiles used for each state for the platform. The wildfire diurnal profiles are similar but vary according to the average meteorological conditions in each state. For all agricultural burning, the diurnal temporal profile used reflected the fact that burning occurs during the daylight. This puts most of the emissions during the workday and suppresses the emissions during the middle of the night. This diurnal profile was used for each day of the week for all agricultural burning emissions in all states.

Figure 3-22. Prescribed and Wildfire diurnal temporal profiles



3.4 Spatial Allocation

The methods used to perform spatial allocation are summarized in this section. For the modeling platform, spatial factors are typically applied by county and SCC. Spatial allocation was performed for each of the modeling grids shown in Section 3.1. To accomplish this, SMOKE used national 12-km spatial surrogates and a SMOKE area-to-point data file. For the U.S., the EPA updated surrogates to use circa 2020 data. The U.S., Mexican, and Canadian 12-km surrogates cover the entire CONUS domain 12US1 shown in Figure 3-1. While highlights of information are provided below, the file *Surrogate_specifications_2021_platform_US_Can_Mex.xlsx* documents the complete configuration for generating the surrogates and can be referenced for more details.

3.4.1 Spatial Surrogates for U.S. emissions

There are more than 90 spatial surrogates available for spatially allocating U.S. county-level emissions to the 12-km grid cells used by the air quality model. As described in Section 3.4.2, an area-to-point approach overrides the use of surrogates for airport refueling sources.

The surrogates for the platform are based on a variety of geospatial data sources, including the American Community Survey (ACS) for census-related data, the National Land Cover Database (NLCD)

Onroad surrogates are based on average annual daily traffic counts (AADT) from the highway monitoring performance system (HPMS).

U.S. Surrogate datasets used for this platform include:

- County boundaries used for all surrogates were use the 2020 TIGER boundaries.
- Oil and gas surrogates represent activity from the year 2021.
- ACS-based surrogates use the 2020 ACS.
- NLCD-based surrogates use NLCD 2019.
- Animal specific livestock waste surrogates were derived from National Pollutant Discharge Elimination System (NPDES) animal operation water permits and Food and Agriculture Organization (FAO) gridded livestock count data.
- Surrogates for fuel stations, asphalt surfaces, and unpaved roads are based on data from the OpenStreetMap database.
- Gravel and lead mines use separate surrogates based on the more general United States Geological Survey (USGS) mining surrogate.
- Residential wood combustion surrogates are based on ACS data.

When developing modeling platforms, EPA routinely updates surrogates to utilize updated versions of the underlying surrogate databases or to use a different source of data when it is deemed more representative for a particular source category. In the 2020 platform, NLCD-based surrogates were updated from using the 2011 National Land Cover Database (NLCD) to use the 2019 National Land Cover Database. During these updates, EPA also examined the Residential Wood Combustion (RWC) surrogates that were based on the NLCD. This was done to see if there are other sources of spatial data that could improve the geographic representation of the RWC sector when disaggregating the county-level emissions provided by the emissions inventory to grid cells. For the RWC sector prior to the 2020 platform, the spatial surrogate used was #300 computed from “NLCD Low Intensity development” (i.e., land areas with 20-49% impervious surface). This surrogate was initially selected for RWC to capture geographic areas where there may be houses but generally in less developed spaces. However, this surrogate does not differentiate by development or structure type. The result is that RWC emissions can end up concentrated around roads, commercial, and other low to moderately developed grid cells.

In the 2020 platform, housing data provided by the American Community Survey (ACS) were used. The particular attributes used are: single family detached, single family attached, dual family and mobile home and combinations of these, depending on the particular RWC specific source category. Using types of housing seemed more reflective of where RWC emissions would be located. However, a downside of using the ACS housing data is that the census shapes are broad (particularly in rural areas), so the emissions can appear more spread out in some areas than when using the NLCD-based surrogates. When comparing the two approaches for RWC surrogates (NLCD vs ACS), the ACS-based surrogates looked reasonable, and in fact better than the NLCD Low intensity development surrogate. A comparison of the PM_{2.5} emissions gridded with each of these approaches is shown in Figure 3-23 and Figure 3-24. In the future, a goal is to further improve the resolution surrogates, as such, the use of building structure data weighted by the ACS will be examined for future platform updates.

Figure 3-23. 2020 Residential Wood Combustion Emissions using NLCD Low Intensity Surrogate

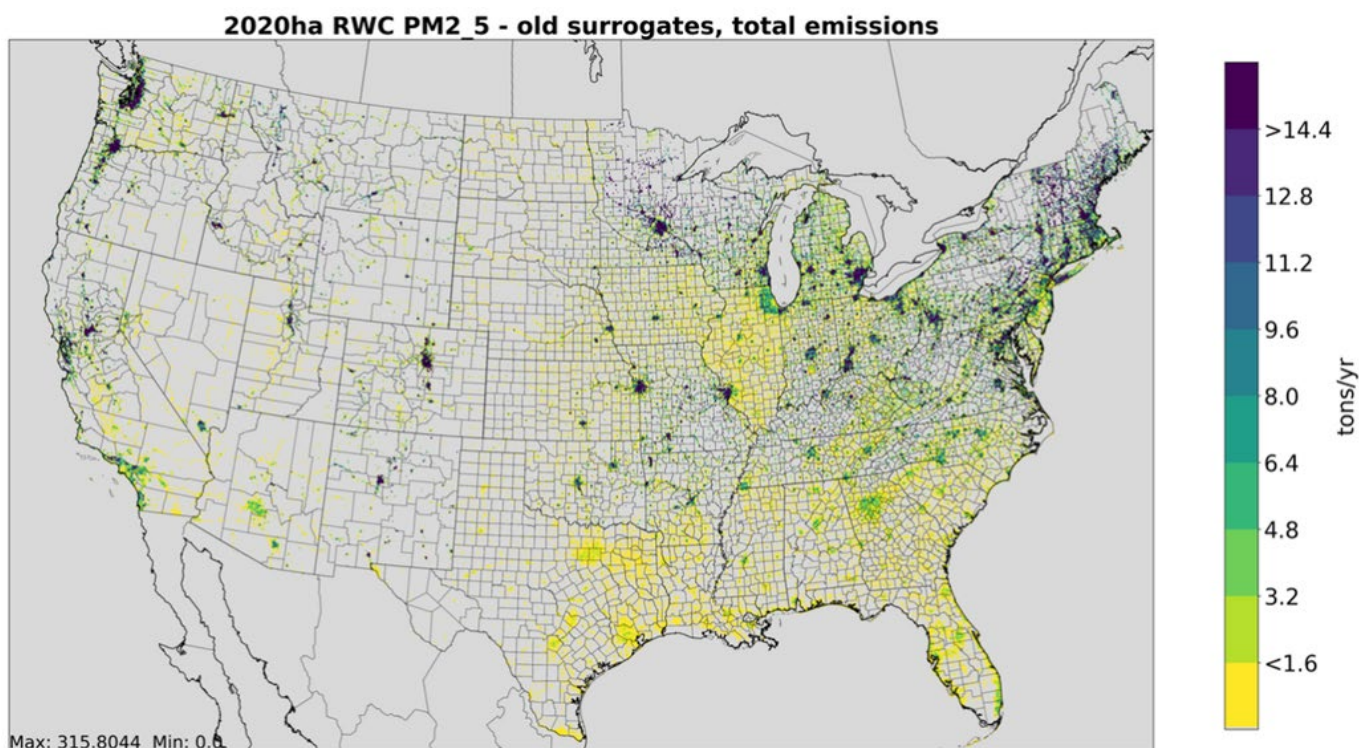
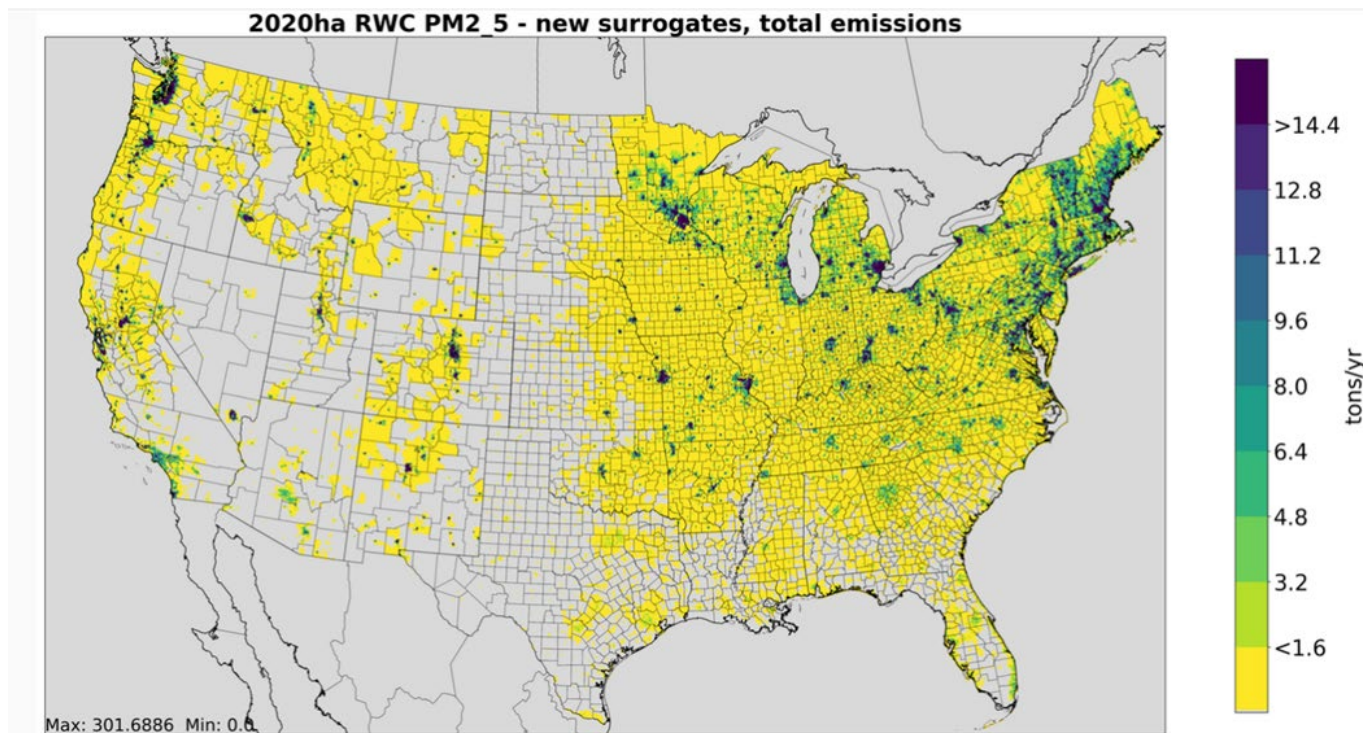


Figure 3-24. 2020 Residential Wood Combustion Emissions using ACS-based Surrogate



Surrogates for the U.S. were generated using the Surrogate Tools DB with the Java-based Surrogate tools used to perform gapfilling and normalization where needed. The tool and documentation for the original Surrogate Tool are available at <https://www.cmascenter.org/sa->

[tools/documentation/4.2/SurrogateToolUserGuide_4_2.pdf](#), and the tool and documentation for the Surrogate Tools DB is available from https://www.cmascenter.org/surrogate_tools_db/. Table 3-17 lists the codes and descriptions of the surrogates. Surrogate names and codes listed in *italics* are not directly assigned to any sources in the platform, but they are sometimes used to gapfill other surrogates. When the source data for a surrogate have no values for a particular county, gap filling is used to provide values for the spatial surrogate in those counties to ensure that no emissions are dropped when the spatial surrogates are applied to the emission inventories. The Shapefiles used to develop the US surrogates along with the attributes and filters used are shown in Table 3-18.

Table 3-17. U.S. Surrogates available for the 2021 modeling platforms

Code	Surrogate Description	Code	Surrogate Description
N/A	Area-to-point approach (see 3.6.2)	650	Refineries and Tank Farms
100	Population	<i>669</i>	<i>All Abandoned Wells</i>
<i>110</i>	<i>Housing</i>	<i>6691</i>	<i>All Abandoned Oil Wells</i>
135	Detached Housing	<i>6692</i>	<i>All Abandoned Gas Wells</i>
136	Single and Dual Unit Housing	<i>6693</i>	<i>All Abandoned CBM Wells</i>
137	Single + Dual Unit + Manufactured Housing	6694	All Abandoned Oil Wells - Plugged
150	Residential Heating - Natural Gas	6695	All Abandoned Gas Wells - Plugged
170	Residential Heating - Distillate Oil	<i>6696</i>	<i>All Abandoned CBM Wells - Plugged</i>
180	Residential Heating - Coal	6697	All Abandoned Oil Wells - Unplugged
190	Residential Heating - LP Gas	6698	All Abandoned Gas Wells - Unplugged
205	Extended Idle Locations	670	Spud Count - CBM Wells
239	Total Road AADT	671	Spud Count - Gas Wells
240	Total Road Miles	<i>672</i>	<i>Gas Production at Oil Wells</i>
242	All Restricted AADT	<i>673</i>	<i>Oil Production at CBM Wells</i>
244	All Unrestricted AADT	674	Unconventional Well Completion Counts
<i>258</i>	<i>Intercity Bus Terminals</i>	676	Well Count - All Producing
259	Transit Bus Terminals	<i>677</i>	<i>Well Count - All Exploratory</i>
<i>260</i>	<i>Total Railroad Miles</i>	678	Completions at Gas Wells
261	NTAD Total Railroad Density	679	Completions at CBM Wells
271	NTAD Class 1 2 3 Railroad Density	681	Spud Count - Oil Wells
300	NLCD Low Intensity Development	683	Produced Water at All Wells
304	NLCD Open + Low	6831	Produced Water at CBM Wells
305	NLCD Low + Med	6832	Produced Water at Gas Wells
306	NLCD Med + High	6833	Produced Water at Oil Wells
307	NLCD All Development	685	Completions at Oil Wells
308	NLCD Low + Med + High	<i>686</i>	<i>Completions at All Wells</i>
309	NLCD Open + Low + Med	687	Feet Drilled at All Wells
310	NLCD Total Agriculture	689	Gas Produced - Total
<i>318</i>	<i>NLCD Pasture Land</i>	691	Well Counts - CBM Wells
319	NLCD Crop Land	692	Spud Count - All Wells

320	NLCD Forest Land	693	Well Count - All Wells
321	NLCD Recreational Land	694	Oil Production at Oil Wells
340	NLCD Land	695	Well Count - Oil Wells
350	NLCD Water	696	Gas Production at Gas Wells
401	FAO 2010 Cattle	697	Oil Production at Gas Wells
4011	FAO 2010 Large Cattle Operations	698	Well Count - Gas Wells
4012	NPDES 2020 Beef Cattle	699	Gas Production at CBM Wells
4013	NPDES 2020 Dairy Cattle	711	Airport Areas
402	FAO 2010 Pig	801	Port Areas
4021	NPDES 2020 Swine	850	Golf Courses
403	FAO 2010 Chicken	860	Mines
4031	NPDES 2020 Chicken	861	Sand and Gravel Mines
404	FAO 2010 Goat	862	Lead Mines
4041	NPDES 2020 Goat	863	Crushed Stone Mines
405	FAO 2010 Horse	900	OSM Fuel
406	FAO 2010 Sheep	901	OSM Asphalt Surfaces
4071	NPDES 2020 Turkey	902	OSM Unpaved Roads
508	Public Schools		

Table 3-18. Shapefiles used to develop U.S. Surrogates

Code	Surrogate	Weight Shapefile	Weight Attribute	Filter Function
100	Population	ACS_2020_5YR_BG_pop_hu	POP2020	
110	Housing	ACS_2020_5YR_BG_pop_hu	HU2020	
135	Detached Housing	ACS_2020_5YR_BG_pop_hu	detachedh	
136	Single and Dual Unit Housing	ACS_2020_5YR_BG_pop_hu	lttriunit	
137	Single + Dual Unit + Manufactured Housing	ACS_2020_5YR_BG_pop_hu_mobile	sngdlmobl	
150	Residential Heating - Natural Gas	ACS_2020_5YR_BG_pop_hu	UTIL_GAS	
170	Residential Heating - Distillate Oil	ACS_2020_5YR_BG_pop_hu	FUEL_OIL	
180	Residential Heating - Coal	ACS_2020_5YR_BG_pop_hu	COAL	
190	Residential Heating - LP Gas	ACS_2020_5YR_BG_pop_hu	LP_GAS	
205	Extended Idle Locations	pil_2019_06_24	rev_truck	rev_truck>0
239	Total Road AADT	hpms2017_v3_04052020	aadt	moves2014 IN ('02','03','04','05')
240	Total Road Miles	hpms2017_v3_04052020	NONE	moves2014 IN ('02','03','04','05')
242	All Restricted AADT	hpms2017_v3_04052020	aadt	moves2014 IN ('02','04')

Code	Surrogate	Weight Shapefile	Weight Attribute	Filter Function
244	All Unrestricted AADT	hpms2017_v3_04052020	aadt	moves2014 IN ('03','05')
259	Transit Bus Terminals	ntad_2016_ipcd	NONE	bus_t=1
260	Total Railroad Miles	tiger_2014_rail	NONE	
261	NTAD Total Railroad Density	ntad_2014_rail_fixed	dens	RAILTYPE IN (1,2,3)
271	NTAD Class 1 2 3 Railroad Density	ntad_2014_rail_fixed	dens	RAILTYPE=1
300	NLCD Low Intensity Development	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE=22
304	NLCD Open + Low	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (21,22)
305	NLCD Low + Med	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (22,23)
306	NLCD Med + High	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (23,24)
307	NLCD All Development	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (21,22,23,24)
308	NLCD Low + Med + High	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (22,23,24)
309	NLCD Open + Low + Med	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (21,22,23)
310	NLCD Total Agriculture	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (81,82)
318	NLCD Pasture Land	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE=81
319	NLCD Crop Land	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE=82
320	NLCD Forest Land	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (41,42,43)
321	NLCD Recreational Land	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE IN (21,31,41,42,43,52,71)
340	NLCD Land	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE != 11
350	NLCD Water	nlcd_2019_land_cover_l48_20210604_500m_ll	NONE	GRIDCODE=11
401	FAO 2010 Cattle	fao_Cattle_2010_Da_nlcdproj_masked	DN	
4011	FAO 2010 Large Cattle Operations	fao_LargeCattle_2010_Da_nlcdproj_masked	DN	
4012	NPDES 2020 Beef Cattle	livestock_npdes_state_permits_subset	Population	Animal = 'Beef'
4013	NPDES 2020 Dairy Cattle	livestock_npdes_state_permits_subset	Population	Animal = 'Dairy'
402	FAO 2010 Pig	fao_Pig_2010_Da_nlcdproj_masked	DN	
4021	NPDES 2020 Swine	livestock_npdes_state_permits_subset	Population	Animal = 'Swine'

Code	Surrogate	Weight Shapefile	Weight Attribute	Filter Function
403	FAO 2010 Chicken	fao_Chicken_2010_Da_nlcdproj_masked	DN	
4031	NPDES 2020 Chicken	livestock_npdes_state_permits_subset	Population	Animal = 'Chicken'
404	FAO 2010 Goat	fao_Goat_2010_Da_nlcdproj_masked	DN	
4041	NPDES 2020 Goat	livestock_npdes_state_permits_subset	Population	Animal = 'Goat'
405	FAO 2010 Horse	fao_Horse_2010_Da_nlcdproj_masked	DN	
406	FAO 2010 Sheep	fao_Sheep_2010_Da_nlcdproj_masked	DN	
4071	NPDES 2020 Turkey	livestock_npdes_state_permits_subset	Population	Animal = 'Turkey'
508	Public Schools	public_schools_2018_2019	TOTAL	
650	Refineries and Tank Farms	eia_2015_us_oil	NONE	
669	All Abandoned Wells	ABANDONED_WELLS_ALL_COUNTS_669_2021	ACTIVITY	
6696	All Abandoned CBM Wells – Plugged	ABANDONED_WELLS_CBM_PLUGGED_6696_2021	ACTIVITY	
6693	All Abandoned CBM Wells	ABANDONED_WELLS_CBM_PLUGGED_UNPLUGGED_6693_2021	ACTIVITY	
6695	All Abandoned Gas Wells – Plugged	ABANDONED_WELLS_GAS_PLUGGED_6695_2021	ACTIVITY	
6692	All Abandoned Gas Wells	ABANDONED_WELLS_GAS_PLUGGED_UNPLUGGED_6692_2021	ACTIVITY	
6698	All Abandoned Gas Wells – Unplugged	ABANDONED_WELLS_GAS_UNPLUGGED_6698_2021	ACTIVITY	
6694	All Abandoned Oil Wells – Plugged	ABANDONED_WELLS_OIL_PLUGGED_6694_2021	ACTIVITY	
6691	All Abandoned Oil Wells	ABANDONED_WELLS_OIL_PLUGGED_UNPLUGGED_6691_2021	ACTIVITY	
6697	All Abandoned Oil Wells – Unplugged	ABANDONED_WELLS_OIL_UNPLUGGED_6697_2021	ACTIVITY	
670	Spud Count - CBM Wells	SPUD_CBM_670_2021	ACTIVITY	
671	Spud Count - Gas Wells	SPUD_GAS_671_2021	ACTIVITY	
672	Gas Production at Oil Wells	ASSOCIATED_GAS_PRODUCTION_672_2021	ACTIVITY	
673	Oil Production at CBM Wells	CONDENSATE_CBM_PRODUCTION_673_2021	ACTIVITY	
674	Unconventional Well Completion Counts	COMPLETIONS_UNCONVENTIONAL_674_2021	ACTIVITY	
676	Well Count - All Producing	TOTAL_PROD_WELL_676_2021	ACTIVITY	
677	Well Count - All Exploratory	TOTAL_EXPL_WELL_677_2021	ACTIVITY	
678	Completions at Gas Wells	COMPLETIONS_GAS_678_2021	ACTIVITY	
679	Completions at CBM Wells	COMPLETIONS_CBM_679_2021	ACTIVITY	
681	Spud Count - Oil Wells	SPUD_OIL_681_2021	ACTIVITY	
683	Produced Water at All Wells	PRODUCED_WATER_ALL_683_2021	ACTIVITY	

Code	Surrogate	Weight Shapefile	Weight Attribute	Filter Function
6831	Produced Water at CBM Wells	PRODUCED_WATER_CBM_6831_2021	ACTIVITY	
6832	Produced Water at Gas Wells	PRODUCED_WATER_GAS_6832_2021	ACTIVITY	
6833	Produced Water at Oil Wells	PRODUCED_WATER_OIL_6833_2021	ACTIVITY	
685	Completions at Oil Wells	COMPLETIONS_OIL_685_2021	ACTIVITY	
686	Completions at All Wells	COMPLETIONS_ALL_686_2021	ACTIVITY	
687	Feet Drilled at All Wells	FEET_DRILLED_687_2021	ACTIVITY	
689	Gas Produced - Total	TOTAL_GAS_PRODUCTION_689_2021	ACTIVITY	
691	Well Counts - CBM Wells	CBM_WELLS_691_2021	ACTIVITY	
692	Spud Count - All Wells	SPUD_ALL_692_2021	ACTIVITY	
693	Well Count - All Wells	TOTAL_WELL_693_2021	ACTIVITY	
694	Oil Production at Oil Wells	OIL_PRODUCTION_694_2021	ACTIVITY	
695	Well Count - Oil Wells	OIL_WELLS_695_2021	ACTIVITY	
696	Gas Production at Gas Wells	GAS_PRODUCTION_696_2021	ACTIVITY	
697	Oil Production at Gas Wells	CONDENSATE_GAS_PRODUCTION_697_2021	ACTIVITY	
698	Well Count - Gas Wells	GAS_WELLS_698_2021	ACTIVITY	
699	Gas Production at CBM Wells	CBM_PRODUCTION_699_2021	ACTIVITY	
711	Airport Areas	airport_area	area	
801	Port Areas	Ports_2014NEI	area_sqmi	
850	Golf Courses	usa_golf_courses_2019_10	NONE	
860	Mines	usgs_mrds_active_mines	NONE	
861	Sand and Gravel Mines	usgs_mrds_active_mines	NONE	CAT='Gravel'
862	Lead Mines	usgs_mrds_active_mines	NONE	CAT='Lead'
863	Crushed Stone Mines	usgs_mrds_active_mines	NONE	CAT='Stone'
900	OSM Fuel	osm_fuel_points_us_mar2023	NONE	
901	OSM Asphalt Surfaces	osm_asphalt_surfaces_us_mar2023	NONE	
902	OSM Unpaved Roads	osm_unpaved_roads_us_mar2023	NONE	

The 'Data Shapefile' used for all of the U.S. surrogates except for those based on HPMS data is cb_2020_us_county_500k, while the HPMS-based surrogates use hpms2017_v3_04052020. Similarly, most surrogates use the GEOID as the Data attribute while the HPMS surrogates use FIPS. The gapfilling configuration for the surrogates is shown in Table 3-19. If there are no entries for a county for the primary surrogate, the values for the county from the secondary surrogate are used. If there are also no entries for the secondary surrogate, the values for the tertiary surrogate are used, with the quarternary surrogate being the final fallback. Typically, only surrogates that should have values for all counties are selected as the quarternary surrogate. This process is used to limit any emissions that could be dropped if there are emissions in the inventory in a county for which the primary surrogate does not have values. It is important to note that once gapfilling is performed, SMOKE does not know that emissions for that

county were from a secondary, tertiary or quarternary surrogate and any reports will assign the emissions in gapfilled counties to the primary surrogate.

Table 3-19. Surrogates used to gapfill U.S. Surrogates

SURROGATE CODE	SURROGATE	SECONDARY SURROGATE	TERTIARY SURROGATE	QUARTERNARY SURROGATE
100	Population			
110	Housing	Population		
135	Detached Housing	NLCD Low Intensity Development		
136	Single and Dual Unit Housing	NLCD Low Intensity Development		
137	Single + Dual Unit + Manufactured Housing	NLCD Low Intensity Development	NLCD Land	
150	Residential Heating - Natural Gas	Population		
170	Residential Heating - Distillate Oil	Housing		
180	Residential Heating – Coal	Housing		
190	Residential Heating - LP Gas	Housing		
205	Extended Idle Locations	Total Road Miles		
239	Total Road AADT	Total Road Miles		
240	Total Road Miles			
242	All Restricted AADT	Total Road Miles		
244	All Unrestricted AADT	Total Road Miles		
259	Transit Bus Terminals	Population	NLCD Land	
260	Total Railroad Miles	Total Road Miles	Population	
261	NTAD Total Railroad Density	Total Railroad Miles	Total Road Miles	Population
271	NTAD Class 1 2 3 Railroad Density	NTAD Total Railroad Density	Total Railroad Miles	Total Road Miles
300	NLCD Low Intensity Development	Housing	Population	NLCD Land
304	NLCD Open + Low	Housing	Population	NLCD Land
305	NLCD Low + Med	Housing	Population	NLCD Land
306	NLCD Med + High	Housing	Population	NLCD Land
307	NLCD All Development	Housing	Population	NLCD Land
308	NLCD Low + Med + High	Housing	Population	NLCD Land
309	NLCD Open + Low + Med	Housing	Population	NLCD Land
310	NLCD Total Agriculture	NLCD Open + Low	NLCD Land	
318	NLCD Pasture Land	Housing	NLCD Land	
319	NLCD Crop Land	Housing	NLCD Land	
320	NLCD Forest Land	Housing	NLCD Land	
321	NLCD Recreational Land	Housing	NLCD Land	
340	NLCD Land			
350	NLCD Water			
401	FAO 2010 Cattle	NLCD Total Agriculture	NLCD Open + Low	

SURROGATE CODE	SURROGATE	SECONDARY SURROGATE	TERTIARY SURROGATE	QUARTERNARY SURROGATE
4011	FAO 2010 Large Cattle Operations	FAO 2010 Cattle	NLCD Total Agriculture	NLCD Open + Low
4012	NPDES 2020 Beef Cattle	FAO 2010 Cattle	NLCD Total Agriculture	NLCD Open + Low
4013	NPDES 2020 Dairy Cattle	FAO 2010 Large Cattle Operations	NLCD Total Agriculture	NLCD Open + Low
402	FAO 2010 Pig	NLCD Total Agriculture	NLCD Open + Low	
4021	NPDES 2020 Swine	FAO 2010 Pig	NLCD Total Agriculture	NLCD Open + Low
403	FAO 2010 Chicken	NLCD Total Agriculture	NLCD Open + Low	
4031	NPDES 2020 Chicken	FAO 2010 Chicken	NLCD Total Agriculture	NLCD Open + Low
404	FAO 2010 Goat	NLCD Total Agriculture	NLCD Open + Low	
4041	NPDES 2020 Goat	FAO 2010 Goat	NLCD Total Agriculture	NLCD Open + Low
405	FAO 2010 Horse	NLCD Total Agriculture	NLCD Open + Low	
406	FAO 2010 Sheep	NLCD Total Agriculture	NLCD Open + Low	
4071	NPDES 2020 Turkey	NLCD Total Agriculture	NLCD Open + Low	
508	Public Schools	Population	NLCD Land	
650	Refineries and Tank Farms	NLCD Low + Med	Population	NLCD Land
669	All Abandoned Wells	Well Count - All Wells	NLCD Open + Low	
6696	All Abandoned CBM Wells - Plugged	All Abandoned CBM Wells	Well Count - All Wells	NLCD Open + Low
6693	All Abandoned CBM Wells	Well Count - All Wells	NLCD Open + Low	
6695	All Abandoned Gas Wells - Plugged	All Abandoned Gas Wells	Well Count - All Wells	NLCD Open + Low
6692	All Abandoned Gas Wells	Well Count - All Wells	NLCD Open + Low	
6698	All Abandoned Gas Wells - Unplugged	All Abandoned Gas Wells	Well Count - All Wells	NLCD Open + Low
6694	All Abandoned Oil Wells - Plugged	All Abandoned Oil Wells	Well Count - All Wells	NLCD Open + Low
6691	All Abandoned Oil Wells	Well Count - All Wells	NLCD Open + Low	
6697	All Abandoned Oil Wells - Unplugged	All Abandoned Oil Wells	Well Count - All Wells	NLCD Open + Low
670	Spud Count - CBM Wells	Spud Count - All Wells	Well Count - All Wells	
671	Spud Count - Gas Wells	Well Count - Gas Wells	Well Count - All Wells	
672	Gas Production at Oil Wells	NLCD Open + Low	Well Count - Oil Wells	Well Count - All Wells
673	Oil Production at CBM Wells	Well Count - CBM Wells	Well Count - All Wells	NLCD Open + Low
674	Unconventional Well Completion Counts	Completions at All Wells	Well Count - All Wells	NLCD Open + Low
676	Well Count - All Producing	Well Count - All Wells	NLCD Open + Low	
677	Well Count - All Exploratory	Well Count - All Wells	NLCD Open + Low	
678	Completions at Gas Wells	Spud Count - All Wells	Well Count - All Wells	NLCD Open + Low

SURROGATE CODE	SURROGATE	SECONDARY SURROGATE	TERTIARY SURROGATE	QUARTERNARY SURROGATE
679	Completions at CBM Wells	Spud Count - All Wells	Well Count - All Wells	NLCD Open + Low
681	Spud Count - Oil Wells	Well Count - Oil Wells	Well Count - All Wells	NLCD Open + Low
683	Produced Water at All Wells	Completions at All Wells	Well Count - All Wells	NLCD Open + Low
6831	Produced Water at CBM Wells	Well Counts - CBM Wells	Well Count - All Wells	NLCD Open + Low
6832	Produced Water at Gas Wells	Well Count - Gas Wells	Well Count - All Wells	NLCD Open + Low
6833	Produced Water at Oil Wells	Well Count - Oil Wells	Well Count - All Wells	NLCD Open + Low
685	Completions at Oil Wells	Spud Count - All Wells	Well Count - All Wells	NLCD Open + Low
686	Completions at All Wells	Well Count - All Exploratory	Well Count - All Wells	NLCD Open + Low
687	Feet Drilled at All Wells	Well Count - All Exploratory	Well Count - All Wells	NLCD Open + Low
689	Gas Produced - Total	Well Count - All Wells	NLCD Open + Low	
691	Well Counts - CBM Wells	Completions at CBM Wells	Well Count - All Wells	NLCD Open + Low
692	Spud Count - All Wells	Completions at All Wells	Well Count - All Wells	NLCD Open + Low
693	Well Count - All Wells	NLCD Open + Low		
694	Oil Production at Oil Wells	Completions at Oil Wells	Well Count - All Wells	NLCD Open + Low
695	Well Count - Oil Wells	Completions at Oil Wells	Well Count - All Wells	NLCD Open + Low
696	Gas Production at Gas Wells	Completions at Gas Wells	Well Count - All Wells	NLCD Open + Low
697	Oil Production at Gas Wells	Well Count - Gas Wells	Well Count - All Wells	NLCD Open + Low
698	Well Count - Gas Wells	Completions at Gas Wells	Well Count - All Wells	NLCD Open + Low
699	Gas Production at CBM Wells	Well Counts - CBM Wells	Well Count - All Wells	NLCD Open + Low
711	Airport Areas	Population	NLCD Land	
801	Port Areas	NLCD Water		
850	Golf Courses	Housing	Population	NLCD Land
860	Mines	NLCD Open + Low	NLCD Land	
861	Sand and Gravel Mines	Mines	NLCD Open + Low	NLCD Land
862	Lead Mines	Mines	NLCD Open + Low	NLCD Land
863	Crushed Stone Mines	Mines	NLCD Open + Low	NLCD Land
900	OSM Fuel	Total Road AADT	Total Road Miles	
901	OSM Asphalt Surfaces	NLCD All Development		
902	OSM Unpaved Roads	NLCD Open + Low		

For the onroad sector, the on-network (RPD) emissions were spatially allocated differently from other off-network processes (i.e. RPV, RPP, RPHO, RPS, RPH). Surrogates for on-network processes are based on AADT data and off network processes (including the off-network idling included in RPHO) are based on land use surrogates as shown in Table 3-20. Emissions from the extended (i.e., overnight) idling of trucks were assigned to surrogate 205, which is based on locations of overnight truck parking spaces. The underlying data for this surrogate were updated during the development of the 2016 platforms to include additional data sources and corrections based on comments received and these updates were carried into this platform.

Table 3-20. Off-Network Mobile Source Surrogates

Source type	Source Type name	Surrogate ID	Description
11	Motorcycle	307	NLCD All Development
21	Passenger Car	307	NLCD All Development
31	Passenger Truck	307	NLCD All Development
32	Light Commercial Truck	308	NLCD Low + Med + High
41	Other Bus	306	NLCD Med + High
42	Transit Bus	259	Transit Bus Terminals
43	School Bus	508	Public Schools
51	Refuse Truck	306	NLCD Med + High
52	Single Unit Short-haul Truck	306	NLCD Med + High
53	Single Unit Long-haul Truck	306	NLCD Med + High
54	Motor Home	304	NLCD Open + Low
61	Combination Short-haul Truck	306	NLCD Med + High
62	Combination Long-haul Truck	306	NLCD Med + High

For the oil and gas sources in the np_oilgas sector, the spatial surrogates were updated to those shown in Table 3-21 using 2021 data consistent with what was used to develop the nonpoint oil and gas emissions. The exploration and production of oil and gas have increased in terms of quantities and locations over the last seven years, primarily through the use of new technologies, such as hydraulic fracturing. Census-tract, 2-km, and 4-km sub-county Shapefiles were developed, from which the 2020 oil and gas surrogates were generated. All spatial surrogates for np_oilgas are developed based on known locations of oil and gas activity for year 2021.

The primary activity data source used for the development of the oil and gas spatial surrogates was data from ENVERUS [formerly Drilling Info (DI) Desktop's HPDI] database (ENVERUS, 2023). This database contains well-level location, production, and exploration statistics at the monthly level. Due to a proprietary agreement with ENVERUS, individual well locations and ancillary production cannot be made publicly available, but aggregated statistics are allowed. These data were supplemented with data from state Oil and Gas Commission (OGC) websites (Alaska, Arizona, Idaho, Illinois, Indiana, Kentucky, Louisiana, Michigan, Mississippi, Missouri, Nevada, Oregon, Pennsylvania, and Tennessee). In cases when the desired surrogate parameter was not available (e.g., feet drilled), data for an alternative surrogate parameter (e.g., number of spudded wells) were downloaded and used. Under that methodology, both completion date and date of first production from HPDI were used to identify wells completed during 2020. The spatial surrogates were gapfilled using fallback surrogates as shown in Table 3-19. All gapfilling was performed with the Surrogate Tool.

Table 3-21. Spatial Surrogates for Oil and Gas Sources

Surrogate Code	Surrogate Description
669	<i>All Abandoned Wells</i>
6691	<i>All Abandoned Oil Wells</i>
6692	<i>All Abandoned Gas Wells</i>
6693	<i>All Abandoned CBM Wells</i>
6694	All Abandoned Oil Wells – Plugged
6695	All Abandoned Gas Wells – Plugged
6696	<i>All Abandoned CBM Wells – Plugged</i>
6697	All Abandoned Oil Wells – Unplugged
6698	All Abandoned Gas Wells – Unplugged
670	Spud Count - CBM Wells
671	Spud Count - Gas Wells
672	Gas Production at Oil Wells
673	Oil Production at CBM Wells
674	Unconventional Well Completion Counts
676	Well Count - All Producing
677	Well Count - All Exploratory
678	Completions at Gas Wells
679	Completions at CBM Wells
681	Spud Count - Oil Wells
683	Produced Water at All Wells
685	Completions at Oil Wells
686	Completions at All Wells
687	Feet Drilled at All Wells
689	Gas Produced – Total
691	Well Counts - CBM Wells
692	Spud Count - All Wells
693	Well Count - All Wells
694	Oil Production at Oil Wells
695	Well Count - Oil Wells
696	Gas Production at Gas Wells
697	Oil Production at Gas Wells
698	Well Count - Gas Wells
699	Gas Production at CBM Wells
6831	Produced water at CBM wells
6832	Produced water at gas wells
6833	Produced water at oil wells

Table 3-22 shows the CAP emissions (i.e., NH₃, NO_x, PM_{2.5}, SO₂, and VOC) by sector assigned to each spatial surrogate.

Table 3-22. Selected 2021 CAP emissions by sector for U.S. Surrogates (short tons in 12US1)

Sector	ID	Description	NH3	NOX	PM2_5	SO2	VOC
afdust	240	Total Road Miles	0	0	302,957	0	0
afdust	306	NLCD Med + High	0	0	41,167	0	0
afdust	308	NLCD Low + Med + High	0	0	122,726	0	0
afdust	310	NLCD Total Agriculture	0	0	502,702	0	0
afdust	861	Sand and Gravel Mines	0	0	271	0	0
afdust	863	Crushed Stone Mines	0	0	291	0	0
afdust	902	OSM Unpaved Roads	0	0	945,886	0	0
afdust	4012	NPDES 2020 Beef Cattle	0	0	191,883	0	0
afdust	4013	NPDES 2020 Dairy Cattle	0	0	15,033	0	0
afdust	4021	NPDES 2020 Swine	0	0	658	0	0
afdust	4031	NPDES 2020 Chicken	0	0	5,071	0	0
afdust	4071	NPDES 2020 Turkey	0	0	1,959	0	0
fertilizer	310	NLCD Total Agriculture	1,275,333	0	0	0	0
livestock	405	FAO 2010 Horse	31,969	0	0	0	2,558
livestock	406	FAO 2010 Sheep	18,776	0	0	0	1,502
livestock	4012	NPDES 2020 Beef Cattle	731,315	0	0	0	58,505
livestock	4013	NPDES 2020 Dairy Cattle	583,465	0	0	0	46,677
livestock	4021	NPDES 2020 Swine	875,338	0	0	0	70,027
livestock	4031	NPDES 2020 Chicken	483,224	0	0	0	38,658
livestock	4041	NPDES 2020 Goat	19,231	0	0	0	1,538
livestock	4071	NPDES 2020 Turkey	81,326	0	0	0	6,506
nonpt	100	Population	454	0	0	0	36
nonpt	135	Detatched Housing	0	16,359	81,108	2,724	18,946
nonpt	150	Residential Heating - Natural Gas	44,524	214,626	2,669	1,436	12,680
nonpt	170	Residential Heating - Distillate Oil	1,499	25,521	3,165	624	1,086
nonpt	180	Residential Heating - Coal	0	2	1	7	2
nonpt	190	Residential Heating - LP Gas	127	36,460	150	164	1,435
nonpt	239	Total Road AADT	0	0	0	0	6,536
nonpt	244	All Unrestricted AADT	0	0	0	0	90,591
nonpt	271	NTAD Class 1 2 3 Railroad Density	0	0	0	0	2,074
nonpt	300	NLCD Low Intensity Development	2,860	3,417	17,009	400	26,432
nonpt	306	NLCD Med + High	17,672	242,209	372,991	84,968	131,292
nonpt	307	NLCD All Development	76,463	28,172	126,918	10,917	81,342
nonpt	308	NLCD Low + Med + High	958	156,265	18,261	4,907	9,688
nonpt	310	NLCD Total Agriculture	517	311	504	31	440
nonpt	319	NLCD Crop Land	0	0	95	70	292
nonpt	320	NLCD Forest Land	0	11	31	0	44
nonpt	650	Refineries and Tank Farms	0	0	0	0	90,120

Sector	ID	Description	NH3	NOX	PM2_5	SO2	VOC
nonpt	711	Airport Areas	0	0	0	0	367
nonpt	801	Port Areas	0	0	0	0	2,351
nonpt	900	OSM Fuel	0	0	0	0	209,354
nonpt	4011	FAO 2010 Large Cattle Operations	0	0	0	0	295,993
nonroad	136	Single and Dual Unit Housing	99	14,624	2,927	38	89,912
nonroad	261	NTAD Total Railroad Density	3	1,572	157	1	326
nonroad	304	NLCD Open + Low	6	1,635	147	4	5,729
nonroad	305	NLCD Low + Med	5	852	1,020	2	21,807
nonroad	306	NLCD Med + High	376	157,391	9,005	249	98,471
nonroad	307	NLCD All Development	112	29,271	16,129	44	184,048
nonroad	308	NLCD Low + Med + High	595	221,128	18,228	232	43,333
nonroad	309	NLCD Open + Low + Med	133	21,811	1,303	51	50,334
nonroad	310	NLCD Total Agriculture	357	235,108	16,563	161	25,572
nonroad	320	NLCD Forest Land	15	1,986	407	7	3,472
nonroad	321	NLCD Recreational Land	79	13,265	4,898	28	178,845
nonroad	350	NLCD Water	202	113,666	4,161	94	236,941
nonroad	850	Golf Courses	13	2,119	122	5	5,946
nonroad	860	Mines	2	2,382	221	1	439
np_oilgas	670	Spud Count - CBM Wells	0	0	0	0	19
np_oilgas	671	Spud Count - Gas Wells	0	0	0	0	2,835
np_oilgas	674	Unconventional Well Completion Counts	43	31,853	683	16	1,473
np_oilgas	678	Completions at Gas Wells	0	6,337	133	2,412	22,964
np_oilgas	679	Completions at CBM Wells	0	1	0	37	124
np_oilgas	681	Spud Count - Oil Wells	0	0	0	0	34,357
np_oilgas	683	Produced Water at All Wells	0	38	0	0	719
np_oilgas	685	Completions at Oil Wells	0	249	0	1,057	30,700
np_oilgas	687	Feet Drilled at All Wells	0	161,321	3,701	88	6,197
np_oilgas	689	Gas Produced - Total	0	457	25	2	45,845
np_oilgas	691	Well Counts - CBM Wells	0	18,985	298	7	16,036
np_oilgas	692	Spud Count - All Wells	0	540	31	1	68
np_oilgas	693	Well Count - All Wells	0	0	0	0	2
np_oilgas	694	Oil Production at Oil Wells	0	3,428	5	26,923	829,684
np_oilgas	695	Well Count - Oil Wells	0	160,298	4,293	107,599	630,183
np_oilgas	696	Gas Production at Gas Wells	0	43,089	233	1,147	410,444
np_oilgas	697	Oil Production at Gas Wells	0	247	0	0	41,163
np_oilgas	698	Well Count - Gas Wells	0	301,795	4,476	225	537,243
np_oilgas	699	Gas Production at CBM Wells	0	24	2	0	5,101
np_oilgas	6694	All Abandoned Oil Wells - Plugged	0	0	0	0	148
np_oilgas	6695	All Abandoned Gas Wells - Plugged	0	0	0	0	71
np_oilgas	6697	All Abandoned Oil Wells - Unplugged	0	0	0	0	161,402
np_oilgas	6698	All Abandoned Gas Wells - Unplugged	0	0	0	0	14,988
np_oilgas	6831	Produced water at CBM wells	0	0	0	0	900

Sector	ID	Description	NH3	NOX	PM2_5	SO2	VOC
np_oilgas	6832	Produced water at gas wells	0	0	0	0	15,339
np_oilgas	6833	Produced water at oil wells	0	0	0	0	68,474
np_solvents	100	Population	0	0	0	0	1,422,267
np_solvents	240	Total Road Miles	0	0	0	0	45,681
np_solvents	306	NLCD Med + High	0	0	0	0	468,793
np_solvents	307	NLCD All Development	0	0	0	0	234,418
np_solvents	308	NLCD Low + Med + High	0	0	0	0	34,900
np_solvents	310	NLCD Total Agriculture	0	0	0	0	171,047
np_solvents	901	OSM Asphalt Surface	0	0	0	0	339,778
onroad	205	Extended Idle Locations	0	33,216	311	15	2,940
onroad	242	All Restricted AADT	58,206	781,682	20,026	2,892	118,369
onroad	244	All Unrestricted AADT	118,185	1,156,599	43,773	5,273	323,939
onroad	259	Transit Bus Terminals	36	1,474	32	1	459
onroad	304	NLCD Open + Low	0	476	12	0	2,571
onroad	306	NLCD Med + High	849	94,910	2,481	68	23,317
onroad	307	NLCD All Development	6,409	171,002	7,199	472	539,749
onroad	308	NLCD Low + Med + High	259	17,306	479	26	27,836
onroad	508	Public Schools	10	1,514	63	1	391
rail	261	NTAD Total Railroad Density	13	23,285	629	16	1,066
rail	271	NTAD Class 1 2 3 Railroad Density	283	420,839	10,354	352	17,301
rwc	135	Detached Housing	6,666	8,591	128,302	3,216	120,590
rwc	137	Single + Dual Unit + Manufactured Housing	15,951	36,199	318,692	8,678	332,453

3.4.2 Allocation method for airport-related sources in the U.S.

There are numerous airport-related emission sources in the NEI, such as aircraft, airport ground support equipment, and jet refueling. The modeling platform includes the aircraft and airport ground support equipment emissions as point sources. For the modeling platform, the EPA used the SMOKE “area-to-point” approach for only jet refueling in the nonpt sector. The following SCCs use this approach: 2501080050 and 2501080100 (petroleum storage at airports), and 2810040000 (aircraft/rocket engine firing and testing). The ARTOPNT file that lists the nonpoint sources to locate using point data was unchanged from the 2005-based platform.

3.4.3 Surrogates for Canada and Mexico emission inventories

The surrogates for Canada to spatially allocate the Canadian emissions are based on the 2020 Canadian inventories and associated data. The spatial surrogate data came from ECCC, along with cross references. The shapefiles they provided were used in the Surrogate Tool (previously referenced) to create spatial surrogates. The Canadian surrogates used for this platform are listed in Table 3-23. The Shapefiles used to compute these surrogates and some configuration information are shown in Table 3-24. Note that the name of most Data Shapefiles have been abbreviated to shorten the table. The complete names and additional details on surrogate computation for Canada and Mexico are available in the file *Surrogate_specifications_2021_platform_US_Can_Mex.xlsx* that is posted in the reports folder for this platform. Mexico surrogates were updated for the 2021 EMP.

The data source for the Mexico population surrogate is the INEGI National Geostatistical Framework's Censo de Población y Vivienda 2020 based on the 2020 GPW v4 (see <https://en.www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463807469>). Other data sources used are Sistema Nacional de Informacion Estadistic y Geografica (SNIEG), US Department of Transportation's (DOT) North American Rail Network Lines, and US DOT's Bureau of Transportation Statistics Border Crossing Data. The Shapefiles and some configuration information used to develop the Mexico surrogates are shown in Table 3-25. The Data Shapefile for all Mexico surrogates is areas_geoestadisticas_municipales_II and the Data Attribute is FIPS. Most of the CAP emissions allocated to the Mexico and Canada surrogates are shown in Table 3-26.

Table 3-23. Canadian Spatial Surrogates

Code	Canadian Surrogate Description	Code	Description
100	Population	925	Manufacturing and Assembly
101	total dwelling	926	Distribution and Retail (no petroleum)
102	<i>urban dwelling</i>	927	Commercial Services
103	<i>rural dwelling</i>	933	Rail-Passenger
104	capped total dwelling	934	Rail-Freight
105	<i>capped meat cooking dwelling</i>	935	Rail-Yard
106	ALL_INDUST	940	PAVED ROADS NEW
113	Forestry and logging	945	<i>Commercial Marine Vessels</i>
116	<i>Total Resources</i>	946	Construction and mining
200	Urban Primary Road Miles	948	<i>Forest</i>
210	Rural Primary Road Miles	949	<i>Combination of Dwelling</i>
211	<i>Oil and Gas Extraction</i>	951	Wood Consumption Percentage
212	Mining except oil and gas	952	<i>Residential Fuel Wood Combustion (PIRD)</i>
220	Urban Secondary Road Miles	955	UNPAVED_ROADS_AND_TRAILS
221	Total Mining	960	<i>TOTBEEF</i>
222	Utilities	961	80110_Broilers
230	Rural Secondary Road Miles	962	80111_Cattle_dairy_and_Heifer
233	<i>Total Land Development</i>	963	80112_Cattle_non-Dairy
240	capped population	964	80113_Laying_hens_and_Pullets
308	Food manufacturing	965	80114_Horses
321	Wood product manufacturing	966	80115_Sheep_and_Lamb
323	Printing and related support activities	967	80116_Swine
324	Petroleum and coal products manufacturing	968	80117_Turkeys
326	Plastics and rubber products manufacturing	969	80118_Goat
327	Non-metallic mineral product manufacturing	970	<i>TOTPOUL</i>
331	Primary Metal Manufacturing	971	80119_Buffalo
340	<i>Construction - Oil and Gas</i>	972	80120_Llama_and_Alpacas
350	<i>Water</i>	973	80121_Deer
412	Petroleum product wholesaler-distributors	974	80122_Elk

Code	Canadian Surrogate Description	Code	Description
448	clothing and clothing accessories stores	975	80123_Wild boars
562	Waste management and remediation services	976	80124_Rabbit
601	SCL:12003 Petroleum Liquids Transportation (PIRD)	977	80125_Mink
602	SCL:12007 Oil Sands In-Situ Extraction and Processing (PIRD)	978	80126_Fox
603	SCL:12010 Light Medium Crude Oil Production (PIRD)	980	TOTSWIN
604	SCL:12011 Well Drilling (PIRD)	981	Harvest_Annual
605	SCL:12012 Well Servicing (PIRD)	982	Harvest_Perennial
606	SCL:12013 Well Testing (PIRD)	983	Synthfert_Annual
607	SCL:12014 Natural Gas Production (PIRD)	984	Synthfert_Perennial
608	SCL:12015 Natural Gas Processing (PIRD)	985	Tillage_Annual
609	SCL:12016 Heavy Crude Oil Cold Production (PIRD)	990	TOTFERT
610	SCL:12018 Disposal and Waste Treatment (PIRD)	996	urban_area
611	SCL:12019 Accidents and Equipment Failures (PIRD)	1251	OFFR_TOTFERT
612	SCL:12020 Natural Gas Transmission and Storage (PIRD)	1252	OFFR_MINES
651	MEIT C1C2 Anchored	1253	OFFR Other Construction not Urban
652	MEIT C1C2 Underway	1254	OFFR Commercial Services
653	MEIT C1C2 Berthed	1255	OFFR Oil Sands Mines
661	MEIT C3 Anchored	1256	OFFR Wood industries CANVEC
662	MEIT C3 Underway	1257	OFFR UNPAVED ROADS RURAL
663	MEIT C3 Berthed	1258	OFFR_Uilities
901	AIRPORT	1259	OFFR total dwelling
902	Military LTO	1260	OFFR_water
903	Commercial LTO	1261	OFFR_ALL_INDUST
904	General Aviation LTO	1262	OFFR Oil and Gas Extraction
905	Air Taxi LTO	1263	OFFR_ALLROADS
921	Commercial Fuel Combustion	1264	OFFR_AIRPORT
923	TOTAL INSTITUTIONAL AND GOVERNEMNT	1265	OFFR_RAILWAY
924	Primary Industry		

Table 3-24. Shapefiles and Attributes used to Compute Canadian Spatial Surrogates

Cod e	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
100	Population	gpr_gda	pruid	da_popdwell_100m_nolakes_1nov17	Pop

Cod e	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
101	total dwelling	gpr_gda	pruid	da_popdwell_100m_nolakes_1nov17	Urdwell
102	urban dwelling	gpr_gda	pruid	da_popdwell_100m_nolakes_1nov17	Uadwell
103	rural dwelling	gpr_gda	pruid	da_popdwell_100m_nolakes_1nov17	Radwell
104	capped total dwelling	gpr_gda	pruid	da_popdwell_100m_nolakes_1nov17	CAP_URDWEL
105	capped meat cooking dwelling	gpr	pruid	da_SimP_100m_pop_dwell_jul2014	Cap_Dwell
106	ALL_INDUST	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	ALL_INDUST
111	Farms	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	FARMS
113	Forestry and logging	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	FORLOG
116	Total Resources	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	TOTRESOURC
1251	OFFR_TOTFERT	gcd	CDID	naesi_fert	TOTFERT
1252	OFFR_MINES	gcd	CDID	mine	MINES
1253	OFFR Other Construction not Urban	gcd	CDID	construction_other	TOTAL
1254	OFFR Commercial Services	gcd	CDID	da2006_pop_labour_SimP_MaxOff_100m_noLake	COMSER
1255	OFFR Oil Sands Mines	gcd	CDID	OS_MinePit_D_v2	
1256	OFFR Wood industries CANVEC	gcd	CDID	wood_industries	WOOD
1257	OFFR UNPAVED ROADS RURAL	gcd	CDID	unpaved_ur	
1258	OFFR_Uilities	gcd	CDID	da2006_pop_labour_SimP_MaxOff_100m_noLake	UTILITIES
1259	OFFR total dwelling	gcd	CDID	da2006_pop_labour_SimP_MaxOff_100m_noLake	DATDWELL20
1260	OFFR_water	gcd	CDID	lu100_valid	
1261	OFFR_ALL_INDUST	gcd	CDID	da2006_pop_labour_SimP_MaxOff_100m_noLake	ALL_INDUST
1262	OFFR Oil and Gas Extraction	gcd	CDID	da2006_pop_labour_SimP_MaxOff_100m_noLake	OILGASEXTR
1263	OFFR_ALLROADS	gcd	CDID	allroads	
1264	OFFR_AIRPORT	gcd	CDID	offroad_osm_airport_locs_spring2017	Movements
1265	OFFR_RAILWAY	gcd	CDID	shp_railway_canvec_jul17_v2	LENGTH
200	Urban Primary Road Miles	gcd_ON4	CDID	NRN_CA_Simp2_16Apr2016_sphere	Class1
210	Rural Primary Road Miles	gcd_ON4	CDID	NRN_CA_Simp2_16Apr2016_sphere	Class2
211	Oil and Gas Extraction	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	OILGASEXTR
212	Mining except oil and gas	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	MINING2

Cod e	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
215	Oil Sands Mines	prov2006	pruid	OS_MinePit_D_v2	
216	Oil Sands Tailing Ponds	prov2006	pruid	OS_WetTailing_D_2015	
217	Oil Sands Plants	prov2006	Pruid	OS_PlantSite_D_2015	
220	Urban Secondary Road Miles	gcd_ON4	CDID	NRN_CA_Simp2_16Apr2016_sphere	Class3
221	Total Mining	prov2006	Pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	TOTALMI3
222	Utilities	prov2006	Pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	UTILITIES
230	Rural Secondary Road Miles	gcd_ON4	CDID	NRN_CA_Simp2_16Apr2016_sphere	Class4
233	Total Land Development	prov2006	Pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	TOTLND
240	capped population	gcd_ON4	CDID	da_popdwell_100m_nolakes_1nov17	CAPURPOP
308	Food manufacturing	prov2006	Pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	FOODMANU
321	Wood product manufacturing	prov2006	Pruid	da2006_SimplifyP_250m_sph here_treesa_Clip	WOODMANU
323	Printing and related support activities	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	PRINTSUPRT
324	Petroleum and coal products manufacturing	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	PETCOLMANU
326	Plastics and rubber products manufacturing	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	PLASTCMANU
327	Non-metallic mineral product manufacturing	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	MINERLMANU
331	Primary Metal Manufacturing	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	METALMANU
340	Construction - Oil and Gas	gpr_gda	pruid	loc_land_UOG2015_CO_v3_Queue_NB_NS	
350	Water	coast	pruid	CONT42_pop_water_Clip_b	pop
412	Petroleum product wholesaler-distributors	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	PETPRWSL
416	Building material and supplies wholesaler-distributors	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	BUILDPRWSL
447	Gasoline stations	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	GASSTOR
448	clothing and clothing accessories stores	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	CLOTHSTOR
482	Rail transportation	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	RAILTRANS
562	Waste management and remediation services	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	WASTEMGMT
901	AIRPORT	gcd	CDID	offroad_osm_airport_locs_spring2017	Movements
902	Military LTO	surg_2017	FAKEFIPS	aviation_runways_spring2017	Military
903	Commercial LTO	surg_2017	FAKEFIPS	aviation_runways_spring2017	Commercial

Cod e	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
904	General Aviation LTO	surg_2017	FAKEFIPS	aviation_runways_spring2017	General_Av
905	Air Taxi LTO	prov2006	pruid	Airport_movements_2006_MultiRingBuffer	SCC2275060
921	Commercial Fuel Combustion	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	COMFUEL
923	TOTAL INSTITUTIONAL AND GOVERNEMNT	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	TOTINSTGOV
924	Primary Industry	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	PRIM1
925	Manufacturing and Assembly	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	MANASSEM
926	Distribtution and Retail (no petroleum)	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	DISRET
927	Commercial Services	prov2006	pruid	da2006_pop_labour_SimP_MaxOff_100m_noLake	COMSER
933	Rail-Passenger	gpr_gda	pruid	shp_railway_canvec_jul17_v2	Passenger
934	Rail-Freight	gpr_gda	pruid	shp_railway_canvec_jul17_v2	Fret
935	Rail-Yard	gpr_gda	pruid	shp_railway_canvec_jul17_v2	Yard
940	PAVED ROADS NEW	gpr	fips	NRN_CA_Simp2_16Apr2016_sphere	PAVEDRD
942	UNPAVED ROADS	prov2006	pruid	unpaved4	
945	Commercial Marine Vessels	lowmedjet_II	CLASS	marine	SO2
946	Construction and mining			MERGE: 0.5*Mining except oil and gas+0.5*Total Land Development	
947	Agriculture Construction and mining			MERGE 0.34*Total Resources + 0.66 * Construction and mining	
948	Forest	prov2006	pruid	treesa_valid	
949	Combination of Dwelling			MERGE: 0.20*urban dwelling+0.80* rural dwelling	
951	Wood Consumption Percentage	gpr	fips	da2006_SimP_100m_WoodCon_1Aug14	WoodComp
955	UNPAVED_ROADS_AND_TRAILS	prov2006	pruid	unpaved5	
960	TOTBEEF	prov2006	pruid	naesi_livestk	TOTBEEF
970	TOTPOUL	prov2006	pruid	naesi_livestk	TOTPOULT
980	TOTSWIN	prov2006	pruid	naesi_livestk	TOTSWINE
990	TOTFERT	prov2006	pruid	naesi_fert	TOTFERT
996	urban_area	prov2006	pruid	ua2001	
961	80110_Broilers	gpr_gda	pruid	animal_nh3_to_agri_slc_80110_valid	QUANTITY
962	80111_Cattle_dairy_and_Heifer	gpr_gda	pruid	animal_nh3_to_agri_slc_80111_valid	QUANTITY

Cod e	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
963	80112_Cattle_non-Dairy	gpr_gda	pruid	animal_nh3_to_agri_slc_80112_valid	QUANTITY
964	80113_Laying_hens_and_Pullets	gpr_gda	pruid	animal_nh3_to_agri_slc_80113_valid	QUANTITY
965	80114_Horses	gpr_gda	pruid	animal_nh3_to_agri_slc_80114_valid	QUANTITY
966	80115_Sheep_and_Lamb	gpr_gda	pruid	animal_nh3_to_agri_slc_80115_valid	QUANTITY
967	80116_Swine	gpr_gda	pruid	animal_nh3_to_agri_slc_80116_valid	QUANTITY
968	80117_Turkeys	gpr_gda	pruid	animal_nh3_to_agri_slc_80117_valid	QUANTITY
969	80118_Goat	gpr_gda	pruid	animal_nh3_to_agri_slc_80118_valid	QUANTITY
971	80119_Buffalo	gpr_gda	pruid	animal_nh3_to_agri_slc_80119_valid	QUANTITY
972	80120_Llama_and_Alpacas	gpr_gda	pruid	animal_nh3_to_agri_slc_80120_valid	QUANTITY
973	80121_Deer	gpr_gda	pruid	animal_nh3_to_agri_slc_80121_valid	QUANTITY
974	80122_Elk	gpr_gda	pruid	animal_nh3_to_agri_slc_80122_valid	QUANTITY
975	80123_Wild boars	gpr_gda	pruid	animal_nh3_to_agri_slc_80123_valid	QUANTITY
976	80124_Rabbit	gpr_gda	pruid	animal_nh3_to_agri_slc_80124_valid	QUANTITY
977	80125_Mink	gpr_gda	pruid	animal_nh3_to_agri_slc_80125_valid	QUANTITY
978	80126_Fox	gpr_gda	pruid	animal_nh3_to_agri_slc_80126_valid	QUANTITY
979	80127_Mules_and_Asses	gpr_gda	pruid	animal_nh3_to_agri_slc_80127_valid	QUANTITY
981	Harvest_Annual	gpr_gda	pruid	harvest_pm10_Annual_to_agri_slc_valid	QUANTITY
982	Harvest_Perennial	gpr_gda	pruid	harvest_pm10_Perennial_to_agri_slc_valid	QUANTITY
983	Synthfert_Annual	gpr_gda	pruid	synth_fert_nh3_Annual_to_agri_slc_valid	QUANTITY
984	Synthfert_Perennial	gpr_gda	pruid	synth_fert_nh3_Perennial_to_agri_slc_valid	QUANTITY
985	Tillage_Annual	gpr_gda	pruid	tillage_pm10_Annual_to_agri_slc_valid	QUANTITY
601	SCL:12003 Petroleum Liquids Transportation (PIRD)	gpr_gda	pruid	scl12003_valid	
602	SCL:12007 Oil Sands In-Situ Extraction and Processing (PIRD)	gpr_gda	pruid	scl12007_valid	NONE
603	SCL:12010 Light Medium Crude Oil Production (PIRD)	gpr_gda	pruid	scl12010_valid	NONE
604	SCL:12011 Well Drilling (PIRD)	gpr_gda	pruid	scl12011_valid	NONE
605	SCL:12012 Well Servicing (PIRD)	gpr_gda	pruid	scl12012_valid	NONE

Code	Surrogate	Data Shapefile	Data Attribute	Weight Shapefile	Weight Attribute
606	SCL:12013 Well Testing (PIRD)	gpr_gda	pruid	scl12013_valid	NONE
607	SCL:12014 Natural Gas Production (PIRD)	gpr_gda	pruid	scl12014_valid	NONE
608	SCL:12015 Natural Gas Processing (PIRD)	gpr_gda	pruid	scl12015_valid	NONE
609	SCL:12016 Heavy Crude Oil Cold Production (PIRD)	gpr_gda	pruid	scl12016_valid	NONE
610	SCL:12018 Disposal and Waste Treatment (PIRD)	gpr_gda	pruid	scl12018_valid	NONE
611	SCL:12019 Accidents and Equipment Failures (PIRD)	gpr_gda	pruid	scl12019_valid	NONE
612	SCL:12020 Natural Gas Transmission and Storage (PIRD)	gpr_gda	pruid	scl12020	NONE
952	Residential Fuel Wood Combustion (PIRD)	gpr_gda	pruid	scl20401_valid	NONE
651	MEIT C1C2 Anchored	lowmedjet_II	CLASS	MEIT_2280002101_2018	fuel
652	MEIT C1C2 Underway	lowmedjet_II	CLASS	MEIT_2280002202_2018	fuel
653	MEIT C1C2 Berthed	lowmedjet_II	CLASS	MEIT_2280002301_2018	fuel
661	MEIT C3 Anchored	lowmedjet_II	CLASS	MEIT_2280003101_2018	fuel
662	MEIT C3 Underway	lowmedjet_II	CLASS	MEIT_2280003200_2018	fuel
663	MEIT C3 Berthed	lowmedjet_II	CLASS	MEIT_2280003301_2018	fuel

Table 3-25. Shapefiles and Attributes used to Compute Mexican Spatial Surrogates

Code	SURROGATE	WEIGHT SHAPEFILE	WEIGHT ATTRIBUTE
10	MEX Population	mex_population_2020	gridcode_Y
22	MEX Total Road Miles	mex_roads	NONE
24	MEX Total Railroads Miles	mex_railroads	NONE
26	MEX Total Agriculture	mex_agriculture	NONE
36	MEX Commercial plus Industrial Land	mex_com_ind_land	NONE
44	MEX Airports Area	mex_airports_area	NONE
45	MEX Airports Point	mex_airports_point	NONE
48	MEX Brick Kilns	mex_brick_kilns	NONE
50	MEX Border Crossings	mex_border_crossings	SUM_Value

Table 3-26. 2021 CAP Emissions Allocated to Mexican and Canadian Spatial Surrogates for 12US1 (short tons)

Code	Mexican or Canadian Surrogate Description	NH₃	NO_x	PM_{2.5}	SO₂	VOC
11	MEX Population	25,943	81,217	1,905	7,606	148,785
14	MEX Residential Heating - Wood	206	1,737	6,171	182	20,519
16	MEX Residential Heating - Distillate Oil	0	5	0	0	0

Code	Mexican or Canadian Surrogate Description	NH ₃	NO _x	PM _{2.5}	SO ₂	VOC
20	MEX Residential Heating - LP Gas	0	2,967	206	25	104
22	MEX Total Road Miles	2,508	279,841	12,582	5,616	86,744
24	MEX Total Railroads Miles	0	22,452	497	198	900
26	MEX Total Agriculture	137,780	11,683	13,655	13,571	2,375
32	MEX Commercial Land	0	73	1,591	0	5,448
34	MEX Industrial Land	44	1,809	652	9	28,743
36	MEX Commercial plus Industrial Land	0	2,630	61	5	111,541
38	MEX Commercial plus Institutional Land	0	714	50	6	43
40	MEX Residential (RES1-4)+Comercial+Industrial+Institutional+Government	0	44	236	7	139,071
42	MEX Personal Repair (COM3)	0	0	0	0	6,729
44	MEX Airports Area	0	2,894	60	307	1,740
48	MEX Brick Kilns	3	237	3,892	169	187
50	MEX Mobile sources - Border Crossing	4	86	3	0	65
100	CAN Population	704	57	223	17	3,912
101	CAN total dwelling	0	0	0	0	107,219
104	CAN Capped Total Dwelling	313	32,274	2,434	1,979	1,654
106	CAN ALL_INDUST	0	0	570	0	0
113	CAN Forestry and logging	83	627	2,934	15	2,717
200	CAN Urban Primary Road Miles	1,559	75,445	2,678	192	7,265
210	CAN Rural Primary Road Miles	596	40,590	1,414	82	2,937
212	CAN Mining except oil and gas	0	0	1,702	0	0
220	CAN Urban Secondary Road Miles	2,926	119,891	5,416	382	19,355
221	CAN Total Mining	0	0	12,915	0	0
222	CAN Utilities	0	2,280	2,628	32	99
230	CAN Rural Secondary Road Miles	1,579	74,961	2,705	199	7,837
240	CAN Total Road Miles	338	45,469	1,178	39	80,841
308	CAN Food manufacturing	0	0	17,395	0	5,168
321	CAN Wood product manufacturing	515	1,689	585	210	8,419
323	CAN Printing and related support activities	0	0	0	0	19,532
324	CAN Petroleum and coal products manufacturing	0	988	1,383	411	6,286
326	CAN Plastics and rubber products manufacturing	0	0	0	0	21,856
327	CAN Non-metallic mineral product manufacturing	0	0	6,946	0	0
331	CAN Primary Metal Manufacturing	0	130	4,563	25	54
412	CAN Petroleum product wholesaler-distributors	0	0	0	0	37,272

Code	Mexican or Canadian Surrogate Description	NH ₃	NO _x	PM _{2.5}	SO ₂	VOC
448	CAN clothing and clothing accessories stores	0	0	0	0	178
562	CAN Waste management and remediation services	2,681	1,245	2,351	2,139	16,053
601	CAN SCL:12003 Petroleum Liquids Transportation (PIRD)	0	0	12	158	6,092
602	CAN SCL:12007 Oil Sands In-Situ Extraction and Processing (PIRD)	0	0	0	0	109
603	CAN SCL:12010 Light Medium Crude Oil Production (PIRD)	0	0	0	0	2
604	CAN SCL:12011 Well Drilling (PIRD)	0	0	0	585	626
605	CAN SCL:12012 Well Servicing (PIRD)	0	0	0	65	69
606	CAN SCL:12013 Well Testing (PIRD)	0	0	0	0	0
607	CAN SCL:12014 Natural Gas Production (PIRD)	0	30	1	0	203
608	CAN SCL:12015 Natural Gas Processing (PIRD)	0	0	0	0	0
611	CAN SCL:12019 Accidents and Equipment Failures (PIRD)	0	0	0	0	95,082
612	CAN SCL:12020 Natural Gas Transmission and Storage (PIRD)	1	735	55	11	402
901	CAN Airport	0	98	9	0	11
921	CAN Commercial Fuel Combustion	193	21,981	2,413	442	954
923	CAN TOTAL INSTITUTIONAL AND GOVERNMENT	0	0	0	0	14,399
924	CAN Primary Industry	0	0	0	0	32,546
925	CAN Manufacturing and Assembly	0	0	0	0	67,574
926	CAN Distribution and Retail (no petroleum)	0	0	0	0	6,649
927	CAN Commercial Services	0	0	0	0	30,535
933	CAN Rail-Passenger	1	3,064	62	1	118
934	CAN Rail-Freight	48	77,088	1,533	43	3,410
935	CAN Rail-Yard	1	4,562	95	1	278
940	CAN Paved Roads New	0	0	25,020	0	0
946	CAN Construction and Mining	43	2,758	156	269	40
951	CAN Wood Consumption Percentage	1,090	12,112	73,726	1,730	102,859
955	CAN UNPAVED_ROADS_AND_TRAILS	0	0	418,718	0	0
961	CAN 80110_Broilers	13,041	0	115	0	12,784
962	CAN 80111_Cattle_dairy_and_Heifer	59,965	0	276	0	40,508
963	CAN 80112_Cattle_non-Dairy	171,295	0	884	0	42,868
964	CAN 80113_Laying_hens_and_Pullets	9,768	0	40	0	10,594
965	CAN 80114_Horses	3,046	0	19	0	1,321
966	CAN 80115_Sheep_and_Lamb	2,200	0	6	0	170

Code	Mexican or Canadian Surrogate Description	NH₃	NO_x	PM_{2.5}	SO₂	VOC
967	CAN 80116_Swine	61,897	0	824	0	9,947
968	CAN 80117_Turkeys	5,046	0	41	0	4,508
969	CAN 80118_Goat	1,743	0	2	0	135
971	CAN 80119_Buffalo	2,175	0	6	0	517
972	CAN 80120_Llama_and_Alpacas	114	0	0	0	0
973	CAN 80121_Deer	19	0	0	0	0
974	CAN 80122_Elk	19	0	0	0	0
975	CAN 80123_Wild boars	35	0	0	0	0
976	CAN 80124_Rabbit	75	0	0	0	1
977	CAN 80125_Mink	285	0	0	0	951
978	CAN 80126_Fox	4	0	0	0	3
981	CAN Harvest_Annual	0	0	24,815	0	0
983	CAN Synthfert_Annual	170,809	3,564	2,114	5,870	129
985	CAN Tillage_Annual	0	0	106,769	0	0
996	CAN urban_area	0	0	3,570	0	0
1251	CAN OFFR_TOTFERT	83	61,875	4,283	57	6,205
1252	CAN OFFR_MINES	1	579	41	1	81
1253	CAN OFFR Other Construction not Urban	67	38,266	4,513	45	10,335
1254	CAN OFFR Commercial Services	46	16,605	2,489	39	38,169
1255	CAN OFFR Oil Sands Mines	0	0	0	0	0
1256	CAN OFFR Wood industries CANVEC	9	3,294	265	7	933
1257	CAN OFFR Unpaved Roads Rural	24	10,153	634	21	27,111
1258	CAN OFFR_Utilityies	8	4,164	214	6	863
1259	CAN OFFR total dwelling	17	6,245	600	15	12,435
1260	CAN OFFR_water	17	4,725	363	25	24,577
1261	CAN OFFR_ALL_INDUST	4	5,000	176	2	880
1262	CAN OFFR Oil and Gas Extraction	1	389	31	0	122
1263	CAN OFFR_ALLROADS	3	1,782	175	2	463
1265	CAN OFFR_CANRAIL	0	65	6	0	12

4 Emission Summaries

Tables 4-1 through 4-3 summarize emissions by sector for the 2021 platform. These summaries are provided at the national level by sector for the contiguous U.S. and for the portions of Canada and Mexico inside the larger 12km domain (12US1) discussed in Section 3.1. Note that totals for the 12US2 domain are not available here, but the sum of the U.S. sectors would be essentially the same and only the Canadian and Mexican emissions would change according to the extent of the grids to the north and south of the continental United States. The afdust sector emissions here represent the emissions *after* application of both the land use (transport fraction) and meteorological adjustments; therefore, this sector is called “afdust_adj” in these summaries. The onroad sector totals are post-SMOKE-MOVES totals, representing air quality model-ready emission totals, and include CARB emissions for California. The cmv sectors include U.S. emissions within state waters only; these extend to roughly 3-5 miles offshore and include CMV emissions at U.S. ports. “Offshore” represents CMV emissions that are outside of U.S. state waters. Canadian CMV emissions are included in the other sector. The total of all US sectors is listed as “Con U.S. Total.” Table 4-4 shows the emissions for key criteria pollutants by sector for Alaska, Hawaii, Puerto Rico, and the Virgin Islands.

State totals and other summaries are available in the reports area on the FTP site for the 2021 platform (<https://gaftp.epa.gov/Air/emismod/2021/>).

Table 4-1. National by-sector CAP emissions for the 2021 platform, 12US1 grid (tons/yr)

Sector	CO	NH3	NOX	PM10	PM2_5	SO2	VOC
afdust_adj				6,027,656	821,738		
airports	333,660	0	83,674	8,521	7,533	9,126	50,041
cmv_c1c2	19,892	68	134,167	3,662	3,548	615	5,116
cmv_c3	10,252	44	81,846	2,507	2,307	5,767	4,687
fertilizer		1,275,333					
livestock		2,824,644					225,971
nonpt	2,173,885	145,073	723,480	711,625	622,905	106,258	1,005,040
nonroad	11,037,304	1,998	816,810	80,205	75,312	917	945,175
np_oilgas	654,275	43	728,663	14,048	13,880	139,514	2,876,480
np_solvents	0	0	0	0	0	0	2,716,884
onroad	14,391,846	183,954	2,258,178	188,833	74,375	8,748	1,039,569
ptegu	467,560	21,482	879,533	125,564	109,306	968,652	26,731
ptagfire	773,523	172,492	33,830	114,547	74,469	13,729	125,668
ptfire-rx	7,825,125	68,537	125,890	1,267,230	1,131,103	80,356	1,586,259
ptfire-wild	17,682,184	178,672	163,750	3,826,054	2,386,263	166,480	4,865,824
ptnonipm	1,226,638	61,712	793,938	350,108	228,948	456,300	726,236
pt_oilgas	174,223	9,095	318,687	12,460	11,916	31,186	195,937
rail	96,705	296	444,124	11,360	10,982	369	18,367
rwc	2,940,341	22,616	44,790	448,615	446,995	11,894	453,043
beis	3,314,764		989,492				28,539,802
CONUS + beis	63,122,176	4,966,059	8,620,851	13,192,993	6,021,580	1,999,910	45,406,832
Canada ag		500,395		6,562	1,875		124,257
Canada oil and gas 2D		8					306,206
Canada afdust				1,028,722	194,713		
Canada pt dust				3,588	443		
Canada area	2,040,850	5,983	317,182	184,382	134,440	14,175	711,153
Canada onroad	1,669,722	6,994	356,236	24,858	13,378	893	118,094
Canada point	1,021,439	18,569	538,357	112,670	42,409	483,703	148,235
Canada fires	18,068,782	259,108	302,681	3,543,123	3,141,541	173,644	5,070,468
Canada cmv_c1c2	3,179	10	20,497	541	525	64	720
Canada cmv_c3	7,750	27	60,418	1,498	1,378	3,331	3,773
Mexico ag		137,778		53,862	11,638		
Mexico area	98,400	26,201	57,960	42,108	20,576	21,937	425,809
Mexico onroad	1,418,503	2,509	350,527	13,377	9,349	5,778	127,181
Mexico point	158,097	979	199,367	90,822	53,973	341,038	32,822
Mexico fires	415,564	6,820	24,903	54,701	45,743	4,240	204,334
Mexico cmv_c1c2	157	0	1,016	27	26	4	42
Mexico cmv_c3	9,601	87	82,079	4,907	4,514	12,970	4,596
Offshore cmv_c1c2	4,445	14	28,377	743	721	88	1,065
Offshore cmv_c3	51,349	309	414,286	17,467	16,069	43,957	25,126
Offshore pt_oilgas	28,548	5	34,658	422	416	321	31,400
Can/Mex/offshore total	24,996,385	965,797	2,788,544	5,184,380	3,693,725	1,106,143	7,335,281

Table 4-2. National by-sector VOC HAP emissions for the 2021 platform, 12US1 grid (tons/yr)

Sector	Acetaldehyd e	Benzene	Formaldehyde	Methanol	Naphthalen e	Acrolein	1,3- Butadiene
airports	1,469	660	4,252	615	728	836	596
cmv_c1c2	22	10	94		9	9	5
cmv_c3	20	10	86		9	9	5
livestock	1,786	486		20,352	0		
nonpt	11,536	6,952	7,434	14,561	500	169	1,036
nonroad	8,524	25,793	21,093	1,191	1,462	1,286	4,369
np_oilgas	4,146	33,081	40,914	2,691	124	2,567	579
np_solvents	72	351	7	15,551	7,804		
onroad	8,979	18,353	11,404	1,451	1,416	817	2,441
ptegu	273	288	2,195	94	27	210	4
ptagfire	10,283	2,064	8,480				841
ptfire-rx	64,740	20,900	126,337	94,386	19,090	25,806	16,037
ptfire-wild	131,650	36,614	234,404	241,636	45,902	40,205	20,944
ptnonipm	5,307	2,756	5,888	49,370	674	827	742
pt_oilgas	2,772	2,098	11,858	1,636	78	1,813	249
rail	1,437	413	4,092		50	294	34
rcw	51,488	13,358	36,066		6,978	1,958	3,632
beis	367,982		504,811	1,959,801			
CONUS + beis	672,484	164,188	1,019,415	2,403,335	84,852	76,804	51,514
Canada ag	1,398	159	0	32,663	0		
Canada oil and gas 2D	0	922	0	0	0		
Canada area	15,577	12,804	13,095	4,031	2,601		
Canada onroad	2,138	5,153	2,955	0	39		
Canada point	1,506	2,006	5,056	10,313	23		
Canada fires	193,822	51,813	385,491	432,821	63,192	58,857	31,274
Canada cmv_c1c2	3	1	13	0	1	1	1
Canada cmv_c3	16	8	69	0	7	7	4
Mexico area	3,099	1,802	2,538	2,663	470		
Mexico onroad	530	3,026	1,283	573	180	91	446
Mexico point	65	1,208	2,587	519	11		
Mexico fires	10,924	2,947	12,434	4,760	576	0	0
Mexico cmv_c1c2	0	0	1	0	0	0	0
Mexico cmv_c3	15	7	64	0	22	8	5
Offshore cmv_c1c2	4	2	20	0	2	2	1
Offshore cmv_c3	94	46	412	0	79	46	25

Offshore pt_oilgas	248	41	595	0	0	0	0
Non-U.S. Total	229,438	81,946	426,612	488,342	67,203	59,013	31,755

Table 4-3. National by-sector Diesel PM and metal emissions for the 2021 platform, 12US1 grid (tons/yr)

Sector	Diesel PM ₁₀	Diesel PM _{2.5}	Chromium Hex	Arsenic	Cadmium	Nickel	Manganese	Ethylene Oxide
airports	26	26	0.00	0.00	0.00	0.00	0.00	0.00
cmv_c1c2	3,662	3,548	0.00	0.09	0.84	2.44	0.01	0.00
cmv_c3	2,507	2,307	0.00	0.06	0.54	1.58	0.01	0.00
nonpt	0	0	0.37	7.78	5.58	37.19	11.59	0.98
nonroad	40,060	38,709	0.01	0.75	0.00	4.49	1.27	0.00
np_oilgas	0	0	0.00	0.06	0.19	0.20	0.12	0.00
onroad	41,285	37,986	0.07	7.23	0.00	15.24	36.30	0.00
ptegu	0	0	3.63	14.58	5.06	47.26	81.97	0.00
ptnonipm	145	140	21.45	22.08	10.26	136.70	485.01	84.29
pt_oilgas	0	0	0.02	0.02	0.28	5.24	2.46	0.00
rail	11,360	10,982	0.00	0.06	0.00	0.21	0.12	0.00
rwc	0	0	0.00	0.00	0.11	0.10	0.83	0.00
Con. U.S. Total	99,045	93,699	25.55	52.70	22.85	250.65	619.68	85.27
Canada								
cmv_c1c2	541	525	0.00	0.01	0.12	0.36	0.00	0.00
Canada cmv_c3	1,498	1,378	0.00	0.04	0.33	0.95	0.00	0.00
Mexico								
cmv_c1c2	27	26	0.00	0.00	0.01	0.02	0.00	0.00
Mexico cmv_c3	4,907	4,514	0.00	0.12	1.07	3.10	0.01	0.00
Offshore								
cmv_c1c2	743	721	0.00	0.02	0.17	0.49	0.00	0.00
Offshore cmv_c3	17,467	16,069	0.00	0.42	3.79	11.04	0.05	0.00
Offshore pt_oilgas	0	0	0.02	0.00	0.26	0.00	0.00	0.00

5 References

- Adelman, Z. 2012. *Memorandum: Fugitive Dust Modeling for the 2008 Emissions Modeling Platform*. UNC Institute for the Environment, Chapel Hill, NC. September 28, 2012.
- Adelman, Z. 2016. *2014 Emissions Modeling Platform Spatial Surrogate Documentation*. UNC Institute for the Environment, Chapel Hill, NC. October 1, 2016. Available at https://gaftp.epa.gov/Air/emismod/2014/v1/spatial_surrogates/.
- Adelman, Z., M. Omary, Q. He, J. Zhao and D. Yang, J. Boylan, 2012. “A Detailed Approach for Improving Continuous Emissions Monitoring Data for Regulatory Air Quality Modeling.” Presented at the 2012 International Emission Inventory Conference, Tampa, Florida. Available from <http://www.epa.gov/ttn/chief/conference/ei20/index.html#ses-5>.
- Appel, K.W., Napelenok, S., Hogrefe, C., Pouliot, G., Foley, K.M., Roselle, S.J., Pleim, J.E., Bash, J., Pye, H.O.T., Heath, N., Murphy, B., Mathur, R., 2018. Overview and evaluation of the Community Multiscale Air Quality Model (CMAQ) modeling system version 5.2. In Mensink C., Kallos G. (eds), *Air Pollution Modeling and its Application XXV*. ITM 2016. Springer Proceedings in Complexity. Springer, Cham. Available at https://doi.org/10.1007/978-3-319-57645-9_11.
- Bash, J.O., Baker, K.R., Beaver, M.R., Park, J.-H., Goldstein, A.H., 2016. Evaluation of improved land use and canopy representation in BEIS with biogenic VOC measurements in California. Available from <http://www.geosci-model-dev.net/9/2191/2016/>.
- Bullock Jr., R. and K. A. Brehme (2002) “Atmospheric mercury simulation using the CMAQ model: formulation description and analysis of wet deposition results.” *Atmospheric Environment* 36, pp 2135–2146. Available at [https://doi.org/10.1016/S1352-2310\(02\)00220-0](https://doi.org/10.1016/S1352-2310(02)00220-0).
- California Air Resources Board (CARB): Final 2015 Consumer & Commercial Product Survey Data Summaries, 2019.
- Coordinating Research Council (CRC), 2017. Report A-100. Improvement of Default Inputs for MOVES and SMOKE-MOVES. Final Report. February 2017. Available at http://crcsite.wpengine.com/wp-content/uploads/2019/05/ERG_FinalReport_CRCA100_28Feb2017.pdf.
- Coordinating Research Council (CRC), 2019. Report A-115. Developing Improved Vehicle Population Inputs for the 2017 National Emissions Inventory. Final Report. April 2019. Available at http://crcsite.wpengine.com/wp-content/uploads/2019/05/CRC-Project-A-115-Final-Report_20190411.pdf.
- Drillinginfo, Inc. 2017. “DI Desktop Database powered by HPDI.” Currently available from <https://www.enverus.com/>.
- England, G., Watson, J., Chow, J., Zielenska, B., Chang, M., Loos, K., Hidy, G., 2007. “Dilution-Based Emissions Sampling from Stationary Sources: Part 2-- Gas-Fired Combustors Compared with Other Fuel-Fired Systems,” *Journal of the Air & Waste Management Association*, 57:1, 65-78,

- DOI: 10.1080/10473289.2007.10465291. Available at <https://www.tandfonline.com/doi/abs/10.1080/10473289.2007.10465291>.
- EPA, 2007a. Control of Hazardous Air Pollutants from Mobile Sources Regulatory Impact Analysis. EPA420-R-07-002. EPA Office of Transportation and Air Quality (OTAQ) Assessment and Standards Division, Ann Arbor, MI. Available online at <https://nepis.epa.gov/Exe/ZyPdf.cgi?Dockey=P1004LNN.PDF>.
- EPA, 2015b. Draft Report Speciation Profiles and Toxic Emission Factors for Nonroad Engines. EPA-420-R-14-028. Available at https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=309339&CFID=83476290&CFTOKEN=35281617.
- EPA, 2015c. Speciation of Total Organic Gas and Particulate Matter Emissions from On-road Vehicles in MOVES2014. EPA-420-R-15-022. Available at <https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100NOJG.pdf>.
- EPA, 2016. SPECIATE Version 4.5 Database Development Documentation, U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Research Triangle Park, NC 27711, EPA/600/R-16/294, September 2016. Available at https://www.epa.gov/sites/production/files/2016-09/documents/speciate_4.5.pdf.
- EPA, 2018. AERMOD Model Formulation and Evaluation Document. EPA-454/R-18-003. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711. Available at https://www3.epa.gov/ttn/scram/models/aermod/aermod_mfed.pdf.
- EPA, 2019. Final Report, SPECIATE Version 5.0, Database Development Documentation, Research Triangle Park, NC, EPA/600/R-19/988. . Available at <https://www.epa.gov/air-emissions-modeling/speciate-51-and-50-addendum-and-final-report>.
- EPA and National Emissions Inventory Collaborative (NEIC), 2019. Technical Support Document (TSD) Preparation of Emissions Inventories for the Version 7.2 North American Emissions Modeling Platform. Available at <https://www.epa.gov/air-emissions-modeling/2016-version-72-technical-support-document>.
- EPA, 2020. Population and Activity of Onroad Vehicles in MOVES3. EPA-420-R-20-023. Office of Transportation and Air Quality. US Environmental Protection Agency. Ann Arbor, MI. November 2020. Available under the MOVES3 section at <https://www.epa.gov/moves/moves-technical-reports>.
- EPA, 2020b. Technical Support document: “Development of Mercury Speciation Factors for EPA’s Air Emissions Modeling Programs, April 2020”. US EPA Office of Air Quality Planning and Standards.
- EPA, 2021. 2017 National Emission Inventory: January 2021 Updated Release, Technical Support Document. U.S. Environmental Protection Agency, OAQPS, Research Triangle Park, NC 27711. Available at: <https://www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-technical-support-document-tsd>.
- EPA, 2021. 2017 National Emissions Inventory (NEI) data, Research Triangle Park, NC, January 2021. <https://www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-data>.

- EPA and NEIC, 2021. Technical Support Document (TSD) Preparation of Emissions Inventories for the 2016v1 North American Emissions Modeling Platform. Available at: <https://www.epa.gov/air-emissions-modeling/2016-version-1-technical-support-document>.
- EPA, 2022a. Technical Support Document EPA's Air Toxics Screening Assessment – 2018 AirToxScreen TSD. Available at: <https://www.epa.gov/AirToxScreen/2018-airtoxscreen-technical-support-document>.
- EPA, 2022b. Technical Support Document: Preparation of Emissions Inventories for the 2019 North American Emissions Modeling Platform. Available at: <https://www.epa.gov/air-emissions-modeling/2019-emissions-modeling-platform-technical-support-document>.
- EPA, 2023. 2020 National Emission Inventory Technical Support Document. U.S. Environmental Protection Agency, OAQPS, Research Triangle Park, NC 27711. Available at: <https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-technical-support-document-tsd>.
- ERG, 2016b. "Technical Memorandum: Modeling Allocation Factors for the 2014 Oil and Gas Nonpoint Tool." Available at https://gaftp.epa.gov/air/emismod/2014/v1/spatial_surrogates/oil_and_gas/.
- ERG, 2017. "Technical Report: Development of Mexico Emission Inventories for the 2014 Modeling Platform." Available at https://gaftp.epa.gov/air/emismod/2016/v1/reports/EPA%205-18%20Report_Clean%20Final_01042017.pdf.
- ERG, 2018. Technical Report: "2016 Nonpoint Oil and Gas Emission Estimation Tool Version 1.0". Available at https://gaftp.epa.gov/air/emismod/2016/v1/reports/2016%20Nonpoint%20Oil%20and%20Gas%20Emission%20Estimation%20Tool%20V1_0%20December_2018.pdf.
- The Freedonia Group: Solvents, Industry Study #3429, 2016.
- Khare, P., and Gentner, D. R.: Considering the future of anthropogenic gas-phase organic compound emissions and the increasing influence of non-combustion sources on urban air quality, *Atmos Chem Phys*, 18, 5391-5413, 10.5194/acp-18-5391-2018, 2018.
- Luecken D., Yarwood G, Hutzell WT, 2019. Multipollutant modeling of ozone, reactive nitrogen and HAPs across the continental US with CMAQ-CB6. *Atmospheric environment*. 2019 Mar 15;201:62-72.
- Mansouri, K., Grulke, C. M., Judson, R. S., and Williams, A. J.: OPERA models for predicting physicochemical properties and environmental fate endpoints, *J Cheminformatics*, 10, 10.1186/s13321-018-0263-1, 2018.
- McCarty, J.L., Korontzi, S., Jutice, C.O., and T. Loboda. 2009. The spatial and temporal distribution of crop residue burning in the contiguous United States. *Science of the Total Environment*, 407 (21): 5701-5712. Available at <https://doi.org/10.1016/j.scitotenv.2009.07.009>.
- MDNR, 2008. "A Minnesota 2008 Residential Fuelwood Assessment Survey of individual household responses". Minnesota Department of Natural Resources. Available from http://files.dnr.state.mn.us/forestry/um/residentialfuelwoodassessment07_08.pdf.

- NCAR, 2016. FIRE EMISSION FACTORS AND EMISSION INVENTORIES, FINN Data. downloaded 2014 SAPRC99 version from <http://bai.acom.ucar.edu/Data/fire/>.
- NEIC, 2019. Specification sheets for the 2016v1 platform. Available from <http://views.cira.colostate.edu/wiki/wiki/10202>.
- NESCAUM, 2006. "Assessment of Outdoor Wood-fired Boilers". Northeast States for Coordinated Air Use Management (NESCAUM) report. Available from http://www.nescaum.org/documents/assessment-of-outdoor-wood-fired-boilers/2006-1031-owb-report_revised-june2006-appendix.pdf.
- NYSERDA, 2012. "Environmental, Energy Market, and Health Characterization of Wood-Fired Hydronic Heater Technologies, Final Report". New York State Energy Research and Development Authority (NYSERDA). Available from: <http://www.nyserda.ny.gov/Publications/Case-Studies/-/media/Files/Publications/Research/Environmental/Wood-Fired-Hydronic-Heater-Tech.ashx>.
- Pouliot, G., H. Simon, P. Bhawe, D. Tong, D. Mobley, T. Pace, and T. Pierce. 2010. "Assessing the Anthropogenic Fugitive Dust Emission Inventory and Temporal Allocation Using an Updated Speciation of Particulate Matter." International Emission Inventory Conference, San Antonio, TX. Available at http://www3.epa.gov/ttn/chief/conference/ei19/session9/pouliot_pres.pdf.
- Pouliot, G. and J. Bash, 2015. Updates to Version 3.61 of the Biogenic Emission Inventory System (BEIS). Presented at Air and Waste Management Association conference, Raleigh, NC, 2015.
- Pouliot G, Rao V, McCarty JL, Soja A. Development of the crop residue and rangeland burning in the 2014 National Emissions Inventory using information from multiple sources. Journal of the Air & Waste Management Association. 2017 Apr 27;67(5):613-22.
- Reichle, L., R. Cook, C. Yanca, D. Sonntag, 2015. "Development of organic gas exhaust speciation profiles for nonroad spark-ignition and compression-ignition engines and equipment", Journal of the Air & Waste Management Association, 65:10, 1185-1193, DOI: 10.1080/10962247.2015.1020118. Available at <https://doi.org/10.1080/10962247.2015.1020118>.
- Reff, A., Bhawe, P., Simon, H., Pace, T., Pouliot, G., Mobley, J., Houyoux. M. "Emissions Inventory of PM_{2.5} Trace Elements across the United States", Environmental Science & Technology 2009 43 (15), 5790-5796, DOI: 10.1021/es802930x. Available at <https://doi.org/10.1021/es802930x>.
- Sarwar, G., S. Roselle, R. Mathur, W. Appel, R. Dennis, "A Comparison of CMAQ HONO predictions with observations from the Northeast Oxidant and Particle Study", Atmospheric Environment 42 (2008) 5760–5770). Available at <https://doi.org/10.1016/j.atmosenv.2007.12.065>.
- Schauer, J., G. Lough, M. Shafer, W. Christensen, M. Arndt, J. DeMinter, J. Park, "Characterization of Metals Emitted from Motor Vehicles," Health Effects Institute, Research Report 133, March 2006. Available at <https://www.healtheffects.org/publication/characterization-metals-emitted-motor-vehicles>.
- Seltzer, K. M., Pennington, E., Rao, V., Murphy, B. N., Strum, M., Isaacs, K. K., and Pye, H. O. T., 2021: "Reactive organic carbon emissions from volatile chemical products", Atmos. Chem. Phys. 21,

- 5079–5100, 2021. <https://doi.org/10.5194/acp-21-5079-2021> and <https://acp.copernicus.org/articles/21/5079/2021/>.
- Skamarock, W., J. Klemp, J. Dudhia, D. Gill, D. Barker, M. Duda, X. Huang, W. Wang, J. Powers, 2008. A Description of the Advanced Research WRF Version 3. NCAR Technical Note. National Center for Atmospheric Research, Mesoscale and Microscale Meteorology Division, Boulder, CO. June 2008. Available at: http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3_bw.pdf.
- Swedish Environmental Protection Agency, 2004. Swedish Methodology for Environmental Data; Methodology for Calculating Emissions from Ships: 1. Update of Emission Factors.
- U.S. Bureau of Labor and Statistics, 2020. Producer Price Index by Industry, retrieved from FRED, Federal Reserve Bank of St. Louis, available at: <https://fred.stlouisfed.org/categories/31>, access date: 21 August 2020.
- U.S. Census Bureau: Paint and Allied Products - 2010, MA325F(10), 2011. <https://www.census.gov/data/tables/time-785-series/econ/cir/ma325f.html>.
- U.S. Census Bureau, Economy Wide Statistics Division: County Business Patterns, 2018. <https://www.census.gov/programs-surveys/cbp/data/datasets.html>.
- U.S. Department of Transportation and the U.S. Department of Commerce, 2015. 2012 Commodity Flow Survey, EC12TCF-US. <https://www.census.gov/library/publications/2015/econ/ec12tcf-us.html>.
- U.S. Energy Information Administration, 2019. The Distribution of U.S. Oil and Natural Gas Wells by Production Rate, Washington, DC. <https://www.eia.gov/petroleum/wells/>.
- Wang, Y., P. Hopke, O. V. Rattigan, X. Xia, D. C. Chalupa, M. J. Utell. (2011) “Characterization of Residential Wood Combustion Particles Using the Two-Wavelength Aethalometer”, Environ. Sci. Technol., 45 (17), pp 7387–7393. Available at <https://doi.org/10.1021/es2013984>.
- Weschler, C. J., and Nazaroff, W. W.: Semivolatile organic compounds in indoor environments, Atmos Environ, 42, 9018-9040, 2008.
- Wiedinmyer, C., Y. Kimura, E. C. McDonald-Buller, L. K. Emmons, R. R. Buchholz, W. Tang, K. Seto, M. B. Joseph, K. C. Barsanti, A. G. Carlton, and R. Yokelson, Volume 16, issue 13, GMD, 16, 3873–3891, 2023. <https://gmd.copernicus.org/articles/16/3873/2023/>.
- Wiedinmyer, C., S.K. Akagi, R.J. Yokelson, L.K. Emmons, J.A. Al-Saadi³, J. J. Orlando¹, and A. J. Soja. (2011) “The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning”, Geosci. Model Dev., 4, 625-641. <http://www.geosci-model-dev.net/4/625/2011/> doi:10.5194/gmd-4-625-2011.
- Yarwood, G., R. Beardsley, Y. Shi, and B. Czader: Revision 5 of the Carbon Bond 6 Mechanism (CB6r5). Presented at the Annual CMAS Conference, Chapel Hill, NC, 2020.
- Zhu, Henze, et al, 2013. “Constraining U.S. Ammonia Emissions using TES Remote Sensing Observations and the GEOS-Chem adjoint model”, Journal of Geophysical Research: Atmospheres, 118: 1-14. Available at <https://doi.org/10.1002/jgrd.50166>.

United States Environmental Protection Agency	Office of Air Quality Planning and Standards Air Quality Assessment Division Research Triangle Park, NC	Publication No. EPA-454/B-24-011 October 2024
---	---	--
