Database for Air Quality and Noise Analysis (DANA) Tool Version 2.0 User Guide

January 2023

U.S. Department of Transportation
Federal Highway Administration

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Quality Assurance Statement

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. The FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HEP-23-007		

4. Title and Subtitle

Database for Air Quality and Noise Analysis (DANA) Tool Version 2.0 User Guide

5. Report Date

January 2023

6. Performing Organization Code

7. Author(s)

Sophie Kaye, Aaron Hastings, and William Chupp, Volpe Center; Richard A. Margiotta and Richard Ge, Cambridge Systematics, Inc.

8. Performing Organization Report No.

9. Performing Organization Name and Address

U.S. Department of Transportation
John A. Volpe National Transportation Systems Center
Office of the Secretary of Transportation - Research
Environmental Measurement and Modeling Division, V-324
Cambridge, MA 02142-1093

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

HW58F119

12. Sponsoring Agency Name and Address

U.S. Department of Transportation
Federal Highway Administration
Office of Planning, Environment, and Realty
1200 New Jersey Avenue, SE
Washington, DC 20590

13. Type of Report and Period Covered

Final Report

14. Sponsoring Agency Code

HEP

15. Supplementary Notes

FHWA Review Team: David Kall (COTR), Cecilia Ho, Aileen Varela-Margolles

16. Abstract

The Database for Air Quality and Noise Analysis (DANA) tool version 2.0 provides traffic-related inputs to the Motor Vehicle Emission Simulator (MOVES) vehicle emissions model and the Traffic Noise Model Aide (TNMAide). DANA provides real-world measurements of traffic conditions for use in environmental analyses. In the past, analysts relied almost exclusively on transportation models to generate base year traffic data, an often-cited shortcoming of transportation emission and noise analyses. By having these data already compiled, environmental analysts are spared the task of assembling the data.

DANA creates two air quality-related data sets and a combined traffic data for a spreadsheet-based noise tool, TNMAide. The datasets are:

- Link-Level Data Set — Three output files are provided: 1) hour-by-hour detailed output file (parquet format) with combined traffic data and emission rates by vehicle type for each NPMRDS link (NHS roadways); 2) sample output file (csv format) with the first 1000 and last 1000 rows of the detailed output data; and 3) summary output file (csv format) with average speed over the specified time period, the average daily traffic, and the cumulative emissions over the specified time period of each of the pollutants (CO, NOX, VOC, PM2.5, PM10, CO2, CH4, N2O, and CO2eq) for each NPMRDS link
- MOVES County-Level Input Data Set—the following MOVES input types are provided: Average Speed Distribution; Vehicle Type VMT, Road Type Distribution; Hour VMT Fraction; Day VMT Fraction; and Month VMT Fraction.
- Traffic data summaries for TNMAide—TNMAide is a separate spreadsheet tool that provides: Single Noise Worst Case Hour Analysis and 24-Hour Traffic Distribution for Noise Analysis based on one year of real-world traffic data provided by the DANA Tool.

This volume is the first in a series. The other volume in the series is:

- FHWA-HEP-21-021 Volume II: TNMAide – Technical Description and User's Guide

17. Key Words	18. Distribution Statement
Traffic Data, Vehicle classification data, HPMS data, National Performance Management Research Data Set, Travel Monitoring and Analysis System, Transportation noise analysis, MOVES county level inputs, emissions analysis, Speed distributions, emission rates	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	N/A

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

23CFR772	Title 23, Part 772 of the Code of Federal Regulations
AADT	Annual Average Daily Traffic
AMS	Analysis, Modeling, and Simulation
ATDM	Active Transportation and Demand Management
CMAQ	Congestion Mitigation and Air Quality Improvement Program
CO	Carbon Monoxide
CRC	Coordinating Research Council
DANA	Database for Air Quality and Noise Analysis
DMA	Dynamic Mobility Applications
DOT	Department of Transportation
EPA	Environmental Protection Agency
FHWA	Federal Highway Administration
FRA	Federal Railroad Administration
GIS	Geographic Information System
GLRTOC	Great Lakes Regional Transportation Operations Coalition
GUI	Graphic User Interface
HPMS	Highway Performance Monitoring System
K-S	Kolmogorov-Smirnov
Ldn	Day-night average sound level
Lden	Day-evening-night average sound level
LOS	Level of Service
MAP-21	Moving Ahead for Progress in the 21st Century Act
MOVES	MOtor Vehicle Emission Simulator
MPO	Metropolitan Planning Organization
NCHRP	National Cooperative Highway Research Program
NEI	National Emissions Inventory
NHS	National Highway System
NO_x	Oxides of nitrogen, a collective term for all compounds of nitrogen and oxygen, including nitrogen monoxide and nitrogen dioxide.
NPMRDS	National Performance Management Research Data Set

$PM_{2.5}$	Particulate matter with a diameter less than 2.5 micrometers
PM_{10}	Particulate matter with a diameter less than 10 micrometers
REMEL	Reference energy mean noise emission level
TMAS	Traffic Monitoring Analysis System
TMC	Traffic Message Channel
TNM	Traffic Noise Model
TNMAide	Traffic Noise Model Aide
TRB	Transportation Research Board
UCR	Urban Congestion Report
VHT	Vehicle-Hours Traveled
VMT	Vehicle-Miles Traveled
VOC	Volatile Organic Compound

1 DATABASE FOR AIR QUALITY AND NOISE ANALYSIS TOOL VERSION 2.0 USER MANUAL

1.1 BACKGROUND

Database for Air Quality and Noise Analysis (DANA) is a tool created by the Federal Highway Administration (FHWA) to combine traffic data from existing data sources into a single database and process the combined data into properly formatted inputs to the United States Environmental Protection Agency's (EPA) Motor Vehicle Emission Simulator (MOVES) model and the FHWA's Traffic Noise Model Aide (TNMAide). DANA version 1.0 was released in July 2021 and was the result of more than two years of research on "National Traffic Dataset Applications for Air Quality and Noise Analysis." Data updates for DANA 1.0 were released in March 2022 and included 2020 TMAS and HPMS data. DANA version 2.0 was released in January 2023 and is the current release documented in this user guide. Improvements for DANA 2.0 focused on improving the accuracy of the calculations, but it also included usability improvements, bug fixes, and the latest TMAS data (2021). See Section 1.2 for a full change log.

DANA provides real-world measurements of traffic conditions for use in environmental analyses. In the past, analysts relied almost exclusively on transportation models to generate base year traffic data, an often-cited shortcoming of transportation emission and noise analyses. By having these data already compiled, environmental analysts are spared the task of assembling the data. Finally, DANA helps ensure that environmental analyses use a consistent set of traffic data and processing methods across the entire country. FHWA provides the DANA tool as a resource to stakeholders and use of the tool is voluntary; however, using it may not satisfy all regulatory requirements.

Some possible uses for the DANA tool include:

- Assisting in completing noise analyses for NEPA documents,
- Planning analyses to identify highway projects and mitigation measures, and
- MOVES county-level runs completed for various purposes (mobile source air toxics analysis, greenhouse gas analysis, etc.).

Note that the DANA tool processes historical traffic data and does not include the capability to forecast traffic for future years.

The DANA tool integrates the National Performance Management Research Data Set (NPMRDS), Highway Performance Monitoring System (HPMS), and Travel Monitoring Analysis System (TMAS) national traffic datasets into a single composite dataset, which is useful for many applications, but is specifically used to produce:

1. **Link-Level Data Set**—For every NPMRDS link (National Highway System (NHS) roadways) for every hour of the specified time period, the following traffic data are compiled: speeds and travel times for all vehicles, passenger vehicles, and trucks; the hourly percent of the MOVES vehicle types; the hourly percent of TNM vehicle types; and total VMT. Emission rates (grams-per-mile) for VOC, CO, NO_x, PM_{10}, $PM_{2.5}$, as well as the four greenhouse gas pollutants (CO_2, CH_4, N_2O, and CO_2 equivalent) for each vehicle type are

also included; these MOVES-based emission rates come from EPA's National Emissions Inventory (NEI)[1] and use the representative county approach to reflect local conditions.

2. **MOVES County-Level Input Data Set**—For the counties provided in the inputs, the following MOVES input types are produced: Average Speed Distribution; Vehicle Type VMT, Road Type Distribution; Hour VMT Fraction; Day VMT Fraction; and Month VMT Fraction.

3. **Traffic data summaries for TNMAide**—TNMAide is a separate spreadsheet tool developed as part of the same research project. Full-year traffic data summaries (including annual average daily traffic (AADT), speeds, and percent distributions by TNM vehicle type) output by the DANA tool serve as inputs to TNMAide. TNMAide provides: Single Noise Worst Case Hour Analysis and 24-Hour Traffic Distribution for Noise Analysis based on one year of real-world traffic data provided by the DANA tool. More information can be found in the document entitled "TNMAide: Technical Description and User Guide."

[1] 2017 NEI based on MOVES2014b

1.2 CHANGE LOG

The following updates were made between DANA Tool versions 1.0 and 2.0:

1.2.1 Calculation accuracy

1.2.1.1 Process 0

- Adds 'route_id' as a unique parameter to search for when filling in HPMS fields during the TMAS stations/HPMS link matching process, increasing the number of TMAS stations included in the final TMAS datasets distributed with the DANA tool
- Improved memory usage to allow Process 0 to run with large TMAS source data files

1.2.1.2 Process 1

- Improved Tier 1 matching between NPMRDS and TMAS data by shrinking the match region to be more precise
- Improved Tier 2 matching between NPMRDS and TMAS data by removing leading zeros to facilitate roadway name string matching
- Now uses vehicle type-specific NPMRDS speed data rather than all vehicle average speed

1.2.1.3 Process 2

- Now uses VMT-based HPMS data for roadway classes 6 and 7 where available, rather than roadway mileage
- Now aggregates to MOVES road types later in calculations to retain higher level of detail

1.2.2 Usability improvements

- Added capability to read HPMS shapefiles directly, without having to reformat
- Changed format of output data from processes 2 and 3 to match MOVES templates
- Added output data directory selection
- Added output data sub-directories for each process
- Added date selection option for processing specific time periods less than one year
- Added automatic text file export of Progress Log tab content
- Changed NEI2017_RepresentativeEmissionsRates file from .csv to .parquet for more manageable file size and quicker processing

1.2.3 Bug fixes

- Added files missing from lib folder
- Fixed Process 1 Progress Log typographical errors
- Identified and referenced proper input data columns to facilitate Tier 2 matching between NPMRDS and TMAS data in Process 1

1.2.4 New input data provided[2]

- 2015-2020 TMAS data that have been updated to reflect Process 0 improvements
- 2021 TMAS data

[2] 2020 TMAS, HPMS County Road Mileage, and Statewide Functional Class VMT provided as part of the February 2022 data update (no accompanied changes to the DANA Tool)

1.3 SETUP

1.3.1 Computer Requirements

Given the large dataset that the tool processes, high-capacity hardware should be used to run the DANA tool. The following hardware configuration is recommended for running the tool:

- High-capacity computer including a powerful computer processing unit (CPU) (e.g., Intel Xeon) and at least 16 GB of free random access memory (RAM) [3]
- Sufficient hard disk space to store output data
- Windows Operating System

1.3.2 Installation

There are 9 total installers[4] to unpack the DANA tool and associated input files:

- DANA_Installer.zip, which will install the DANA tool executable and some of the smaller default input files
- 7 different DANA_TMAS[year]_Installer.zip files, which will extract the larger TMAS input data files within the proper folder structure[5] ("year" indicates the year represented by the TMAS data in that installer, 2015-2021 data now included)
- DANA_EXAMPLE_DATA_Installer.zip, which will extract sample data from Middlesex County, MA to use for self-training on the DANA tool

The order in which the installers are executed has no effect on their functionality. All installers can be accessed here: https://www.fhwa.dot.gov/environment/air_quality/methodologies/dana. There is a potential for your network's anti-virus software to block the installer executable(s), in which case you should contact your IT department to unblock the executable(s).

Once downloaded, double clicking on any installer will launch a similar setup wizard, exemplified for DANA_Installer.exe in Figure 1. The first screen will ask if the tool should be installed for the current user or all users. Installing for all users will change the installation directory and may require administrative privileges.

[3] Ideal specification would be 32 GB or more RAM

[4] The sample data is only for self-training and is not required if users download their own local HPMS and NPMRDS data.

[5] See Section 1.5.2 for details

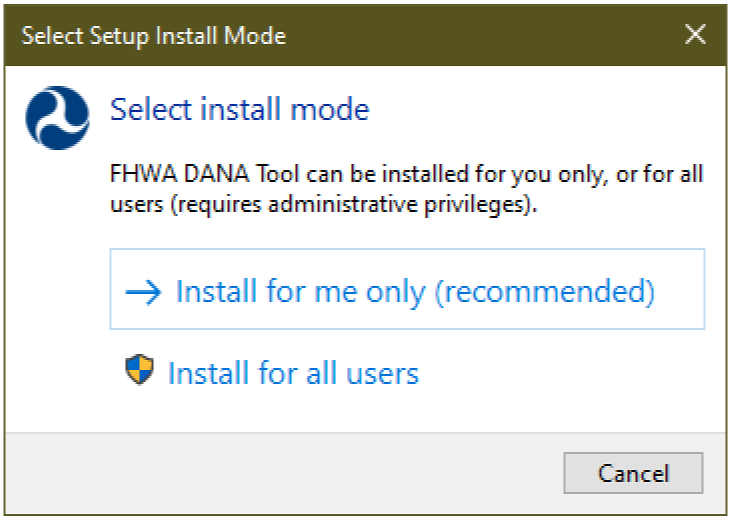


Figure 1. Screenshot. Setup wizard opening window.

After choosing for whom to install the tool, the next screen shows the installation directory, as in Figure 2. If satisfied with the default extraction directory, click “Next”. Otherwise, choose an alternate directory by clicking “Browse…” before clicking “Next”. Note that if one of the other installers has been run previously, a message may appear to warn that the installation folder already exists. This can safely be closed by clicking “Yes” to continue installing in the existing DANA Tool installation folder.

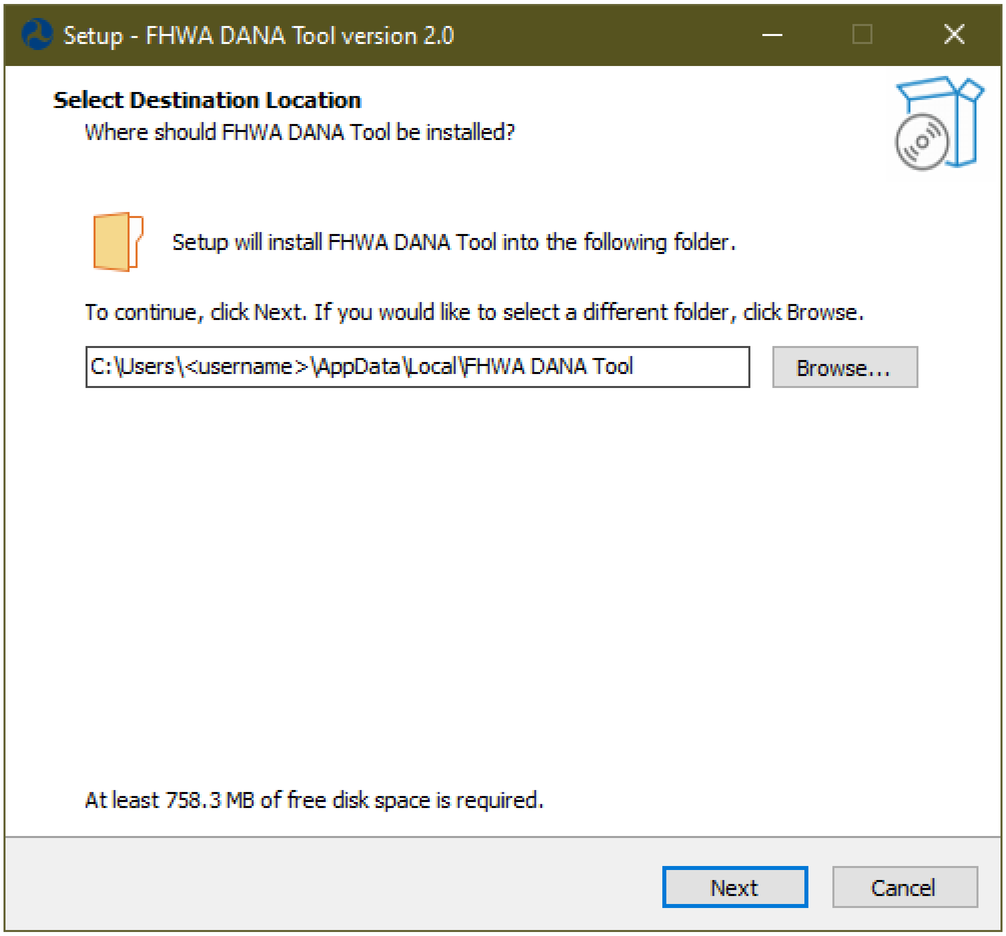


Figure 2. Screenshot. Setup wizard installation directory.

The following window will prompt the user to decide whether to create a desktop shortcut, as shown in Figure 3. Check the box to create a shortcut or leave the box unchecked to skip the step during the installation. Note that this decision point will only appear for the DANA Tool installer, not for the input data installers. Click “Next” to continue.

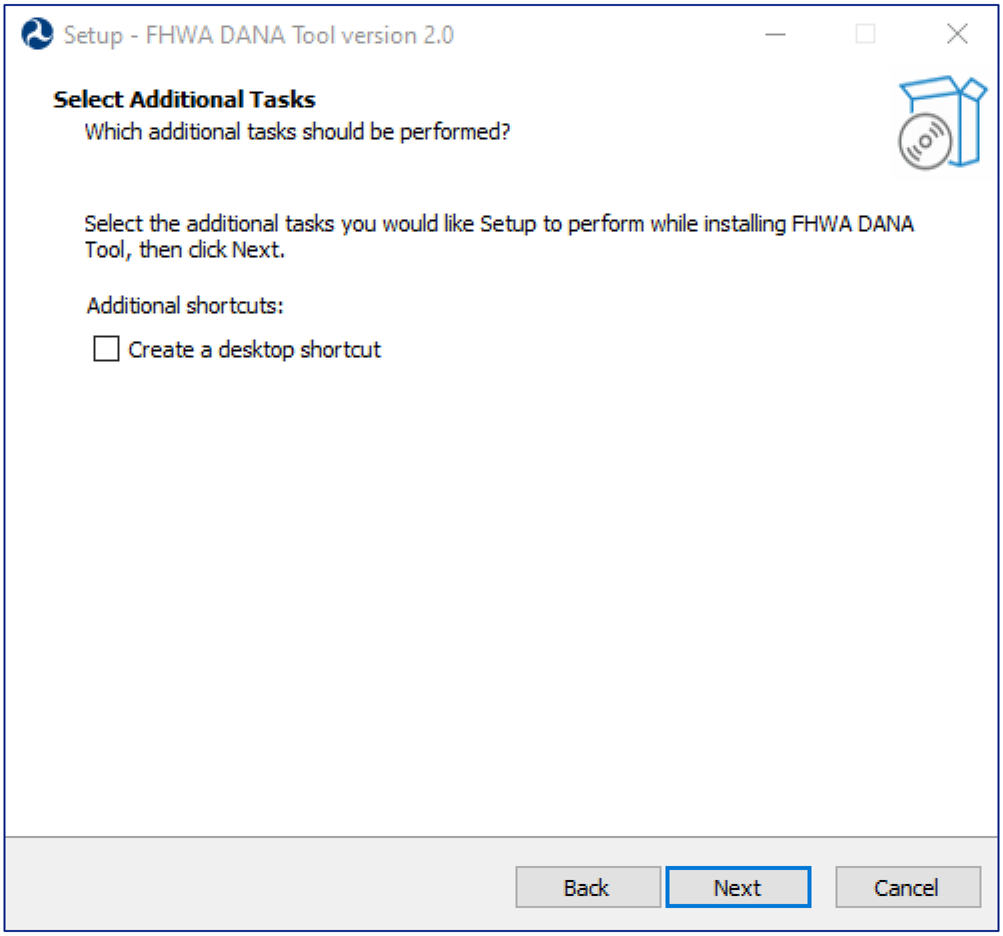


Figure 3. Screenshot. Setup wizard Select Start Menu Shortcut Folder window.

The following window reviews your selections from the previous two windows, as shown in Figure 4. If dissatisfied, click "Back" to return to a previous window. Otherwise, click "Install".

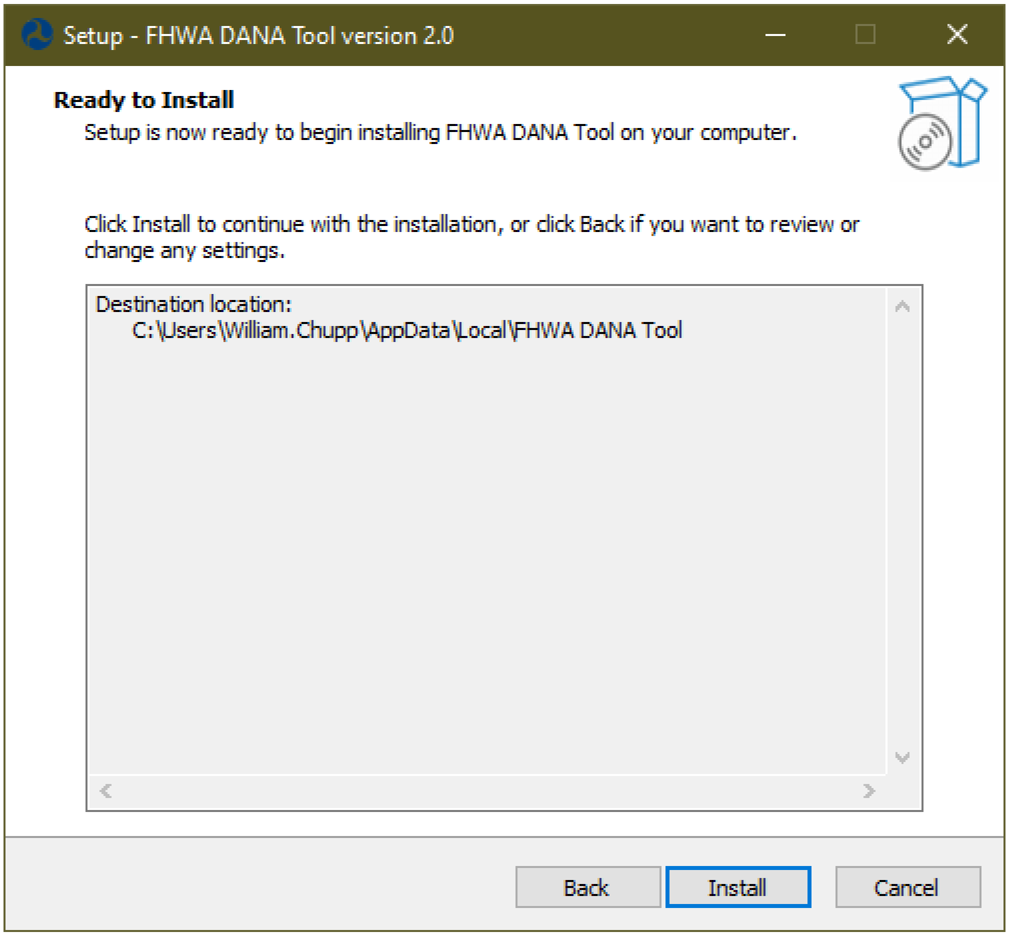


Figure 4. Screenshot. Setup wizard Ready to Install window.

The following window displays a progress bar of the installation and/or unpacking process, as shown in Figure 5.

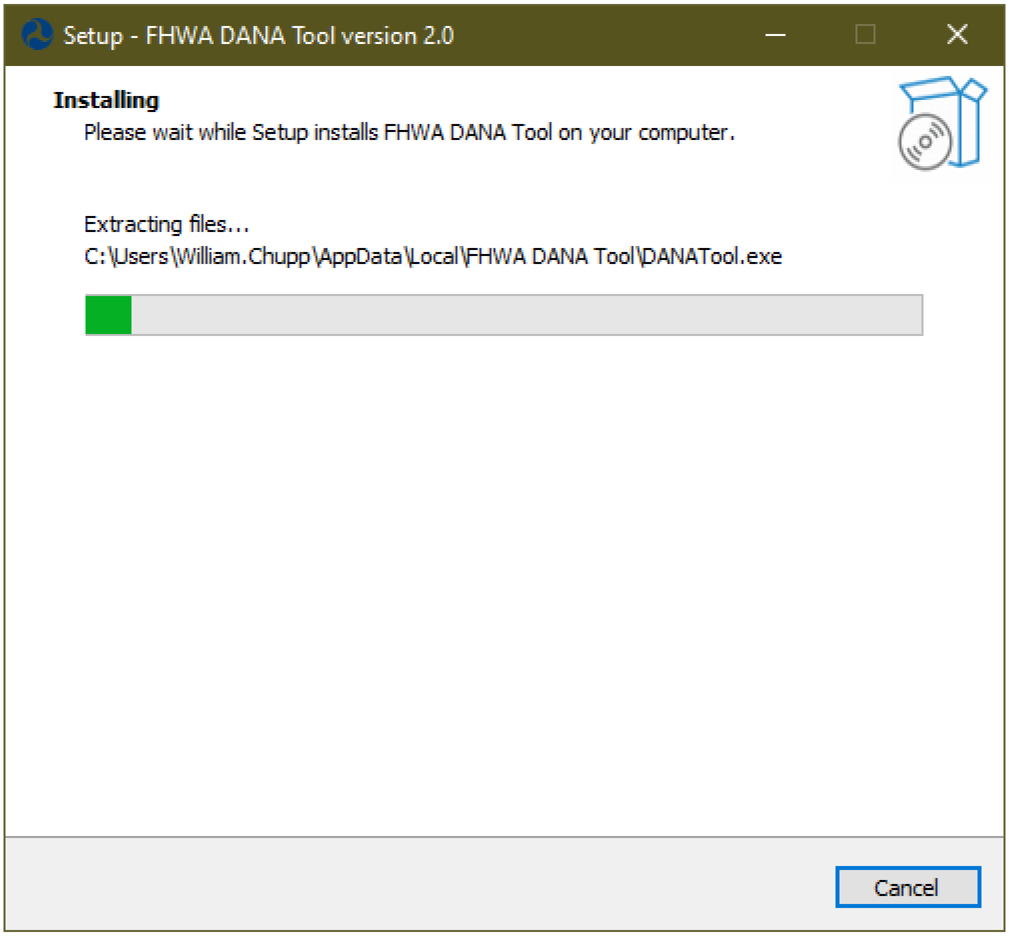


Figure 5. Screenshot. Setup wizard Installing window.

When installation and unpacking have finished, a message will display, as shown in Figure 6. Click “Finish” to close the setup wizard.

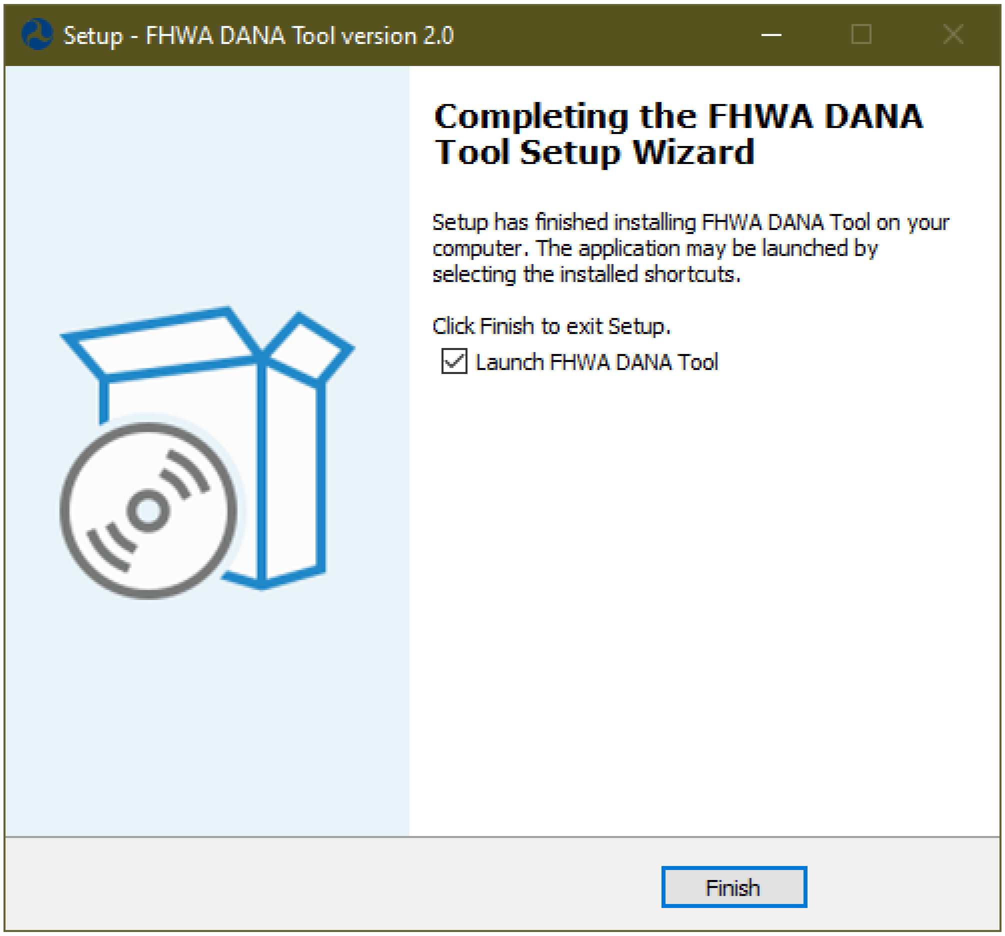


Figure 6. Screenshot. Setup wizard final window.

1.4 GRAPHIC USER INTERFACE OPERATIONS

1.4.1 Operating the Main GUI

The following steps should be used to operate the main GUI:

1. Double click the DANA tool executable to launch the main GUI, as shown in Figure 7.
2. Specify the desired directory for the output files, as shown in Figure 8. Note that the file path chosen cannot contain any spaces.
3. Select which state[6] to analyze from the dropdown menu, as shown in Figure 9.
4. If running Process 1, choose whether the output data should reflect the entire temporal extent of the NPMRDS input data by maintaining the default first radio button selection ("Auto-detect date range…"), or if the output data should reflect a subset of the temporal extent of the NPMRDS input data by choosing the second radio button option ("Or, select a date range…") and selecting a valid date range, as illustrated in Figure 10. See Section 1.6.3 for more details.
5. Locate the input data files for the desired process by clicking the selection buttons in the relevant process section of the GUI, exemplified in Figure 11.[7] If one or more of the required input files are detected from a previous run, the available file(s) will automatically appear to the right of the corresponding selection button(s).[8]
6. Click the "Run Process X" button at the top of each process section of the GUI to execute that process.

Initiating a process will open the Progress Log tab (Figure 12) and the user can navigate between tabs via click at any time. While a process is running, all "Run Process X" buttons on the Data Processing tab will be disabled (as indicated by their greyed out appearance) and a new "Process X Running" status message will appear to the right of the "Cancel Process X" button in the relevant GUI section (Figure 13). If desired, processing can be halted mid-computation by clicking the "Cancel Process X" button.

After a process has finished running, the GUI will not advance to the subsequent process on its own. The user must initiate each process using the respective buttons, after selecting all required input files. Internal checks are implemented to ensure all required input files are provided. Warning reports will appear to prompt for missing selections, as demonstrated in Figure 14.

---

[6] The DANA Tool only contains data for 50 states and D.C. Puerto Rico and other territories are not found in the dropdown list.

[7] See Section 1.5 for details.

[8] See Section 1.4 for details.

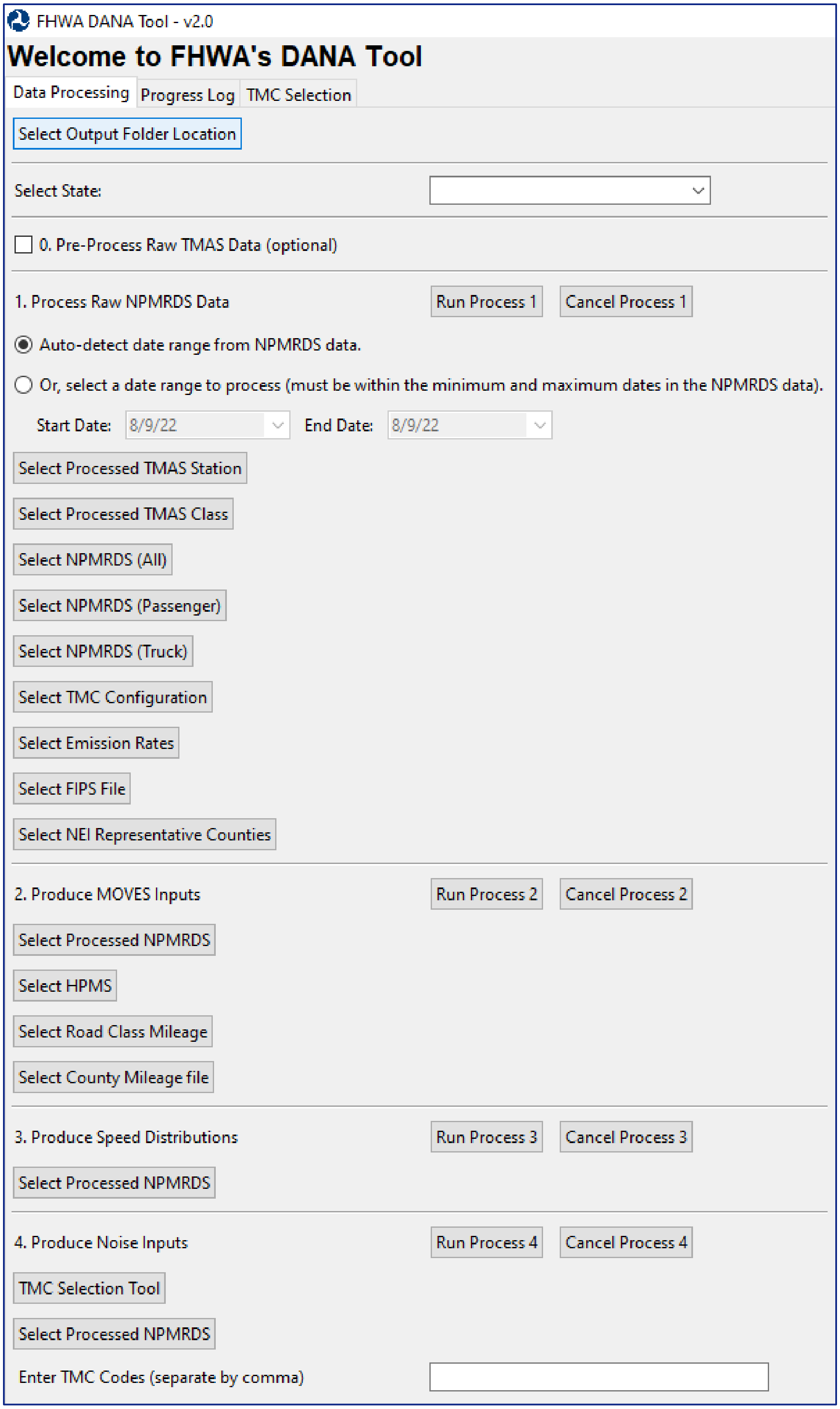


Figure 7. Screenshot. Main graphic user interphase

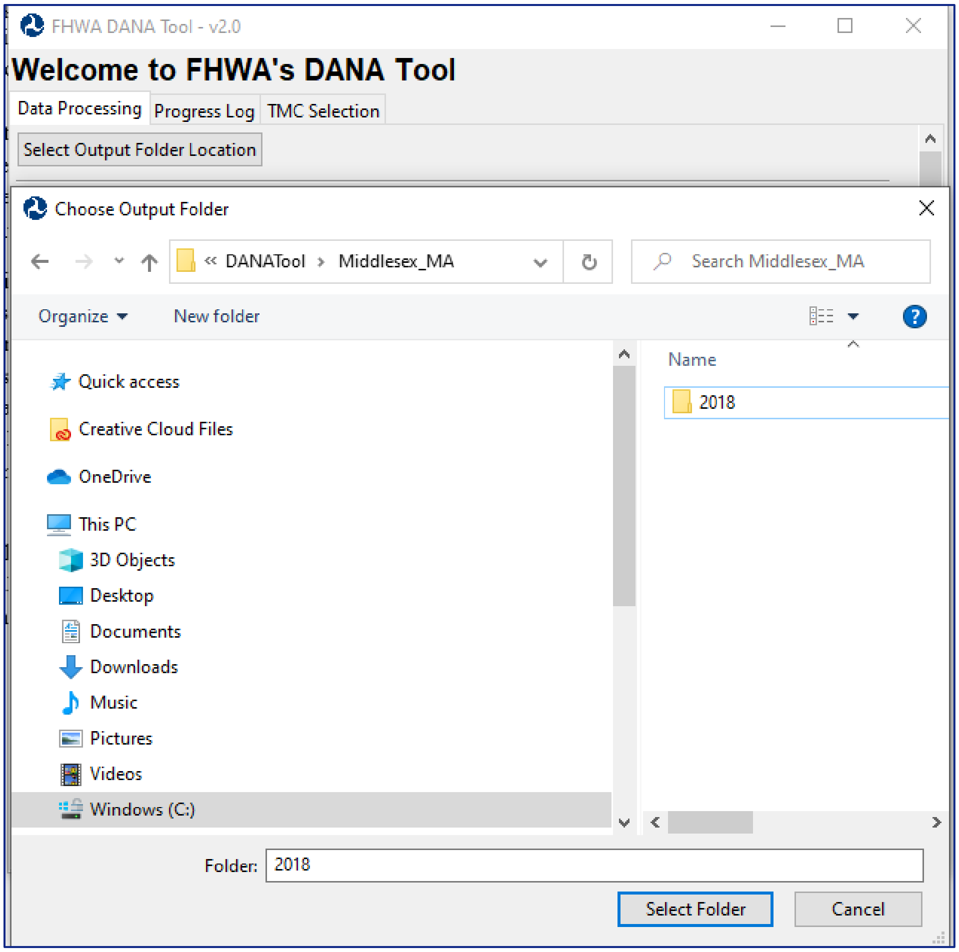


Figure 8. Screenshot. Output folder selection

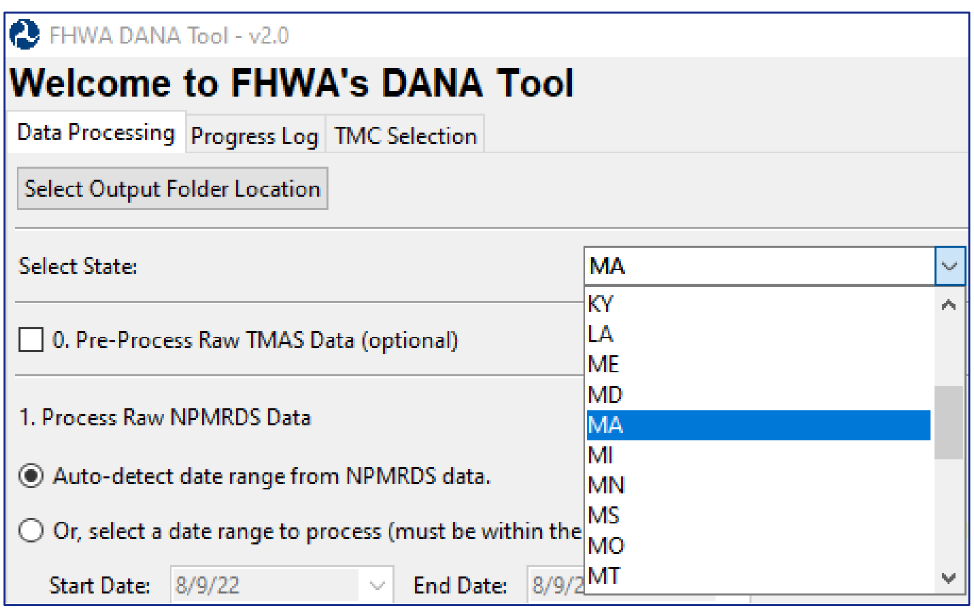


Figure 9. Screenshot. State selection

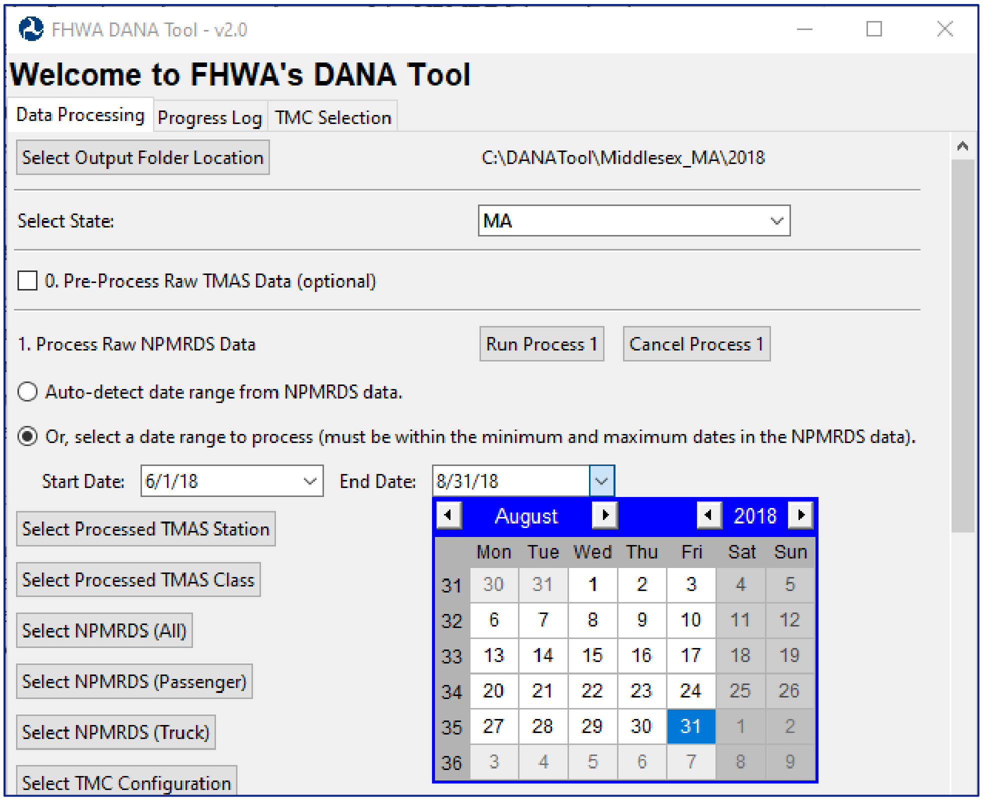


Figure 10. Screenshot. Date range selection

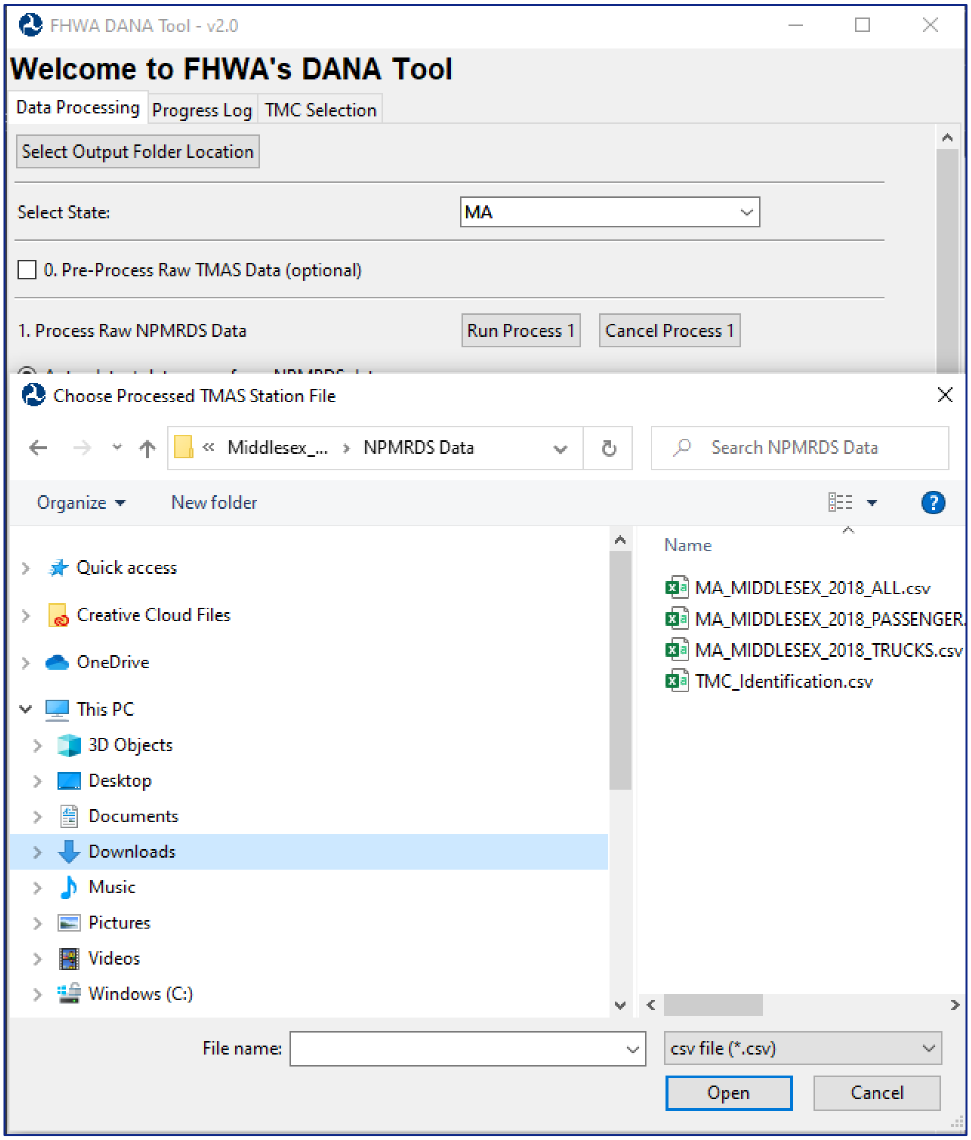


Figure 11. Screenshots. Input file selection

```
********** Produce MOVES VMT Inputs **********
Processing State VMT data
Reading in State HPMS
Reading in Highway Statistics VM2
Reading in HPMS County Rd Mileage
Processing VMT for functional system 1-5
Processing VMT for rural functional system 6 and urban and rural 7
  took: 3.209
Reading Composite Dataset
  took: 118.525
Developing monthly VMT Fractions dataset
  took: 8.070
Developing daily VMT Fractions dataset
  took: 9.595
Developing hourly VMT Fractions dataset
  took: 3.822
Developing Regional VMT summaries
  took: 1.504
Developing RoadType VMT summaries
  took: 2.284
Outputs saved in C:/DANATool/Middlesex_MA/2018/Process2_MOVES_VMT_Distributions/
**********Process Completed**********

*** 2. Produce MOVES Inputs has finished running ***
```

Figure 12. Screenshot. Progress Log GUI tab with script status and runtime messages

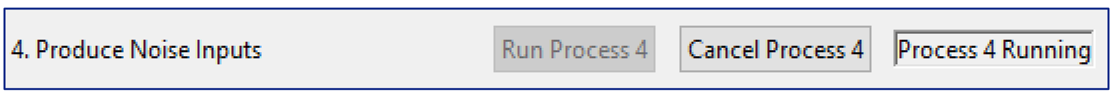


Figure 13. Screenshot. DANA GUI appearance during processing

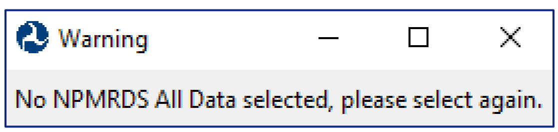


Figure 14. Screenshot. DANA missing input file warning

It is possible to operate the DANA Tool without a mouse using the following keyboard commands:

- Tab – move between the items in the tool (buttons, drop-down lists, etc.)
- Down Arrow – open a drop-down list
- Up and down arrows – move between items in a drop-down list
- Enter – select an item in a drop-down list
- Spacebar – press a button or check a checkbox
- Esc – close a drop down list or file browser without making a selection
- Ctrl +Tab – move between tabs at the top of the tool; navigate out of the progress log back to the tabs at the top

1.4.2 Progress Log Tab Output

Script running status and runtime messages will be displayed in the Progress Log tab of the main GUI, as exemplified in Figure 12. At the completion or termination of each process, the contents of this tab are exported to progress_log.txt in the root folder of the selected output file directory. Note that this file is cumulative such that it encompasses the progress for every process run while the DANA Tool is open. The messages generally show the step currently executing from the process script and time spent processing the previous steps. The actual processing times will vary based on the process, county input data size, and the available computational resources of the user computer.

Occasionally, warning messages may appear, signified by phrases such as “UserWarning”, “FutureWarning”, or “SettingWithCopyWarning”. While the warning messages should be reviewed, they will not cause errors in the data processing. However, error messages signified by phrases such as “IndexError”, ValueError”, “KeyError”, or “ParserError” should be investigated, as they likely halt data processing or indicate improper output data. An example error message indicating halted data processing is shown in Figure 15.

```
********** Produce MOVES VMT Inputs **********
Processing State VMT data
Reading in State HPMS
Reading in Highway Statistics VM2
Traceback (most recent call last):
  File "NTD_05_main_GUI.py", line 280, in process_handler
  File "lib\NTD_02_MOVES.py", line 91, in MOVES
IndexError: index 0 is out of bounds for axis 0 with size 0

*** 2. Produce MOVES Inputs has finished running ***
```

Figure 15. Screenshot. Progress Log GUI tab with error message

In some cases, Python does not free up used memory between runs of the DANA tool. This is indicated by an error message that states “C error: out of memory”. In such situations, close the DANA tool GUI to completely free up the RAM and restart.

1.4.3 Technical Support

All inquiries regarding the tool itself and the input data should be directed to DANAhelp@dot.gov for dissemination to the proper point of contact based on the issue.

1.5 DATA FLOW

The DANA tool consists of seven production scripts and several data files in a set of folders. The general data processes and associated Python scripts are as follows:

Table 1. Data Processes and Pythons Scripts

Process #	Process Name	Script Name
0	Process Raw TMAS Data (optional)	NTD_00_TMAS.py
1	Process Raw NPMRDS Data	NTD_01_NPMRDS.py
2	Produce MOVES Inputs	NTD_02_MOVES.py
3	Produce Speed Distributions	NTD_03_SPEED.py
4	Produce Noise Inputs	NTD_04_NOISE.py
N/A	Run Graphic User Interface	NTD_05_main_GUI.py
N/A	Run TMC Selection Tool	NTD_06_selection_GUI.py

The actual processes executed will be based on the user needs and available data e.g., if processed NPMRDS data are available in the default output folder from a previous run, then Process 1 is not required in order to run the subsequent processes. The complete script dependencies are displayed in Table 2.

Table 2. Process Dependencies.

Process	Input File Required	Prior Process Prerequisite
0. Process Raw TMAS Data (optional)	N/A	N/A
1. Process Raw NPMRDS Data	Processed TMAS Station/Class	N/A*
2. Process MOVES Inputs	Processed TMAS Class	Process 1*
3. Produce Speed Distribution	Processed NPMRDS	Process 1*
4. Produce Noise Inputs	Processed NPMRDS	Process 1*

*If user-defined geographically-specific or newly updated TMAS data are used rather than the pre-processed national TMAS dataset provided with the DANA tool, then Process 0 would also be a prerequisite. See Appendix A. Process 0: Process Raw TMAS Data (optional) for further instruction.

Each process requires different input data in their given file types. Completing Processes 1-3 creates inputs for the MOVES application, whereas completing Processes 1 and 4 create inputs for TNMAide. All processes including default and externally downloaded input data as well as Process 1 outputs used as inputs for subsequent processes are demonstrated in Figure 16.

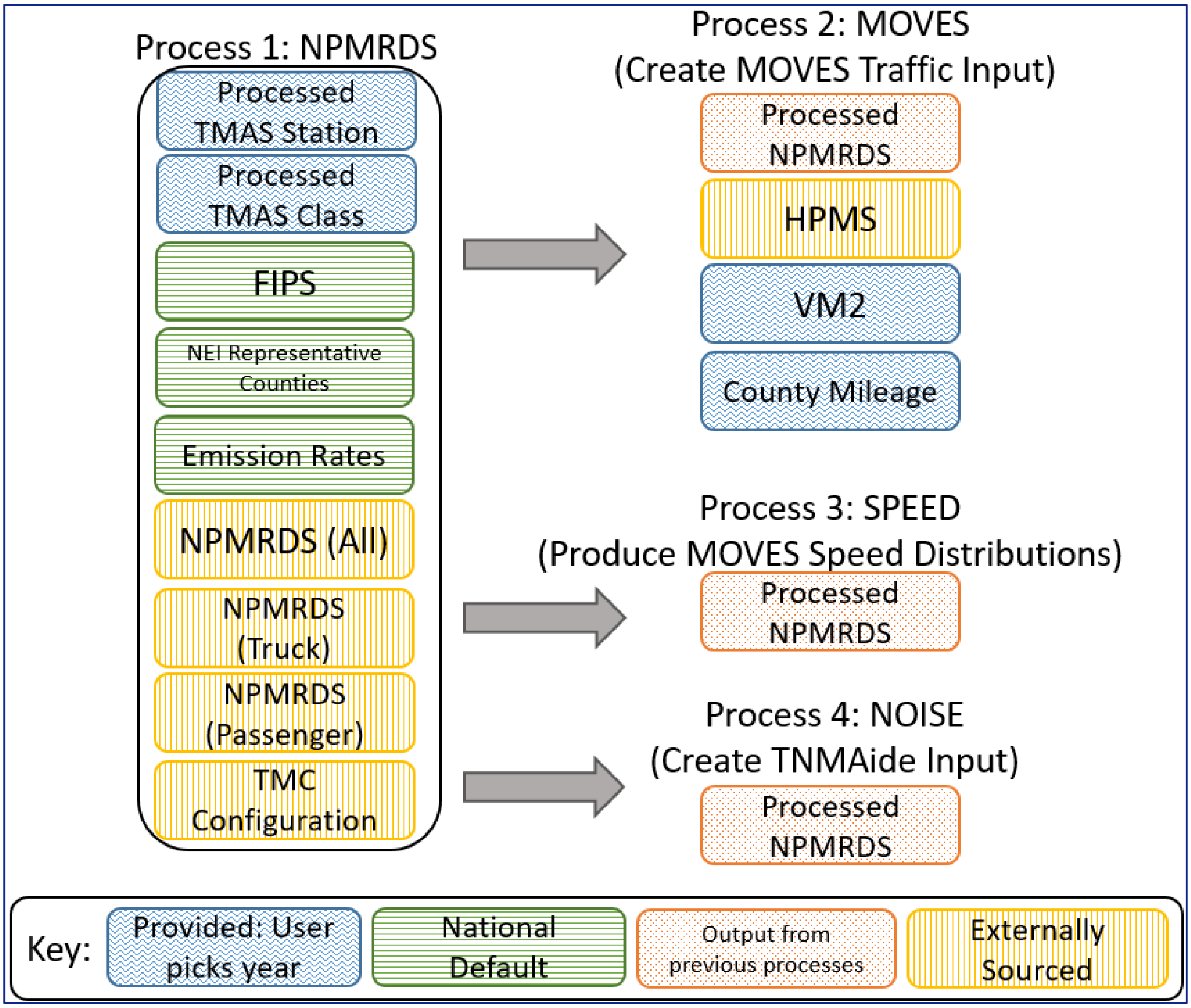


Figure 16. Flowchart. DANA tool input data flow

Figure 17 and Figure 18 show the processing steps involved in the creation of the composite dataset used as the input for Processes 2-4. Figure 17 illustrates Process 0, in which the TMAS data are parsed to decode the vehicle classification data format specified in FHWA's Traffic Monitoring Guide and are combined with the TMAS station and classification count data. Pre-processed TMAS data are provided with the DANA tool and can be found in the Default Input Files\TMAS Data\ directory.

Figure 18 illustrates Process 1, in which the composite link-level dataset is created by joining NPMRDS speed data[9] (which also contain the key HPMS data elements) and TMAS classification count data for the selected state with the national MOVES emission rates. Note that if TMAS data for the selected state and year are unavailable, DANA defaults to the national average TMAS data for that year. HPMS conflated data elements contained in the NPMRDS TMC configuration file are enumerated in Table 3.

[9] Note that the NPMRDS reported speed values represent a harmonic average of all the vehicles logged during that time period. i.e., $n/(1/x_1 + 1/x_2 + 1/x_3 \ldots 1/x_n)$

Table 3. HPMS conflated attributes contained in NPMRDS data

TMC Attributes	HPMS Attributes
tmc	f_system
road	urban_code
direction	faciltype
intersection	structype
state	thrulanes
county	route_numb
zip	route_sign
start_latitude	route_qual
start_longitude	altrtename
end_latitude	aadt
end_longitude	aadt_singl
miles	aadt_combi
road_order	nhs
timezone_name	nhs_pct
type	strhnt_typ
country	strhnt_pct
tmclinear	truck
frc	thrulanes_unidir
border_set	aadt_unidir
isprimary	aadt_singl_unidir
active_start_date	aadt_combi_unidir
active_end_date	

The NPMRDS/TMAS join in Process 1 is performed in four successive "tiers," each one more aggregate than the previous tier. Because the number of NPMRDS links far outnumber the number of TMAS vehicle classification locations, matching is based on the following hierarchy:

- Tier 1: exact roadway match
- Tier 2: state, county and route
- Tier 3: state, urban/rural, and roadway functional class[10]
- Tier 4: national, urban/rural, and roadway functional class

[10] Roadway functional classes are denoted 1-7, which represent the following categories, in ascending order: interstates, freeways and expressways, principal arterials, minor arterials, major collectors, minor collectors, and local roads. Note that DANA does not account for operations on non-roads such as driveways, parking lots, etc. so one cannot use DANA inputs to run MOVES analyses that include "off-network" operations.

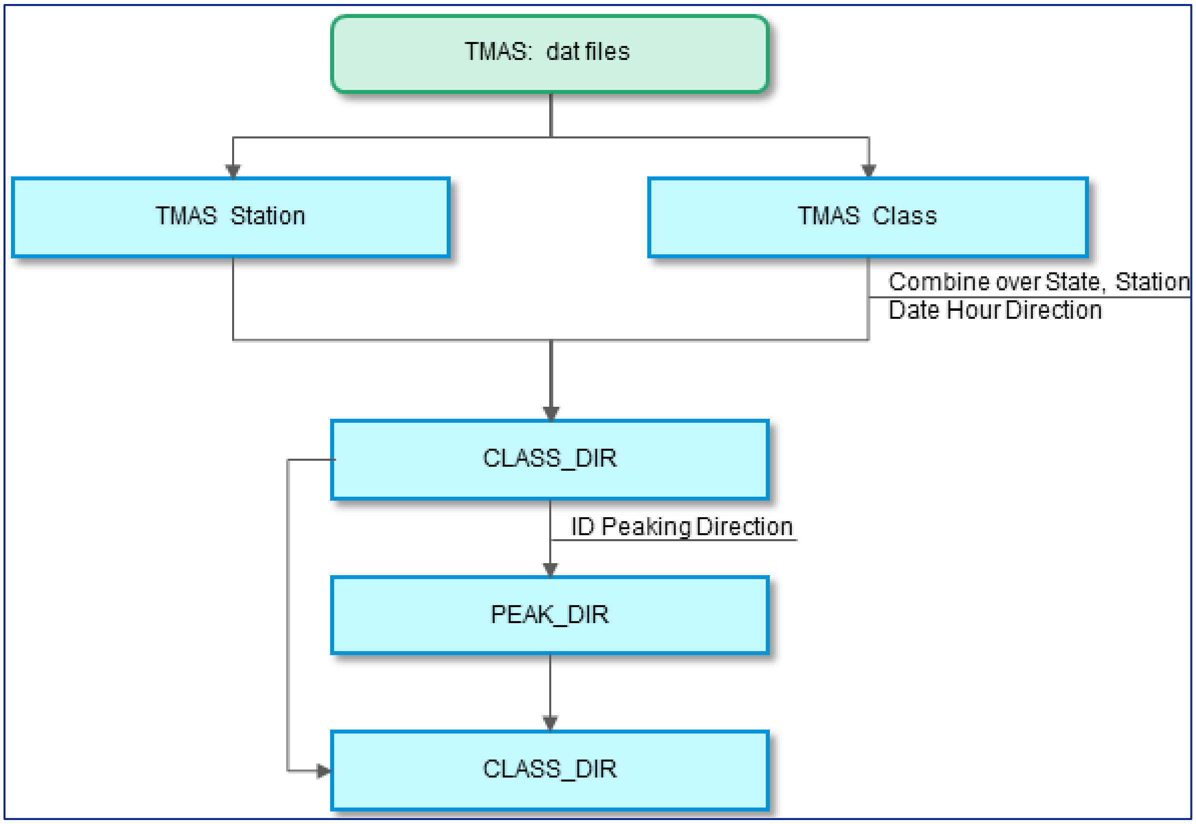


Figure 17. Flowchart. Initial stage composite dataset processing (Process 0).

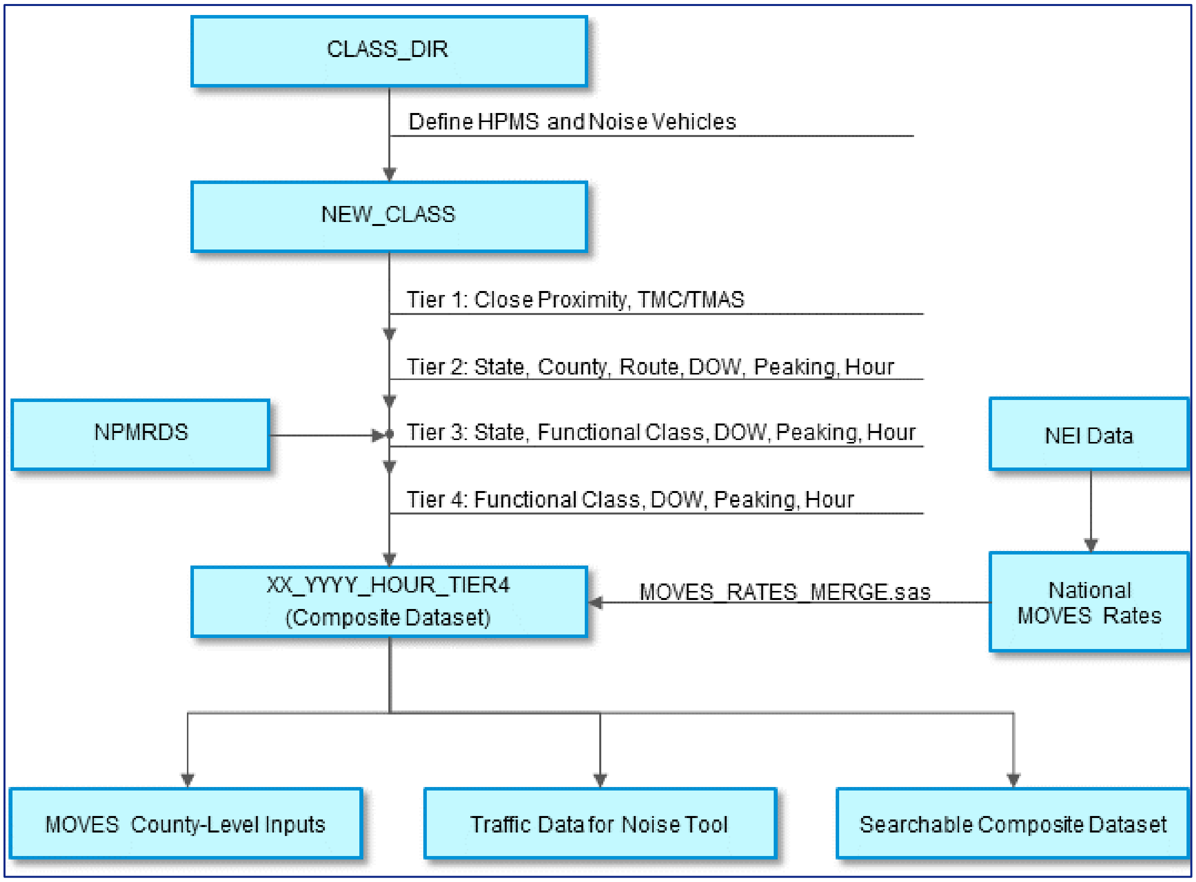


Figure 18. Flowchart. Final stage composite dataset processing (Process 1).

To account for varying traffic patterns throughout the week, the Process 1 output data include the distinction of weekdays from weekends, in which holidays[11] are considered weekends. Process 1 output includes the NPMRDS passenger car and truck speed column in the final link-level output dataset.

In Process 2, the class distributions computed using the TMAS data in Process 1 are applied to the HPMS AADT to determine volume by class. Statewide annual VMT data by functional class (Table VM-2 data) are used as a statewide control total, as those data are the most reliable source for all roadways in the continental United States. County level VMT are adjusted such that they sum to this statewide total as a quality control measure for the resultant MOVES data. DANA version 2.0 also includes enhanced traffic count location validation using the reliable station

[11] DANA detects dates for the following federal holidays each year: New Years Day, Martin Luther King Jr. Day, Presidents Day, Memorial Day, July 4th, Labor Day, Columbus Day, Veterans Day, Thanksgiving, Christmas

latitude and longitude data to match with HPMS roadway geographies.[12] DANA obtains AADT and roadway information (e.g., functional class) from HPMS for the following vehicle types:

Table 4. HPMS vehicle types

HPMS Vehicle Code	Vehicle Type
10	Motorcycles
25	Passenger cars
40	Busses
50	Medium trucks
60	Heavy trucks

Process 3 uses the average speeds from the NPMRDS passenger and truck speed columns in each hour of the day to calculate MOVES average speed bin distributions. There are 15 bins, each of which span a range of 5 miles per hour starting at 2.5 miles per hour, with a 16th bin accounting for all speeds 2.5 miles per hour and below. The speed distributions are adjusted by the HPMS vehicle type distributions to improve accuracy. The NPMRDS passenger and truck speed data is applied to the HPMS and MOVES vehicle types, as follows:

Table 5. NPMRDS, HPMS, and MOVES Vehicle Types

NPMRDS Speed Vehicle Type	HPMS Vehicle Code	MOVES Vehicle Types
Passenger	10	Motorcycles (11)
Passenger	25	Passenger cars, passenger trucks, light commercial history (21, 31, 32)
Truck	40	Transit buses, school buses, other buses (41, 42, 43)
Truck	50	Singe unit short and long-haul trucks, refuse trucks, motor homes (51, 52, 53, 54)
Truck	60	Combination short and long-haul trucks (61, 62)

[12] The only TMAS station data used are the station ID and latitude and longitude coordinates. All other roadway information comes from the more reliable HPMS data.

1.6 INPUT/OUTPUT DATA SPECIFICATIONS

Section 1.6.1 provides instruction on combining input data from various years to achieve the desired output. Section 1.6.2 presents a list that details the file structure of the DANA tool including nested folder hierarchies. Sections 1.6.3-1.6.6 summarize the input data required for each process, including instructions on how to obtain and format externally downloaded data. Appendix B. DANA Tool Input Details provides notes on each input file including the type, the file location, source, and download link, if applicable. The tables in Appendix C. Input Data Dictionaries by Process and Appendix D. Output Data Dictionaries by Process describe each dataset by column including the column name, column description, data type, and an example entry.

1.6.1 Yearly Data Latency

A significant amount of offset exists in the availability of the three data sources that comprise the composite dataset:

- TMAS vehicle classification data are available annually 5 months later (e.g., 2019 data are available in May 2020).
- NPMRDS is available on a monthly basis within five business days after the end of the month (e.g., March's data is availability on April 5).
- HPMS is available annually 10-11 months later (e.g., 2019 data are available in November 2020).
- The NPMRDS contractor conflates key HPMS data items to the NPMRDS Search Results Transportation Management Center (TMC) configuration file annually, but the lag is always two years. Currently, the following data matches are available:
    - 2022 NPMRDS network is matched to 2020 HPMS data
    - 2021 NPMRDS network is matched to 2019 HPMS data
    - 2020 NPMRDS network is matched to 2018 HPMS data
    - 2019 NPMRDS network is matched to 2017 HPMS data,
    - 2018 NPMRDS network is matched to 2016 HPMS data, and
    - 2017 NPMRDS network is matched to 2015 HPMS data.

As noted above, the HPMS data that is conflated onto the NPMRDS data downloads will always represent data from 2 years prior to the speed data in the NPMRDS. This makes it difficult to completely align the years for all of the datasets used in the DANA tool processes. Users should decide what combination of input data years to use based on their specific data processing objectives. Table 3 below outlines three different output data objectives and the corresponding DANA tool processes required to produce the desired dataset. Table 4 provides example input data combinations for each objective to illustrate how the input data years impact the output data. These examples are also FHWA's recommendation as a starting point for the best methods to align years. For processes 2 & 3, which provide MOVES county-level inputs, priority is given to aligning class, speed, and VMT data, such that only the speed distributions are weighted using AADT from a previous year. For processes 1 & 4, it is difficult to align class, speed, and VMT data due to the 2-year offset included in the conflation of HPMS data items to the NPMRDS.

Therefore, it is suggested that input data be aligned with the AADT provided in the offset HPMS data inherent to the NPMRDS input.

Table 6. Required DANA tool processes by output data objective

Output Data Objective	Required Processes
Link-Level Emissions Inventory	Process 1
County-Level MOVES Input	Processes 1-3
Link-Level TNMAide Input	Processes 1 & 4

Table 7. Example input data combinations.

<table>
<tr><th>Process</th><th>Input Data Objective</th><th>DANA Tool Inputs</th><th>DANA Tool Outputs</th></tr>
<tr><td>1-3</td><td>Align class, speed, and VMT distributions</td><td>Process 1:
• 2019 NPMRDS data (which includes 2017 HPMS data)
• 2019 TMAS
Process 2:
• 2019 HPMS
• Process 1 outputs
• 2019 Table VM-2
• 2019 county mileage summaries
Process 3:
• Process 1 outputs</td><td>Process 1:
• Link-level data set: 2019 speeds and vehicle class distributions, 2017 VMT
Process 2:
• MOVES county-level VMT input files based completely on 2019 data
Process 3:
• MOVES speed distribution input files based on 2019 speeds, but weighted using 2017 AADT</td></tr>
<tr><td>1 & 4</td><td>Align input data to AADT in HPMS</td><td>Process 1:
• 2019 NPMRDS data (which includes 2017 HPMS data)
• 2017 TMAS
Process 4:
• Process 1 outputs</td><td>Process 1:
• Link-level data set: 2019 speeds; 2017 vehicle class distributions and VMT
Process 4:
• Traffic summaries for TNMAide: 2019 speeds, 2017 vehicle class distributions and AADT</td></tr>
</table>

1.6.2 Folder Structure

The following list details the file structure of the DANA tool including nested folder hierarchies. Bold text indicates a folder and non-bolded text indicates a file. The images to the right of the folders correspond to the categories shown in Figure 16.

FHWA DANA Tool

1. **Default Input Files**
    - a. **HPMS County Road Mileage** — Provided: User picks year
        - i. County_Road_Mileage_2015.csv
        - ii. County_Road_Mileage_2016.csv
        - iii. County_Road_Mileage_2017.csv
        - iv. County_Road_Mileage_2018.csv
        - v. County_Road_Mileage_2019.csv
        - vi. County_Road_Mileage_2020.csv
    - b. **Statewide Functional Class VMT** — Provided: User picks year
        - i. State_VMT_by_Class_2015.csv
        - ii. State_VMT_by_Class_2016.csv
        - iii. State_VMT_by_Class_2017.csv
        - iv. State_VMT_by_Class_2018.csv
        - v. State_VMT_by_Class_2019.csv
        - vi. State_VMT_by_Class_2020.csv
    - c. **TMAS Data** — Provided: User picks year
        - **i. TMAS 2015**
            1. TMAS_Class_Clean_2015.csv
            2. TMAS_Station_2015.csv
        - **ii. TMAS 2016**
            1. TMAS_Class_Clean_2016.csv
            2. TMAS_Station_2016.csv
        - **iii. TMAS 2017**
            1. TMAS_Class_Clean_2017.csv
            2. TMAS_Station_2017.csv
        - **iv. TMAS 2018**
            1. TMAS_Class_Clean_2018.csv
            2. TMAS_Station_2018.csv
        - **v. TMAS 2019**
            1. TMAS_Class_Clean_2019.csv
            2. TMAS_Station_2019.csv
        - **vi. TMAS 2020**
            1. TMAS_Class_Clean_2020.csv
            2. TMAS_Station_2020.csv
        - **vii. TMAS 2021**
            1. TMAS_Class_Clean_2021.csv
            2. TMAS_Station_2021.csv
    - d. FIPS_County_Codes.csv
    - e. NEI2017_RepresentativeCounties.csv — National Default
    - f. NEI2017_RepresentativeEmissionsRates.parquet

2. **Final Output** (Subfolders and files created after running associated processes, not immediately included in folder structure)

Output from previous processes

   a. **Process1_LinkLevelDataset**
      i. tier1_class.csv
      ii. tier2_class.csv
      iii. tier3_class.csv
      iv. tier4_class.csv
      v. XX_Composite_Emissions.parquet, where XX represents the state abbreviation
      vi. XX_Composite_Emissions_SUMMARY.csv, where XX represents the state abbreviation
      vii. XX_Composite_Emissions_SAMPLE.csv, where XX represents the state abbreviation
   b. **Process2_MOVES_VMT_Distributions**
      i. XX_MONTH_VMT
         1. XX_MONTH_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         2. XX_MONTH_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         3. Etc. for all counties within state
      ii. XX_DAY_VMT
         1. XX_DAY_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         2. XX_DAY_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         3. Etc. for all counties within state
      iii. XX_HOUR_VMT
         1. XX_HOUR_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         2. XX_HOUR_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         3. Etc. for all counties within state
      iv. XX_REGIONAL_VMT
         1. XX_REGIONAL_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         2. XX_REGIONAL_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         3. Etc. for all counties within state
      v. XX_ROADTYPE_VMT
         1. XX_ROADTYPE_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         2. XX_ROADTYPE_VMT_YY.0.csv, where XX represents the state abbreviation and YY.0 represents the county code
         3. Etc. for all counties within state
      vi. XX_MONTH_VMT.csv, where XX represents the state abbreviation
      vii. XX_DAY_VMT.csv, where XX represents the state abbreviation
      viii. XX_HOUR_VMT.csv, where XX represents the state abbreviation

        ix. XX_REGIONAL_VMT.csv, where XX represents the state abbreviation
        x. XX_ROADTYPE_VMT.csv, where XX represents the state abbreviation
    c. **Process3_MOVES_Speed_Distributions**
        i. XX_SPEED_DISTRIBUTION.csv, where XX represents the state abbreviation
        ii. **XX_SPEED_DISTRIBUTION**, where XX represents the state abbreviation
            1. XX_SPEED_DISTRIBUTION_YY.csv, where XX represents the state abbreviation and YY represents the county code

    d. **Process4_TNM_AIDE_Inputs**
        i. XX_Composite_Emissions_select.csv, where XX represents the state abbreviation
    e. **TMC_Selection**
        i. TMCs_X_Y_Z.txt, where X represents the county, Y represents the road name, and Z represents the direction selected by the user
3. **Lib**[13] (Folder for supporting function definition files and processing code)
    a. **ShapeFiles**
        i. XXXXX.ext, where XXXX represents each state name and .ext represents the file extensions (each state has a shapefiles .shp and other database formats in this subfolder)
    b. __init__.py (Blank file which allows importing of entire library into client code)
    c. load_shapes.py (utility that allows the DANA tool to load multiple shapefiles into the library and concatenate them into a single shapefile for processing)
    d. NTD_00_TMAS.py
    e. NTD_01_NPMRDS.py
    f. NTD_02_MOVES.py
    g. NTD_03_SPEED.py
    h. NTD_04_NOISE.py
    i. NTD_06_selection_GUI.py
4. **TMAS_Intermediate_Output** (folder created after running optional Process 0, not immediately included in folder structure)

    Output from previous processes

    a. TMAS_class_clean.csv
    b. TMAS_station_State.csv
5. **User Input Files**
    a. Middlesex_MA

    Externally Sourced

        i. HPMS Data
            1. MA_HPMS_2018.csv
        ii. NPMRDS Data

    Externally Sourced

            1. MA_MIDDLESEX_2018_ALL.csv
            2. MA_MIDDLESEX_2018_PASSENGER.csv
            3. MA_MIDDLESEX_2018_TRUCKS.csv
            4. TMC_Identification.csv
6. DANATool.exe (main executable for running DANA)
7. NTD_05_main_GUI.py (Python script for running the graphic user interface)

[13] Between public releases, the most recent version of the Python data processing scripts can be found at https://github.com/VolpeUSDOT/FHWA-DANATool.

8. unins000.dat (database for uninstalling the DANA tool)
9. unins000.exe (executable for uninstalling the DANA tool)
10. unins001.dat (database for uninstalling the TMAS data)
11. unins001.exe (executable for uninstalling the TMAS data)

1.6.3 Process 1: Process Raw NPMRDS and Emission Rate Data

The following inputs are required:

- Radio button selection regarding the date range of the output data with the following options:
  - Default first radio button stating "Auto-detect date range…", which indicates that the output data should reflect the entire temporal extent of the NPMRDS input data
  - Second radio button stating "Or, select a date range…", which indicates that the output data should reflect a subset of the temporal extent of the NPMRDS input data
    - Requires selection of a valid date range using the start and end date calendar selectors, as illustrated in Figure 10. For example, if the NPMRDS input data encompasses all of 2018, a valid date range would include any portion of that year. Selecting a date range that extends beyond 2018 will prompt the following error message in the Progress Log: "ValueError: Date range is outside minimum or maximum of raw NPMRDS input data."
    - Note that this radio button should not be selected if Processes 2-4 will be run with this Process 1 output file because the processed NPMRDS file used as input to processes 2-4 dictates the date range of outputs for those processes. The Process 2 and 3 outputs containing less than a full year of data are not currently compatible with MOVES and the TNMAide spreadsheet requires a full year of data input from the DANA Process 4 output.
- Single year selection of pre-processed TMAS Station data[14] provided with the DANA tool[15] (TMAS_Station_XXXX.csv, where XXXX represents the year)
- Single year selection of pre-processed TMAS Classification data provided with the DANA tool (TMAS_Class_Clean_XXXX.csv, where XXXX represents the year)
- NPMRDS speed data[16] (passenger vehicles, trucks, and all) – each of these three separate files are obtained from the RITIS website[17] download package and an example download

---

[14] Note that TMAS data for all years have been updated with DANA version 2.0 to reflect Process 0 improvements. If you have previously conducted analyses using TMAS data from DANA version 1.0, the results may be different, even when using the same version of the DANA Tool itself.

[15] Or processed user-defined TMAS data, which should be automatically detected by the main GUI after running Process 0. However, if the file is located outside of the default folder, it will not automatically be detected. In this case, the user can navigate to select the file from the alternate location. This remains true for both the processed TMAS Station and Class data outputs from Process 0.

[16] Note that only TMAS and NPMRDS data for 2021 are available, so 2021 DANA analyses are limited to process 1 and 4 until 2021 HPMS, County Road Mileage, and Statewide Functional Class VMT are available.

[17] https://npmrds.ritis.org/analytics/.

package can be found in the following directory: User Input Files\Middlesex_MA\NPMRDS Data\

- NPMRDS TMC configuration – also obtained from the RITIS website download package[18] (TMC_Identification.csv) – an example file as part of the RITIS download package can be found in the following directory: User Input Files\Middlesex_MA\NPMRDS Data\
- NEI emission rates – pre-processed 2017 emission rates based on EPA's National Emissions Inventory (NEI) are provided as a default file with the DANA tool (NEI2017_RepresentativeEmissionsRates.parquet)[19]
- Federal Information Processing Standard (FIPS) state and county codes – a default file provided with the DANA tool, current as of 2019 (FIPS_County_Codes.csv)
- National Emissions Inventory representative county codes – a default file provided with the DANA tool that identifies the NEI representative county associated with each county in the U.S, based on the 2017 NEI regions (NEI2017_RepresentativeCounties.csv)

1.6.3.1 Downloading NPMRDS Data

NPMRDS data may be obtained by state and local agencies and their contractors from the Regional Integrated Transportation Information System (RITIS) website. After creating a free user account,[20] click either of the highlighted areas in Figure 19 to access the Massive Data Downloader.

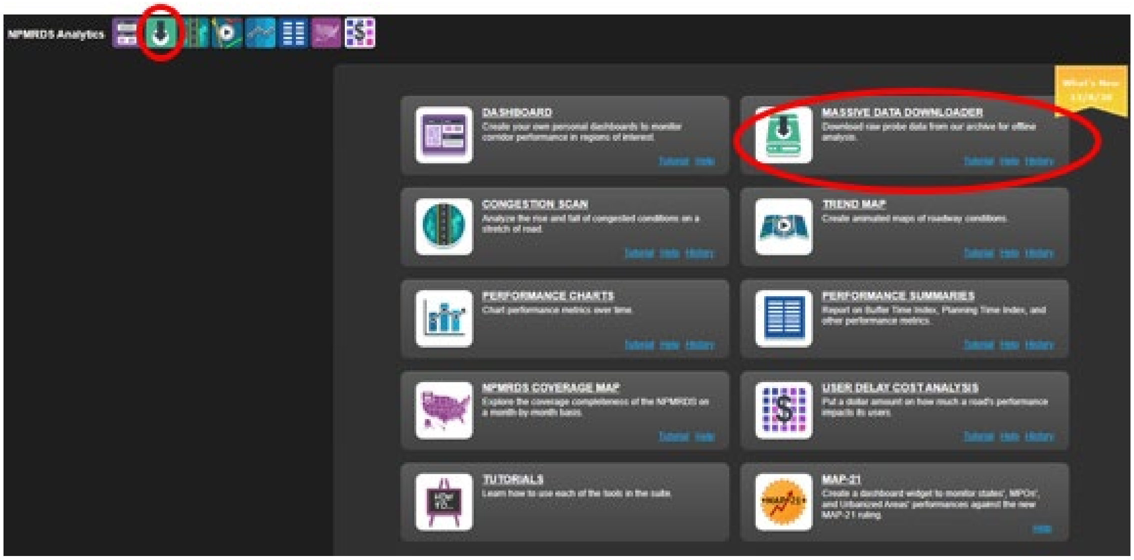


Figure 19. Screenshot. NPMRDS RITIS home screen with Massive Data Downloadeder links highlighted

[18] Note that the DANA tool will attempt to handle non-standard data in the TMC_Identification.csv file; however, if problems occur during processing, this is a good first place to look.

[19] See Section 1.5.3.3 for information on user-supplied emission rates. See Appendix D for more information on default emission rates included.

[20] Please follow the steps from the quick start guide at https://npmrds.ritis.org/static/help/docs/NPMRDSquickstart.pdf or email npmrds@ritis.org for assistance in getting access to the NPMRDS data.

Select the appropriate NPMRDS INRIX year[21] using the "TMC Segments from" dropdown menu. Also select the county of interest using the "Region" tab (exemplified for Middlesex County, Massachusetts in Figure 20), then click the green "Add region" button.

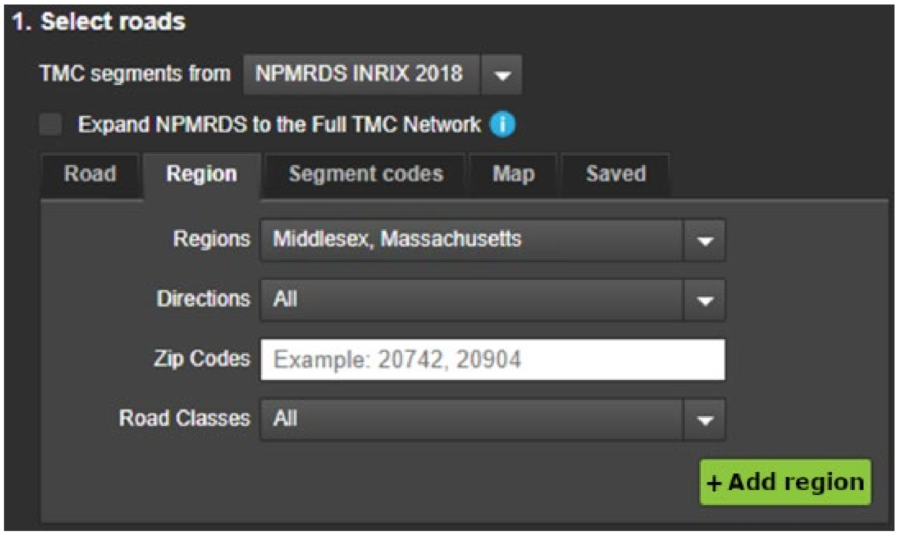


Figure 20. Screenshot. Massive Data Downloader region selection example

Select a date range using the start and end date calendar selectors, as shown in Figure 21 for the full 2018 calendar year. If Processes 2-4 are to be run, users should select a full year of data, as the Process 2 and 3 outputs containing less than a full year of data are not currently compatible with MOVES and TNMAide requires a full year of traffic inputs in order to determine the worst hour. Note that NPMRDS data prior to 2017 are available, but the TMC configuration file in the download package for these data contains no HPMS conflated data. Thus, NPMRDS data prior to 2017 are not digestible by DANA and should not be used.

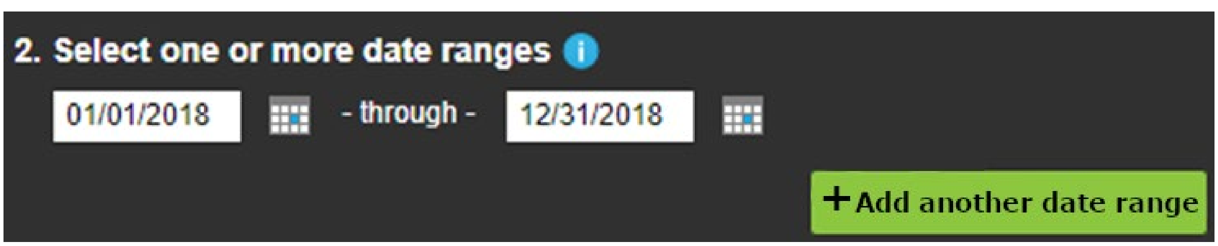


Figure 21. Screenshot. Massive Data Downloader date range selection example

Check to ensure the defaults displayed in Figure 22 are maintained to generate data for every day of the week from midnight to 11:59 PM.

[21] Note that the DANA Tool contains the NPMRDS 2021 shapefile as a back-end input.

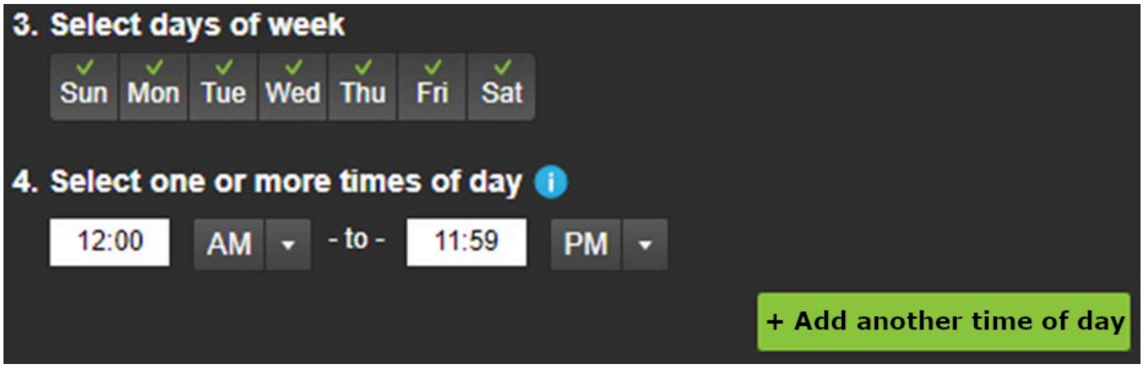


Figure 22. Screenshot. Massive Data Downloader day and time defaults

Select all three vehicle type groups ("passenger vehicles", "trucks", and "trucks and passenger vehicles") as shown in Figure 23.

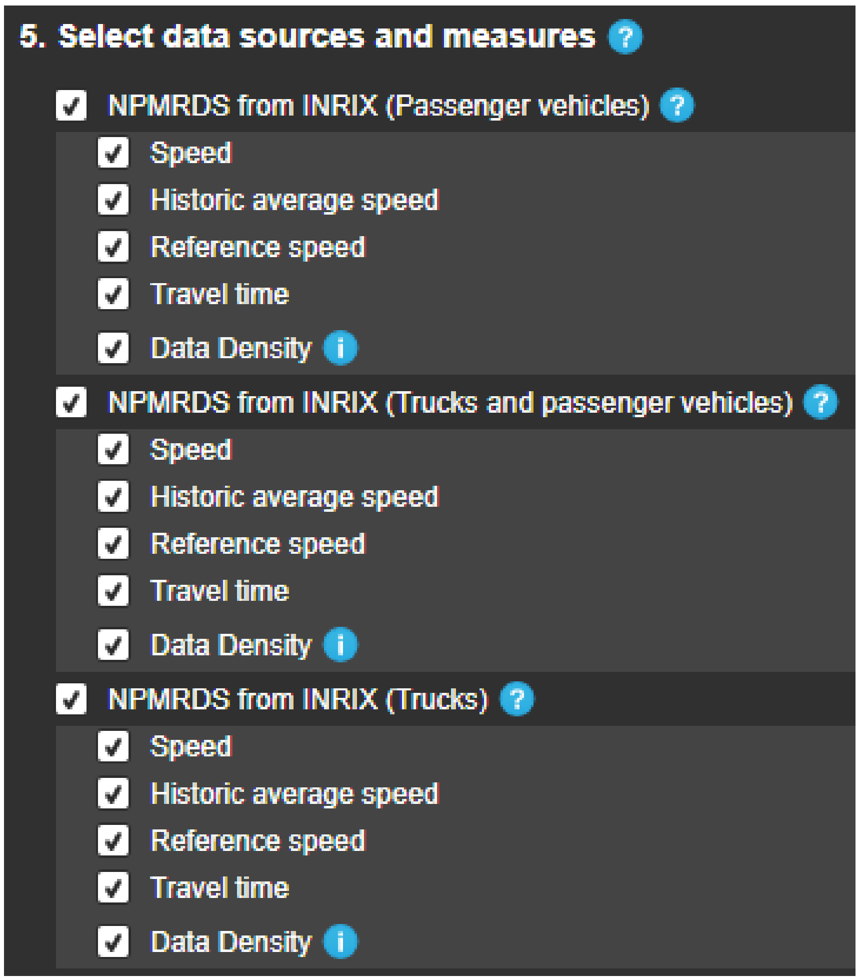


Figure 23. Screenshot. Massive Data Downloader complete vehicle type data source selection

Check to ensure the defaults displayed in Figure 24 are maintained to generate data in units of seconds and to exclude records with null values.

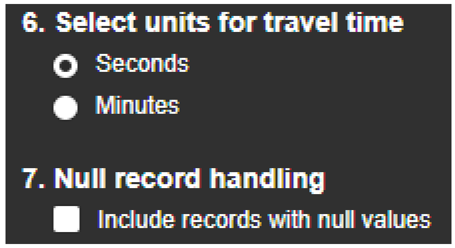


Figure 24. Screenshot. Massive Data Downloader time units and null record handling defaults

Select 1 hour averaging,[22] enter a descriptive title for your download package, and keep the notification box checked before clicking "SUBMIT", as shown in Figure 25.

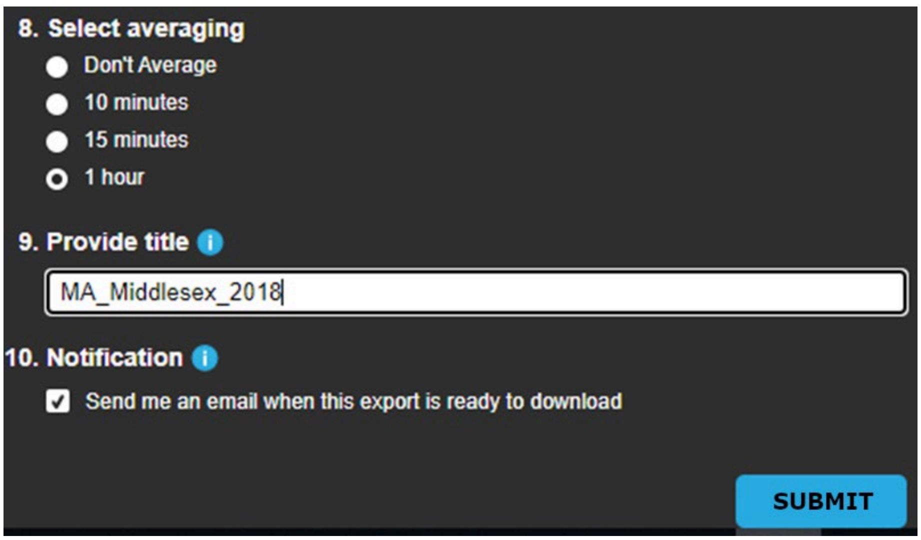


Figure 25. Screenshot. Massive Data Downloader #8-10 inputs

Within a few minutes, you will receive an email from npmrds-analytics-downloader@ritis.org with a link to retrieve your selected data files. Using these downloaded NPMRDS data as inputs, complete execution of the Process 1 script will result in the following output files in the Final_Output\ Process1_LinkLevelDataset directory:

- Hierarchy of results from the TMAS and NPMRDS data merge: [23]
  - Exact roadway match (tier1_class.csv)
  - State, county, and route match (tier2_class.csv)
  - State, urban/rural, and roadway functional class match (tier3_class.csv)
  - National, urban/rural, and roadway functional class match (tier4_class.csv)
- Composite dataset with TMAS and emission rates information wherein the percent vehicle type fractions sum to one across 24 hours of the day and the five vehicle types (PCT_TYPE10-60 and PCT_NOISE_AUTO, etc.). (XX_Composite_Emissions.parquet, where XX represents the state abbreviation)

---

[22] NPMRDS temporal resolution of one hour provides sufficient granularity to produce valid output data, while reducing file size and processing time.
[23] See Figure 17 for more details.

- The first 1000 and last 1000 rows of the parquet link-level dataset (XX_Composite_Emissions_SAMPLE.csv, where XX represents the state abbreviation)
- Aggregation of average speed, AADT, and estimated emissions per roadway mile data for each link in the NPMRDS dataset over the specified time period, as well as the geography of each relevant roadway link in text format for GIS visualization (XX_Composite_Emissions_SUMMARY.csv, where XX represents the state abbreviation)

1.6.3.2 Viewing Parquet Files

Parquet files are not human readable, but there are several tools available to export the compressed data into a usable format. A free, open source Parquet Viewer executable[24] can be downloaded,[25] although it may have trouble viewing larger files. Alternatively, the pyarrow library[26] will support Python scripts to access the data within the parquet file. The parquet.R[27] library serves the same purpose for R scripts.

1.6.3.3 User-Supplied Emissions Rate Data

The user may choose to provide their own emissions rate data, which can be substituted for the default NEI2017_RepresentativeEmissionsRates.parquet input file, provided the substitute file matches the expected data structure of the default. This would also require that the county is either pre-existing in or added to the NEI2017_RepresentativeCounties.csv file. See the Process 1 subsection of Appendix C. Input Data Dictionaries by Process for complete formatting details of both input files.

1.6.4 Process 2: Produce MOVES Inputs

Several sources of HPMS-based data are combined with TMAS data to produce VMT-based MOVES county-level inputs. Note that MOVES includes two roadway categories each with two classes: rural/urban and restricted/unrestricted (i.e., highways that can only be accessed by an on-ramp vs. all other roadways (arterials, connectors, and local streets)). VMT distributions output by the DANA Tool for one or more of these four road types may be zero if the road type does not exist in the input. For example, Washington, D.C. has no area classified as “rural”, so VMT distributions for both restricted and unrestricted rural road types in the DANA Tool output are zero. If VMT for any of these road types are zero in the DANA Tool output, the MOVES RunSpec must be updated to reflect the limited expected road types.

The following inputs are required:

- Cleaned Composite dataset with emission rates from Process 1 (XX_Composite_Emissions.parquet, where XX represents the state abbreviation).
  - Note that this file should have been created using the Process 1 default auto-detected date range based on a full year of NPMRDS Process 1 input data. This

[24] Executable located at https://github.com/mukunku/ParquetViewer/
[25] User documentation located at https://github.com/mukunku/ParquetViewer/wiki
[26] https://arrow.apache.org/docs/python/parquet.html
[27] https://github.com/apache/arrow/blob/master/r/R/parquet.R

input file dictates the date range of the output file and the Process 2 outputs containing less than a full year of data are not currently compatible with MOVES.

- HPMS – the HPMS public release of geospatial data available from FHWA[28] and an example file can be found in the following directory: User Input Files\Middlesex_MA\HPMS Data\
- Single year selection of state annual VMT by roadway classification – 2015-2020 files provided with the DANA tool (State_VMT_by_Class_XXXX.csv, where XXXX represents the year)[29]
- Single year selection of national county mileage summary files – 2015-2020 files included with the DANA tool (County_Road_Mileage_XXXX.csv, where XXXX represents the year)

Using HPMS and formatted VM-2 data as inputs, complete execution of the Process 2 script will result in the following output files in the Final_Output\Process2_MOVES_VMT_Distributions directory:[30]

- State-level monthly VMT fractions (XX_MONTH_VMT.csv, where XX represents the state abbreviation)
- State-level daily VMT fractions (XX_DAY_VMT.csv, where XX represents the state abbreviation)
- State-level hourly VMT fractions (XX_HOUR_VMT.csv, where XX represents the state abbreviation)
- State-level VMT fractions by region (XX_REGIONAL_VMT.csv, where XX represents the state abbreviation)
- State-level VMT fractions by MOVES road type (XX_ROADTYPE_VMT.csv, where XX represents the state abbreviation)

Complete execution of the Process 2 script will also result in one output file for every county containing the NPMRDS TMC links. If only interested in a single county, users can safely ignore outputs from other counties. All county-level files will be located in the Final_Output\Process2_MOVES_VMT_Distributions directory and the appropriate sub-directory noted below:

- County-level monthly VMT fractions (XX_MONTH_VMT_YY.0.csv in the\XX_MONTH_VMT sub-directory, where XX represents the state abbreviation and YY.0 represents the county code)
- County-level daily VMT fractions (XX_DAY_VMT_YY.0.csv in the \XX_DAY_VMT sub-directory, where XX represents the state abbreviation and YY.0 represents the county code)

---

[28]HPMS data for 2017 and prior: https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles_2017.cfm
HPMS data for 2018 and beyond: https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles.cfm.

[29] Non-formated VM-2 tables for other years available for download at https://www.fhwa.dot.gov/policyinformation/statistics.cfm. See Section 1.5.4.4 for more information.

[30] State-level files cannot be used as input to MOVES; rather, county-level files in the subfolders can be used as input to MOVES.

- County-level hourly VMT fractions (XX_HOUR_VMT_YY.0.csv in the \XX_HOUR_VMT sub-directory, where XX represents the state abbreviation and YY.0 represents the county code)
- County-level VMT fractions by region (XX_REGIONAL_VMT_YY.0.csv in the \XX_REGIONAL_VMT sub-directory, where XX represents the state abbreviation and YY.0 represents the county code)
- County-level VMT fractions by MOVES road type (XX_ROADTYPE_VMT_YY.0.csv in the \XX_ROADTYPE_VMT sub-directory, where XX represents the state abbreviation and YY.0 represents the county code)

1.6.4.1 Downloading 2017 and Prior HPMS Data

HPMS data for years 2017 and prior can be obtained from the FHWA Office of Highway Policy Information website[31], by selecting the year of interest and clicking "Download", as shown in Figure 26. Extract the .shp file within the downloaded .zip file. To load this file into the DANA Tool, change the file type to .shp in the file explorer, as shown in Figure 27.

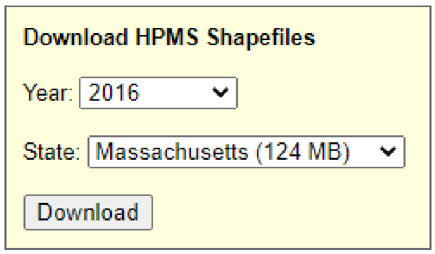


Figure 26. Screenshot. 2011-2017 HPMS download example from FHWA

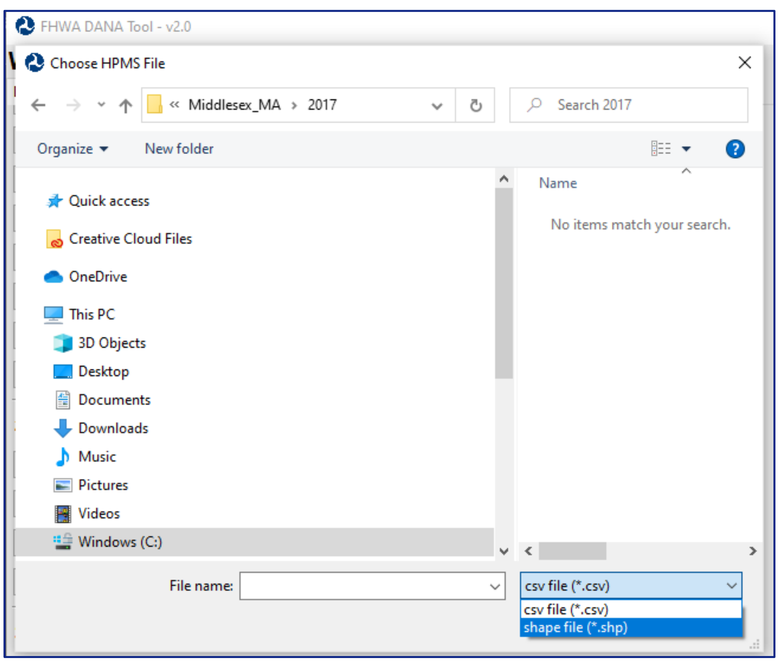


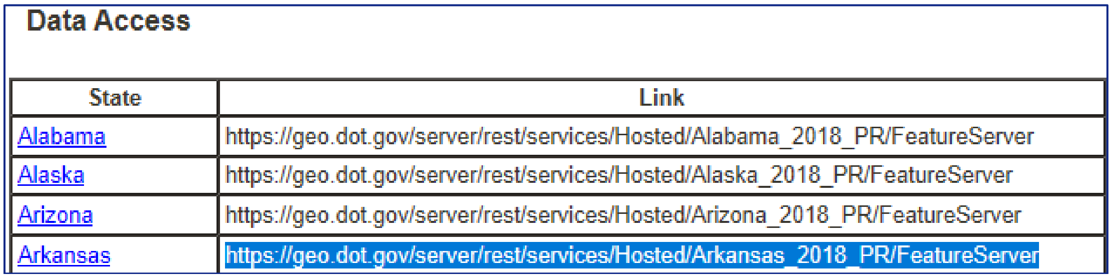
Figure 27. Screenshot. File explorer selection of .shp HPMS input file

[31] https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles_2017.cfm

HPMS data for years 2018 and beyond must be downloaded using GIS software.

1.6.4.2 Downloading 2018 and Beyond HPMS Data via ArcGIS

Select the geodatabase server on the FHWA Office of Highway Policy Information website[32] by scrolling down to the "Data Access" section and copying the link for the state of interest as shown in Figure 28.



Data Access

State	Link
Alabama	https://geo.dot.gov/server/rest/services/Hosted/Alabama_2018_PR/FeatureServer
Alaska	https://geo.dot.gov/server/rest/services/Hosted/Alaska_2018_PR/FeatureServer
Arizona	https://geo.dot.gov/server/rest/services/Hosted/Arizona_2018_PR/FeatureServer
Arkansas	https://geo.dot.gov/server/rest/services/Hosted/Arkansas_2018_PR/FeatureServer

Figure 28. Screenshot. AR 2018 HPMS geodatabase server link from FHWA website

In ArcGIS, click "Catalog", "GIS Servers", "Add ArcGIS Server", and "Use GIS Services". Paste the copied server link from FHWA into the "Server URL" field, as illustrated in Figure 29. These data are public, so Authentication is not necessary. Click "OK".

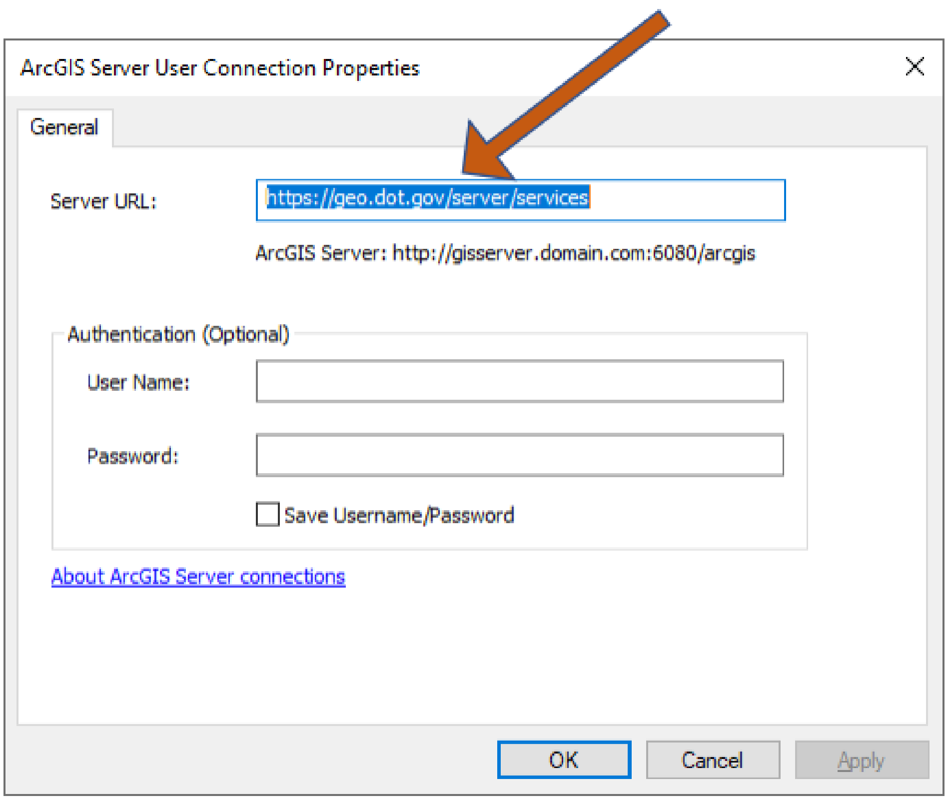


Figure 29. Screenshot. ArcGIS Server User Connection Properties with pasted link

32 https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles.cfm

Click “server on geo.dot.gov (user)” to expand the tree. Click “Hosted” to see the list of files, an excerpt from which is exemplified in Figure 30. Drag and drop the shapefile into ArcMap, as shown in Figure 31. Use the Data Export tool to obtain the shapefile as a .csv file. See Appendix C. Input Data Dictionaries by Process for proper column headers.

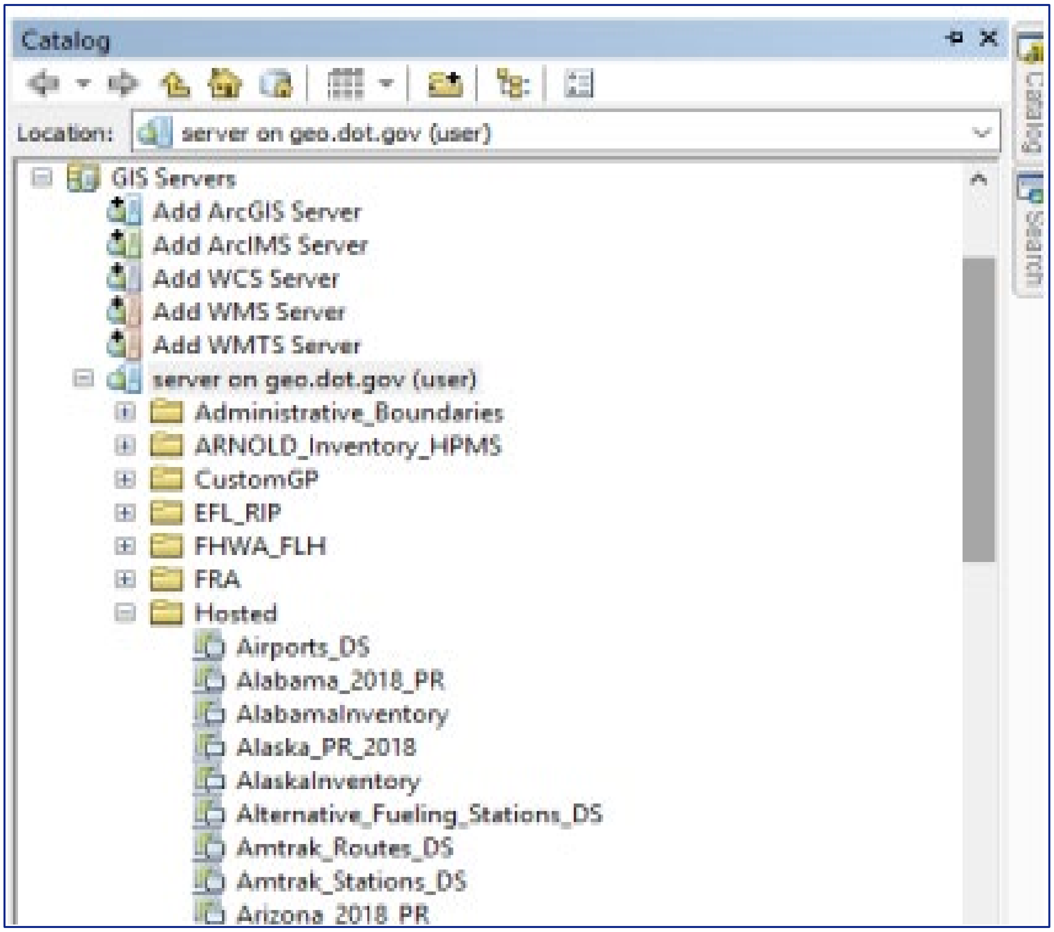


Figure 30. Screenshot. Expanded ArcGIS Server tree

Figure 31. Screenshot. ArcGIS HPMS shapefile display example

1.6.4.3 Downloading 2018 and Beyond HPMS Data via QGIS

In QGIS, right click "ArcGISFeatureServer" from the Browser (shown in Figure 32), and click "New Connection…".

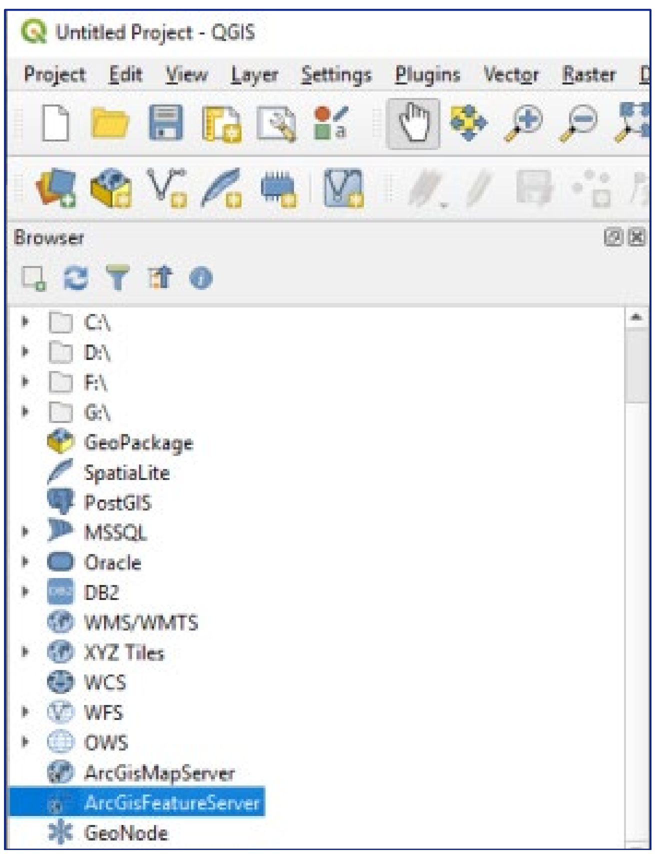


Figure 32. Screenshot. QGIS Browser with ArcGISFeatureServer highlighted

Enter a descriptive server name in the “Name” field, as demonstrated in Figure 33. Enter the following DOT hosted services link in the URL field: https://geo.dot.gov/server/rest/services/Hosted. These data are public, so Authentication is not necessary. Click “OK”.

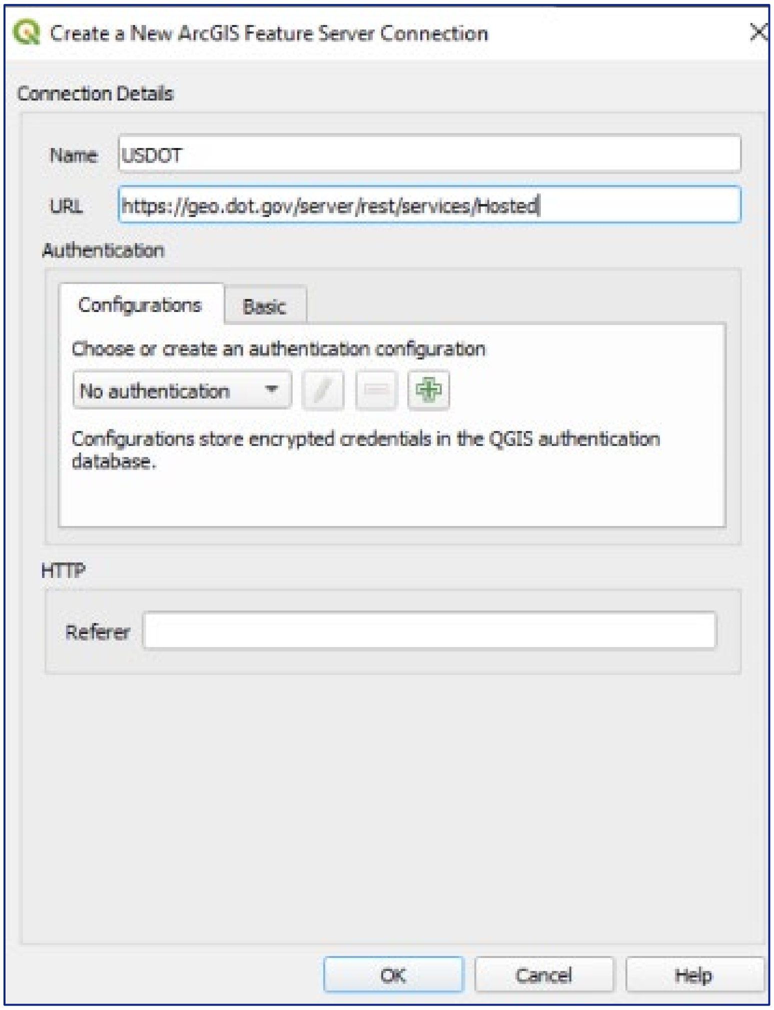


Figure 33. Screenshot. QGIS Server Connection window with completed Name and URL fields

Click on your chosen name for this connection. Figure 34 displays “USDOT” as per the example in Figure 33.

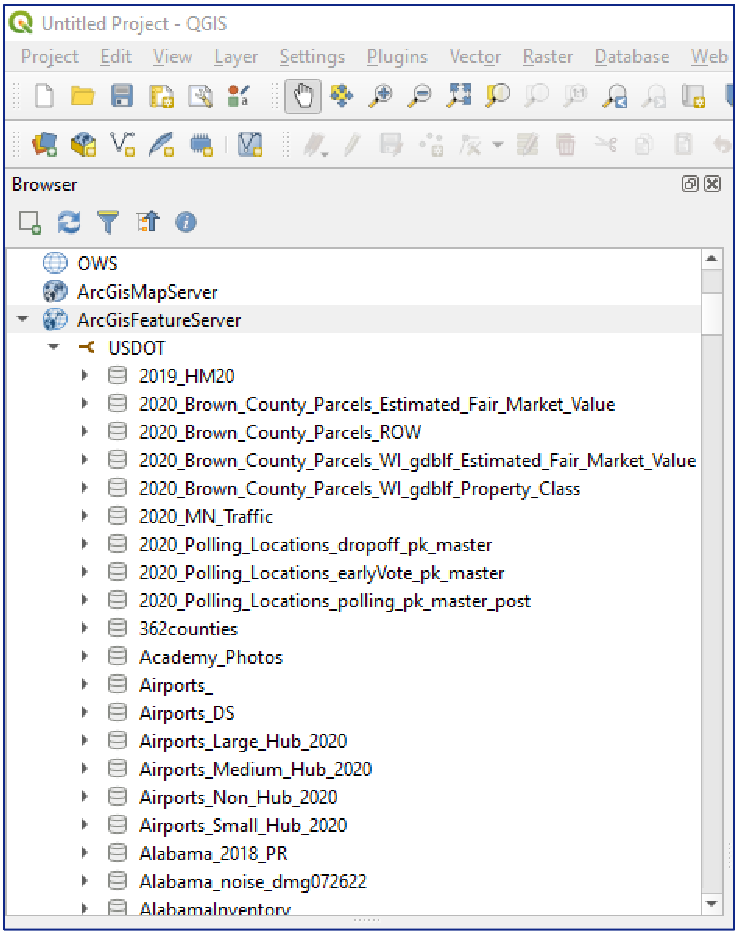


Figure 34. Screenshot. QGIS Browser with newly added server expanded

Files entitled “StateName_2018_PR” (NewJersey_2018_PR, for example) contain the publicly released full extent HPMS where it was required from the state DOTs. Double click on the file of choice and again on the layer beneath. The FeatureClass will be displayed in the Viewer, as demonstrated in Figure 35. This may take several minutes, depending performance factors.

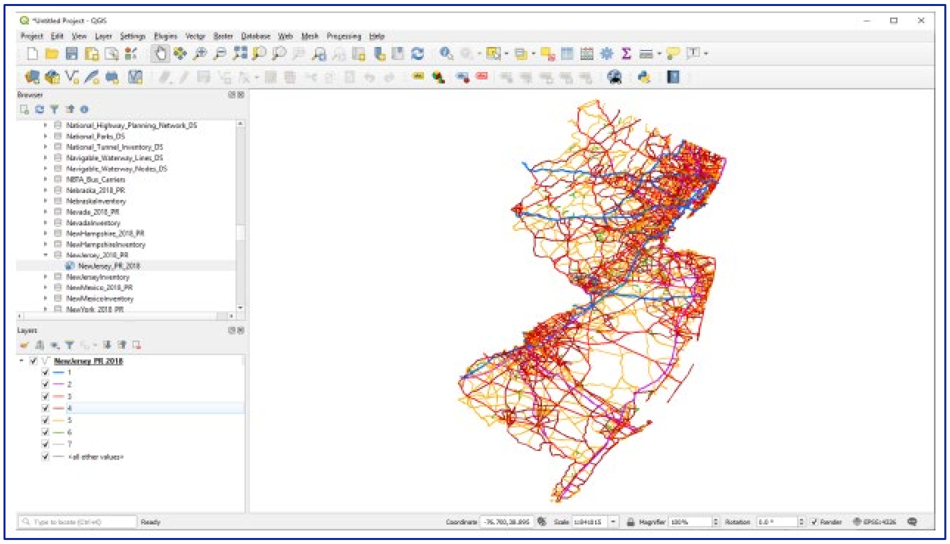


Figure 35. Screenshot. QGIS HPMS shapefile display example

Export this layer as a .csv file, circled in the Save Vector Layer as window in Figure 36. See Appendix C. Input Data Dictionaries by Process for proper column headers.

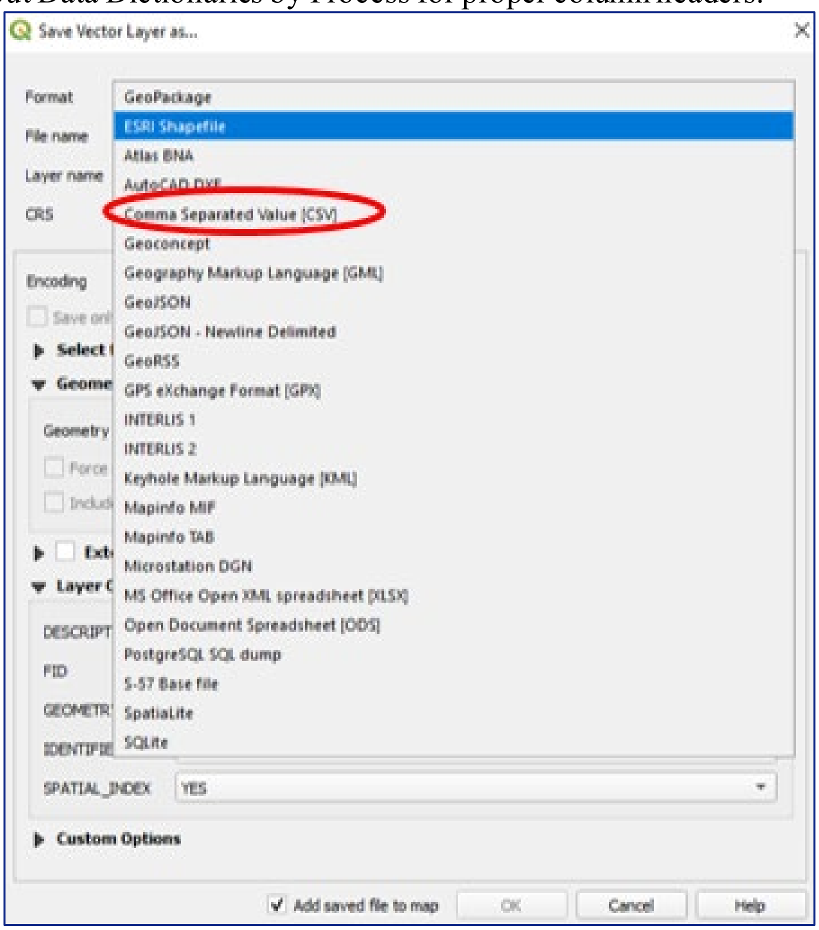


Figure 36. Screenshot. QGIS export options

1.6.4.4 Downloading and Formatting VM-2 Data

VM-2 data for years 2015-2020 are provided with the DANA tool. If other years are desired, such data can be obtained from the FHWA Office of Highway Policy Information website at the following link: https://www.fhwa.dot.gov/policyinformation/statistics.cfm. Use the dropdown menu to choose the year of interest, then click "Go", as shown in Figure 37.

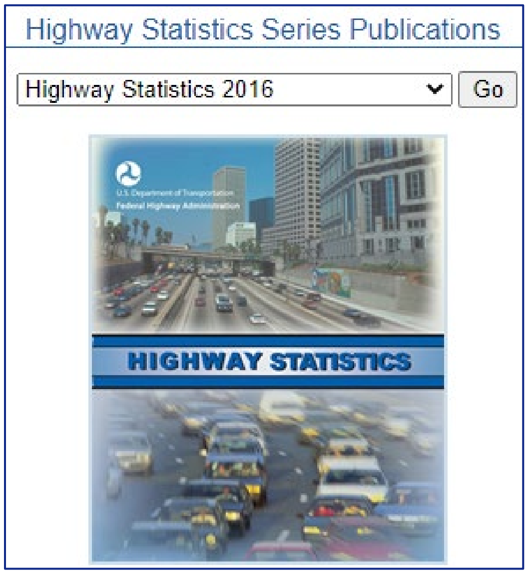


Figure 37. Screenshot. FHWA Highways Statistics year selection

Scroll down to the "5. Highway Travel" section and click the Excel link shown in Figure 38 to download the VM-2 data as an .xls file.

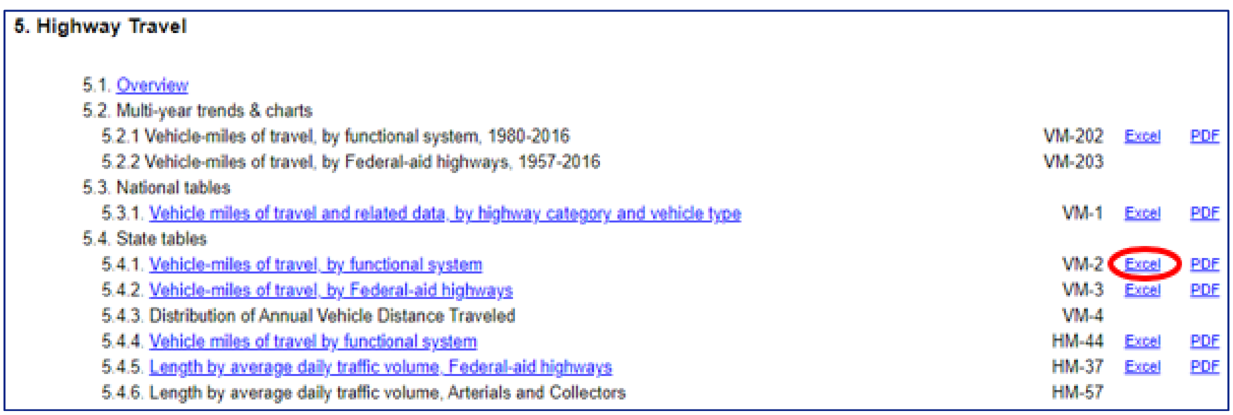


Figure 38. Screenshot. 2016 FHWA Highway Travel Statistics page with the VM-2 Excel download option highlighted

Open the downloaded file in a spreadsheet editor. Excerpts of a raw file from the FHWA website are annotated in Figure 39. Remove the header and footer lines (highlighted in red) and replace all "-" (highlighted in blue) with "0". Ensure that all zero replacements are of a data type indicating numeric values, rather than character strings. Save the formatted file as a .csv for input to the DANA tool. An excerpt of the resulting file is displayed in Figure 40.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
7	FUNCTIONAL SYSTEM TRAVEL - 2016 (1)																	
8	ANNUAL VEHICLE - MILES																	
10	MAY 2020									(MILLIONS)								TABLE VM-2
11		RURAL								URBAN								
12–14	STATE	INTERSTATE	OTHER FREEWAYS AND EXPRESSWAYS	OTHER PRINCIPAL ARTERIAL	MINOR ARTERIAL	MAJOR COLLECTOR	MINOR COLLECTOR	LOCAL	TOTAL	INTERSTATE	OTHER FREEWAYS AND EXPRESSWAYS	OTHER PRINCIPAL ARTERIAL	MINOR ARTERIAL	MAJOR COLLECTOR	MINOR COLLECTOR	LOCAL	TOTAL	TOTAL
15	Alabama	6,333	-	5,489	4,527	4,394	1,357	7,036	29,135	8,988	551	9,853	7,271	4,255	64	9,109	40,092	69,227
16	Alaska	896	-	326	128	307	148	434	2,238	768	-	971	547	244	115	375	3,020	5,259
17	Arizona	6,550	30	3,348	1,446	2,485	459	1,390	15,708	7,572	7,821	8,226	15,582	4,199	147	6,531	50,078	65,786
18	Arkansas	3,995	275	3,710	2,859	3,528	666	2,193	17,227	5,400	965	3,746	4,551	1,841	71	1,954	18,528	35,755
19	California	15,316	4,889	10,163	7,736	8,008	2,524	4,360	52,995	74,066	62,237	57,101	49,120	23,679	313	20,605	287,120	340,115
20	Colorado	4,683	248	4,202	2,057	1,796	764	1,528	15,277	9,375	5,406	9,297	6,342	2,722	46	3,687	36,874	52,152
21	Connecticut	470	296	443	409	813	148	583	3,161	9,886	4,179	3,842	5,169	2,617	240	2,545	28,478	31,639
22	Delaware	-	601	861	317	603	174	444	3,000	1,432	646	2,168	1,098	813	56	965	7,178	10,178
63	West Virginia	2,379	0	2,177	1,445	2,311	348	951	9,610	3,570	93	2,281	1,903	910	27	681	9,466	19,077
64	Wisconsin	6,207	1,252	6,700	4,991	7,756	2,265	4,216	33,386	8,345	3,473	9,320	5,715	3,132	-	2,977	32,962	66,348
65	Wyoming	2,738	-	1,699	565	933	593	662	7,190	561	11	791	472	520	111	551	3,018	10,208
66	U.S. Total	261,644	38,720	197,630	147,457	164,153	42,615	131,633	983,853	575,753	254,877	483,557	418,151	217,681	22,355	305,544	2,277,919	3,261,772
67	Puerto Rico	490	-	210	245	136	0	0	1,082	4,716	912	3,041	3,007	1,935	-	16	13,628	14,710
68	Grand Total	262,134	38,720	197,841	147,703	164,289	42,615	131,633	984,935	580,469	255,789	486,599	421,158	219,616	22,355	305,560	2,291,547	3,276,482
69	(1) Travel for the rural minor collector and rural/urban local functional systems is estimated by the States based on a model or other means and provided																	
70	to the FHWA on a summary basis. Travel for all other systems are estimated from State-provided data in the Highway Performance Monitoring System.																	

Figure 39. Screenshot. Raw VM-2 spreadsheet excerpts

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Alabama	6556	0	5615	4514	4377	1386	6341	28788	9278	550	9747	7275	4235	65	11228	42379	71167
2	Alaska	827	0	309	117	318	178	527	2277	753	0	931	525	247	116	638	3211	5487
3	Arizona	6751	82	3629	1495	2473	508	1405	16343	7888	8539	8070	14980	3659	142	6523	49802	66145
4	Arkansas	4265	327	3727	2878	3623	694	2273	17787	5521	985	3739	4576	1898	76	2094	18888	36675
5	California	16224	5043	10538	8133	9690	1769	7035	58432	75786	62937	54211	48803	24646	808	23173	290364	348796
6	Colorado	4846	274	4415	2099	1876	795	1590	15894	9812	5665	9495	6417	2817	47	3805	38059	53954
7	Connecticut	487	304	435	409	820	147	558	3159	9902	4210	3796	5162	2609	244	2513	28437	31596
8	Delaware	0	321	720	261	574	170	424	2470	1490	723	2353	1167	871	68	1038	7710	10179
9	District of Columbia	0	0	0	0	0	0	0	0	502	376	1055	703	280	0	774	3691	3691
10	Florida	10981	2244	8574	3924	3858	1541	5620	36742	30592	15535	45557	29795	20355	###	39488	185074	221816
11	Georgia	7734	0	6884	6531	6743	1240	3717	32849	24292	3465	17682	20451	8072	552	24093	98607	131456

Figure 40. Screenshot. Formatted VM-2 spreadsheet excerpt

1.6.5 Process 3: Produce Speed Distributions

The composite dataset output from Process 1 is further processed to produce the MOVES county-level speed distribution input. Since NPMRDS data only include roadways on the national highway system (NHS), users should carefully examine resulting speed distributions for unrestricted roadway types (MOVES road types 3 and 5) before deciding whether to use them for MOVES inputs. See Appendix F. Comparison of Speed Distributions Derived from National Performance management Research Data Set (NPMRDS) Versus Streetlight Data for more information.

The following input is required:

- Cleaned Composite dataset with emission rates from Process 1 (XX_Composite_Emissions.parquet, where XX represents the state abbreviation).
    - Note that this file should have been created using the Process 1 default auto-detected date range based on a full year of NPMRDS Process 1 input data. This input file dictates the date range of the output file and the Process 3 outputs containing less than a full year of data are not currently compatible with MOVES.

Complete execution of the Process 3 script will result in the following output file in the Final_Output\Process3_MOVES_Speed_Distributions directory:

- Speed distribution by county, source type, road type, hour day and average speed bin (XX_SPEED_DISTRIBUTION.csv, where XX represents the state abbreviation). This state-level file cannot be used as input to MOVES.

Complete execution of the Process 3 script will also result in the following output file in the Final_Output\Process3_MOVES_Speed_Distributions\ XX_SPEED_DISTRIBUTION directory (where XX represents the state abbreviation):

- Speed distribution for a single county by source type, road type, hour day and average speed bin (XX_SPEED_DISTRIBUTION_YY.csv, where XX represents the state abbreviation and YY represents the county code). This county-level file can be used as input to MOVES.

1.6.6 Process 4: Produce Noise Inputs

The appropriate TMCs are selected out of the composite dataset from Process 1 for use as inputs to the TNMAide spreadsheet tool. Users should provide two[33] TMCs of interest separated by a comma and a space. The tool will conduct basic checks for invalid characters and will output empty data files if the specified format is not followed. TMCs should represent opposite directions along the same roadway section. If the user needs to identify TMCs for input into Process 4, click the "TMC Selection Tool" button (Figure 41) to open the TMC selection GUI tab (Figure 42). This tab lets the user choose a TMC configuration file (from the NPMRDS RITIS download package) and dynamically extracts and displays the lists of county, road and direction in the dropdown boxes for selection. The user can also select TMCs by providing a Google Earth polygon file (.kml). Once all inputs are specified, click the "Select Data" button at the bottom of the tab to create the output file. A window will appear displaying a message indicating completion of the output file creation, exemplified in Figure 43. The output of this tool is a text file TMCs_X_Y_Z.txt, where X represents the county, Y represents the road name, and Z represents the direction selected by the user. This file is created in the TMC_Selection subfolder of the Final Output folder. Data from this file can be copied and pasted into the text box in the Process 4 section of the main GUI with a space added, as shown in Figure 44.

The RITIS Massive Data Downloader[34] used to obtain the NPMRDS data can also be used to visualize the output from the TMC selection tool in order to choose the opposing directions of travel for a single roadway section to analyze in TNMAide. To output TMCs from both

[33] TNMAide requires input data from exactly two TMCs in order to compute the loudest hour. If more than two TMCs are input to Process 4, the user must iterate TNMAide runs.

[34] https://npmrds.ritis.org/analytics/download/

directions of travel in the TMC Selection Tool, leave the third dropdown blank before clicking “Select Data”. After the TMC Selection Tool output has been created (Figure 43), open text file and copy its contents into the Segment codes tab of the RITIS Massive Data Downloader, as shown in Figure 45. Click “Add segments” to view the TMC links on the map. Hovering over each segment in the map will display a tooltip containing the segment code (Figure 46) that can be input into the Process 4 TMC Codes textbox in the DANA Data Processing tab (Figure 44).

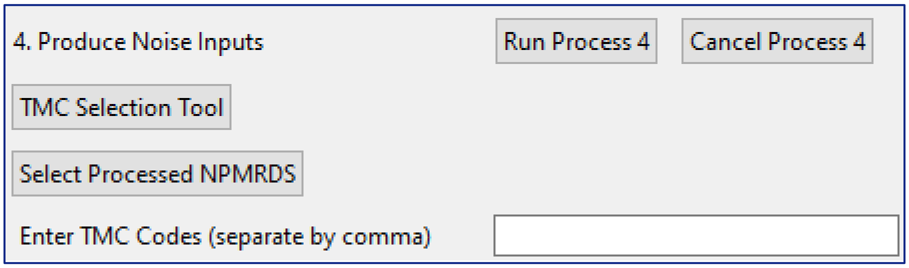


Figure 41. Screenshot. GUI Process 4 section including TMC Selection Tool button

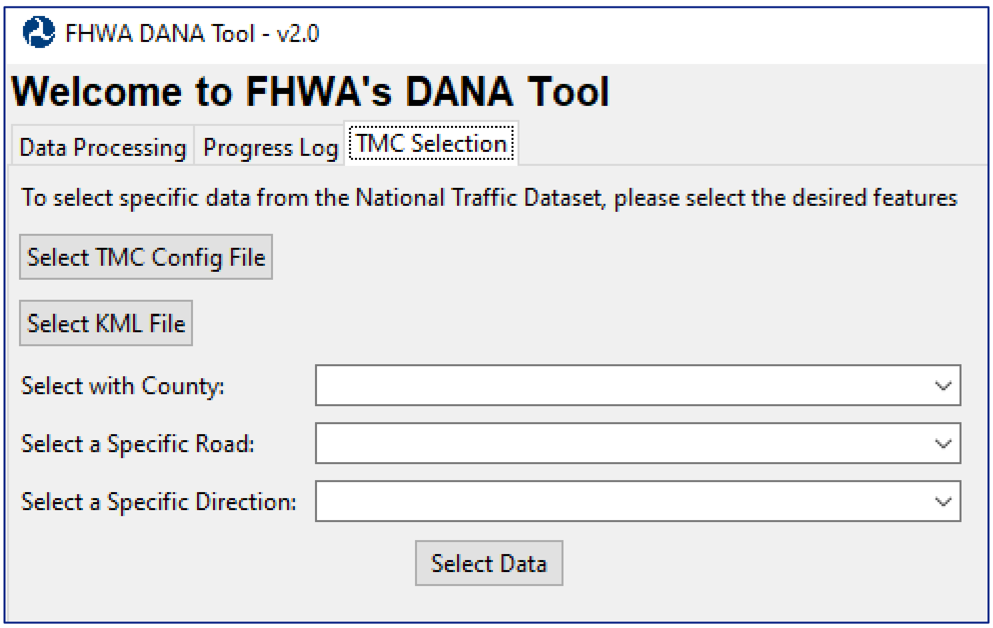


Figure 42. Screenshot. GUI TMC Selection Tab

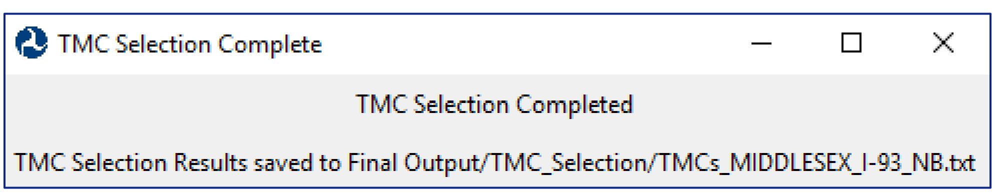


Figure 43. Screenshot. TMC Selection Completion Message

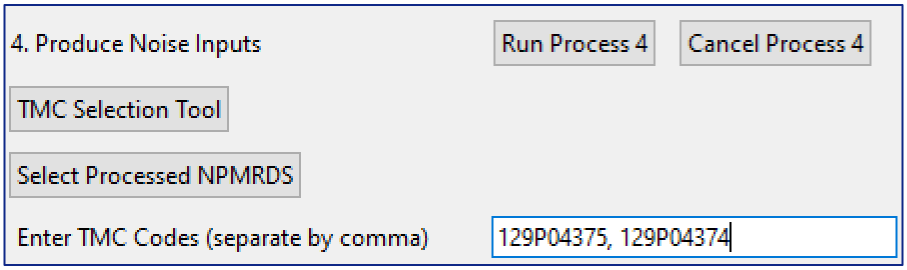


Figure 44. Screenshot. GUI Process 4 section with copied and formatted TMC codes from TMC Selection tab output

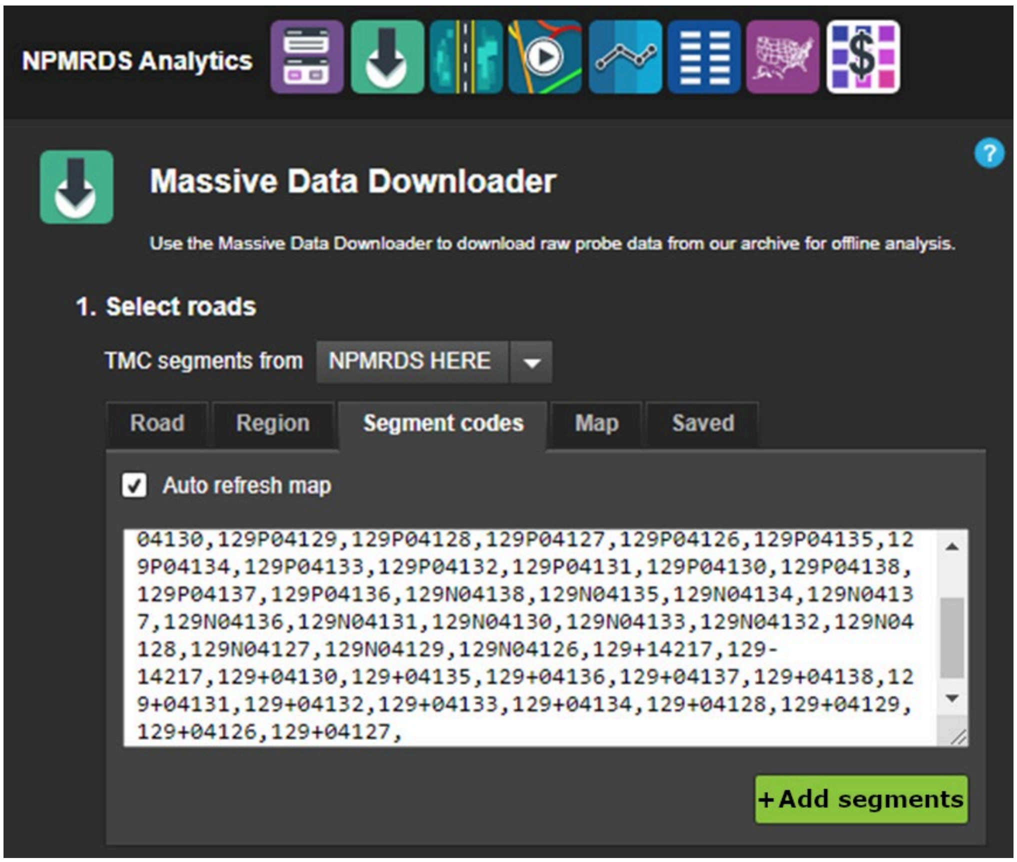


Figure 45. Screenshot. RITIS Massive Data Downloader Segment codes tab containing copied TMC Selection Tool output from all directions of travel

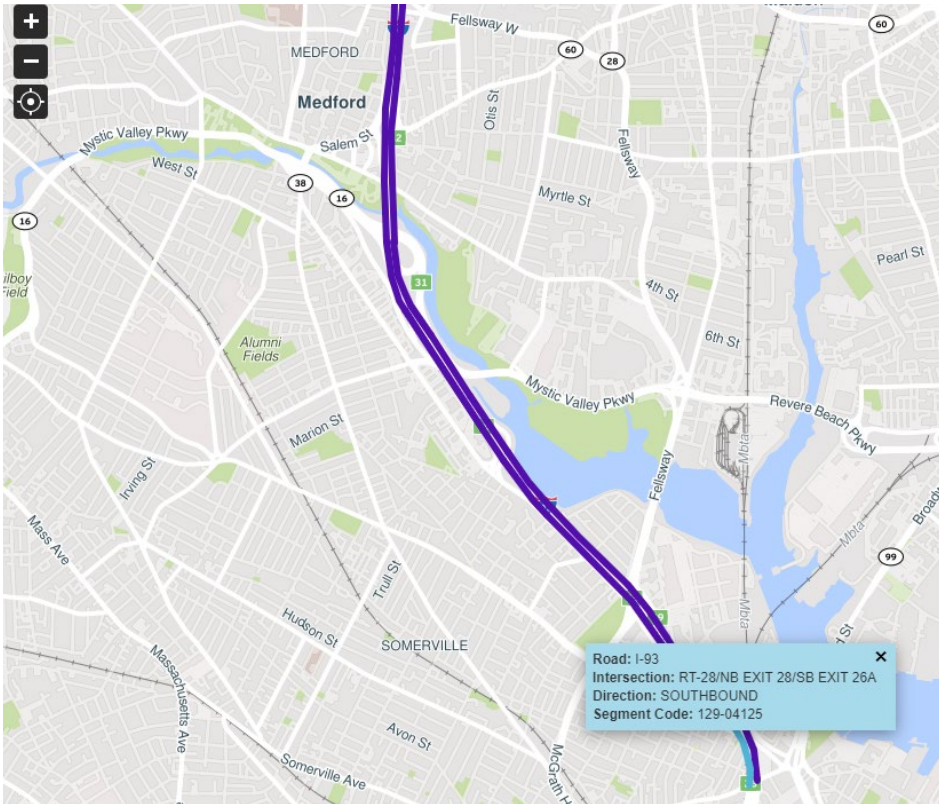


Figure 46. Screenshot. RITIS Massive Data Downloader map viewing copied TMC Selection Tool output from all directions of travel with one segment highlighted

The following inputs are required:

- Cleaned Composite dataset with emission rates from Process 1 [35](XX_Composite_Emissions.parquet, where XX represents the state abbreviation).
    - Note that this file should have been created using the Process 1 default auto-detected date range based on a full year of NPMRDS Process 1 input data. This input file dictates the date range of the output file and the TNMAide spreadsheet requires a full year of data input from the DANA Process 4 output.
- Input TMC list (either manual input or copy from the TMC Selection tool output)

Complete execution of the Process 4 script will result in the following output file in the Final_Output\Process4_TNM_AIDE_Inputs directory:

- Composite dataset results filtered for selected TMCs (XX_Composite_Emissions_select.csv, where XX represents the state abbreviation)

[35] Note that only TMAS and NPMRDS data for 2021 are available, so 2021 DANA analyses are limited to process 1 and 4 until 2021 HPMS, County Road Mileage, and Statewide Functional Class VMT are available.

2 APPENDIX A. PROCESS 0: PROCESS RAW TMAS DATA (OPTIONAL)

Pre-processed TMAS data for the entire country from 2015-2021 are included in the DANA tool for input to Process 1, in which case Process 0 is not required. If processing user-defined geographically-specific or newly updated TMAS data is desired, the following inputs are required:

- TMAS Station Data in *Traffic Monitoring Guide*[36] format
- TMAS Classification Data in *Traffic Monitoring Guide* format
- Federal Information Processing Standard (FIPS) state and county codes – a default file provided with the DANA tool, current as of 2019 (FIPS_County_Codes.csv)
- National Emissions Inventory representative county codes – a default file provided with the DANA tool that identifies the NEI representative county associated with each county in the U.S, based on the 2017 NEI regions (NEI2017_RepresentativeCounties.csv)

Complete execution of the Process 0 script will result in the following output files in the TMAS_Intermediate_Output folder:

- Processed TMAS Station dataset (TMAS_station_State.csv)
- Processed TMAS Classification dataset (tmas_class_clean.csv)

The data flow when executing Process 0 is as follows:

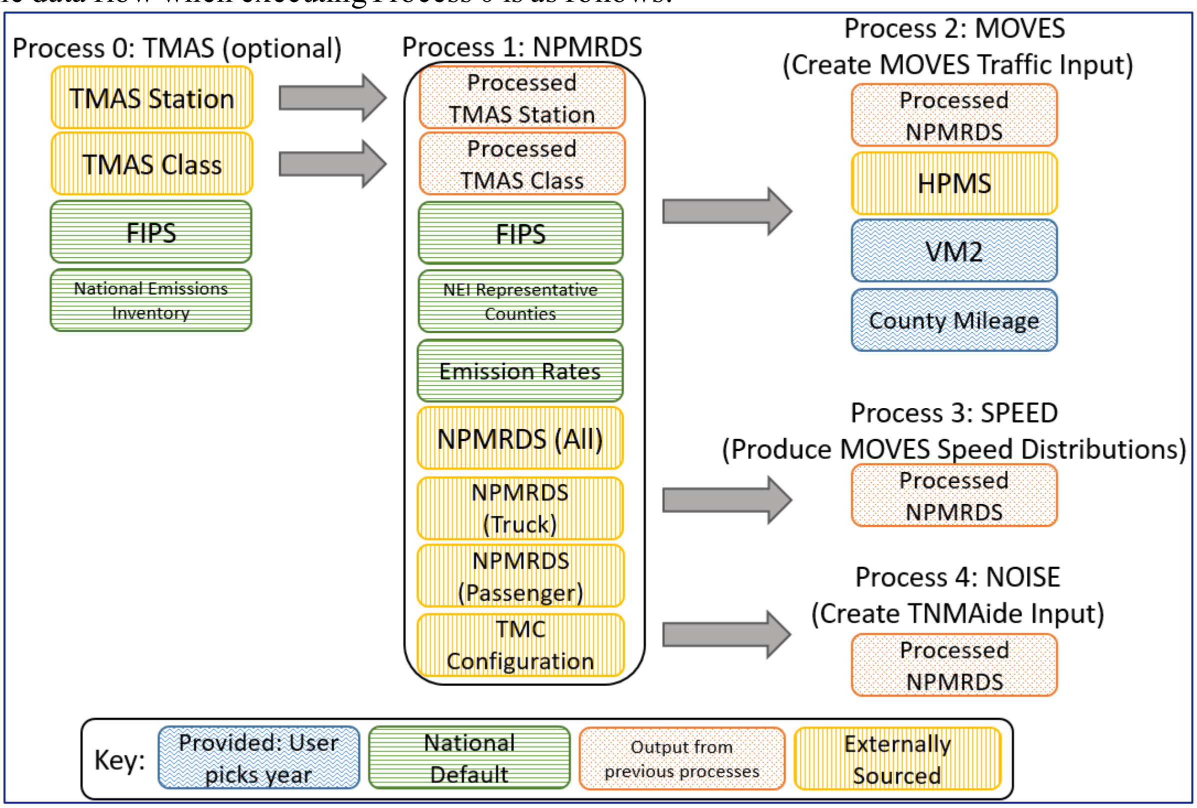


Figure 47. Flowchart. DANA tool input data flow including optional Process 0

[36] https://www.fhwa.dot.gov/policyinformation/tmguide/.

3 APPENDIX B. DANA TOOL INPUT DETAILS

Input Name	Notes	Input Type[37]	File Location	Source	Website for Download
Process 0: Process Raw TMAS Data (optional)					
TMAS Station	If user-defined geographically-specific or newly updated TMAS data are used rather than the pre-processed national TMAS dataset, a TMAS Station File would be required	Externally Sourced	N/A	FHWA Office of Highway Policy Information/Travel Monitoring and Survey Team (HPPI-30)	Additional Information: https://www.fhwa.dot.gov/policyinformation/tmguide/
TMAS Class	If user-defined geographically-specific or newly updated TMAS data are used rather than the pre-processed national TMAS dataset, a TMAS Class File would be required	Externally Sourced	N/A	FHWA Office of Highway Policy Information/Travel Monitoring and Survey Team (HPPI-30)	Additional Information: https://www.fhwa.dot.gov/policyinformation/tmguide/
FIPS	Use the default file provided. This file provides FIPS county codes for every county in the U.S.	National Default	Default Input Files\FIPS_County_Codes.csv	US Census Bureau/American National Standards Institute (ANSI)	https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt
NEI Representative Counties	Use default file provided. This file from NEI provides mapping of all U.S. counties to representative counties with NEI emission rates available.	National Default	Default Input Files\NEI2017_RepresentativeCounties.csv	EPA. 2017 National Emissions Inventory (NEI) Data.	https://gaftp.epa.gov/air/nei/2017/doc/supporting_data/onroad/

[37] The images correspond to those shown in Figure 15.

Process 1: Process Raw NPMRDS Data					
Processed TMAS Station	Use the default file provided for the year of your choice. TMAS counting station files are provided for 2015-2021.[38] Additional years may be provided in the future.	Provided: User picks year	Default Input Files\TMAS Data\TMAS XXXX\TMAS_Station_XXXX.csv, where XXXX represents the year[39]	N/A	N/A
Processed TMAS Class	Use the default file provided for the year of your choice. TMAS classification count files are provided for 2015-2021.[40] Additional years may be provided in the future.	Provided: User picks year	Default Input Files\TMAS Data\TMAS XXXX\TMAS_Class_Clean_XXXX.csv, where XXXX represents the year[41]	N/A	N/A
NPMRDS (All)	Use the NPMRDS speed data file for trucks and passenger vehicles (all) from RITIS download package[42]	Externally Sourced	User Input Files\Middlesex_MA\NPMRDS Data\MA_MIDDLESEX_2018_ALL.csv	FHWA Office of Transportation Management (HOTM-1)	https://npmrds.ritis.org/analytics/download/
NPMRDS (Passenger)	Use the NPMRDS speed data file for passenger vehicles from RITIS download package[42]	Externally Sourced	User Input Files\Middlesex_MA\NPMRDS Data\ MA_MIDDLESEX_2018_PASSENGER.csv	FHWA Office of Transportation Management (HOTM-1)	https://npmrds.ritis.org/analytics/download/
NPMRDS (Truck)	Use the NPMRDS speed data file for trucks from RITIS website download package[42]	Externally Sourced	User Input Files\Middlesex_MA\ NPMRDS Data\ MA_MIDDLESEX_2018_TRUCKS.csv	FHWA Office of Transportation Management (HOTM-1)	https://npmrds.ritis.org/analytics/download/

[38] Note that only TMAS and NPMRDS data for 2021 are available, so 2021 DANA analyses are limited to process 1 and 4 until 2021 HPMS, County Road Mileage, and Statewide Functional Class VMT are available.
[39] Alternatively, if optional Process 0 is executed, output from Process 0 will replace this input file, and will be located in the TMAS_Intermediate_Output folder.
[40] Note that only TMAS and NPMRDS data for 2021 are available, so 2021 DANA analyses are limited to process 1 and 4 until 2021 HPMS, County Road Mileage, and Statewide Functional Class VMT are available.
[41] Alternatively, if optional Process 0 is executed, output from Process 0 will replace this input file, and will be located in the TMAS_Intermediate_Output folder.
[42] See Section 1.5.3.1 for instructions on how to create a download package for a particular geographic area and time period.

Process 1: Process Raw NPMRDS Data					
TMC Configuration	Use the "TMC_Identification.csv" file from RITIS download package[42]	Externally Sourced	User Input Files\Middlesex_MA\NPMRDS Data\TMC_Identification.csv	FHWA Office of Transportation Management (HOTM-1)	https://npmrds.ritis.org/analytics/download/
Emission Rates	Use the default file provided from NEI 2017	National Default	Default Input Files\NEI2017_RepresentativeEmissionsRates.parquet	EPA. 2017 National Emissions Inventory (NEI) Data.	See Appendix E. Development of Default Emissions Rates from 2017 NEI Data
FIPS	Use the default file provided. This file provides FIPS county codes for every county in the U.S.	National Default	Default Input Files\FIPS_County_Codes.csv	US Census Bureau/American National Standards Institute (ANSI)	https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt
NEI Representative Counties	Use default file provided. This file from NEI provides mapping of all U.S. counties to representative counties with NEI emission rates available.	National Default	Default Input Files\NEI2017_RepresentativeCounties.csv	EPA. 2017 National Emissions Inventory (NEI) Data.	https://gaftp.epa.gov/air/nei/2017/doc/supporting_data/onroad/

Process 2: Produce MOVES Inputs					
Processed NPMRDS	Process 1 output becomes input for Process 2	Output from previous processes	Process1_LinkLevelDataset\XX_Composite_Emissions.parquet, where XX represents the state abbreviation	N/A	N/A
HPMS	Use the HPMS public release of geospatial data in shapefile format available from FHWA[43]	Externally Sourced	User Input Files\Middlesex_MA\HPMS Data\MA_HPMS_2018.csv	FHWA Office of Highway Policy Information/Highway System Performance Team (HPPI-20)	Up to and including 2017: https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles_2017.cfm 2018 and later use GIS Server
VM2	Use the default file provided for the year of your choice. VMT data for years 2015-2020 are provided. Use Table VM-2 from *Highway Statistics* if other years are desired[44]	Provided: User picks year	Default Input Files\Statewide Functional Class VMT\State_VMT_by_Class_XXXX.csv, where XXXX represents the year	FHWA Office of Highway Policy Information	https://www.fhwa.dot.gov/policyinformation/statistics.cfm
County Mileage	Use default file provided for the year of your choice. HPMS county mileage summary files are provided for 2015-2020. Additional years may be provided in the future.	Provided: User picks year	Default Input Files\HPMS County Road Mileage\County_Road_Mileage_XXXX.csv, where XXXX represents the year	FHWA Office of Highway Policy Information/Highway System Performance Team (HPPI-20)	N/A

Process 3: Produce Speed Distributions					
Processed NPMRDS	Process 1 output becomes input for Process 3	Output from previous processes	Process1_LinkLevelDataset\XX_Composite_Emissions.parquet, where XX represents the state abbreviation	N/A	N/A

Process 4: Produce Noise Inputs					
Processed NPMRDS	Process 1 output becomes input for Process 4	Output from previous processes	Process1_LinkLevelDataset\XX_Composite_Emissions.parquet, where XX represents the state abbreviation	N/A	N/A
TMC Codes (separate by comma and a space)	Use TMC Selection tab to gather TMC codes for appropriate roadway links	N/A	N/A	N/A	N/A

[43] See Section 1.5.4 for instructions on downloading HPMS data

[44] See Section 1.5.4.4 for instructions on preparing the file for input

4 APPENDIX C. INPUT DATA DICTIONARIES BY PROCESS[45]

4.1 PROCESS 1

Provided: User picks year

TMAS_Class_Clean_XXXX.csv, where XXXX represents the year						
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
STATE	FIPS State Code	int	4	N	N/A	N/A
STATION_ID	Station Identification	int	10091	Y	Y	Matching with relevant station information from TMAS_Station_XXXX.csv
DIR	Direction of Travel Code	int	1	Y	Y	Matches with corresponding NPMRDS roadway Links
DATE	Date of Data	string	7/27/2017	N	N/A	N/A
YEAR	Year of Data	int	17	N	N/A	N/A
MONTH	Month of Data	int	7	Y	Y	Vehicle type distributions are grouped by month in the TMAS tiers
DAY	Day of Data	int	27	N	N/A	N/A
HOUR	Hour of Data	int	6	Y	Y	Vehicle type distributions are grouped by hour of the day in the TMAS tiers
DAY_TYPE	Weekday/Weekend	string	WD	Y	Y	Vehicle type distributions are grouped by weekends or weekdays in the TMAS tiers
PEAKING	Peak in Morning/Afternoon	string	AM	Y	Y	In Tiers 2, 3, 4, the vehicle type distributions are grouped by this indicator which marks whether the link has its peak hours in the morning of the afternoon, which in turn is an indicator of whether the

[45] The images in the top right corner of each table correspond to the categories shown in Figure 15.

						link direction runs inbound or outbound
VOL	Peak Hour Volume	int	986	Y	N	N/A
F_SYSTEM	Functional System Code	int	3	Y	Y	Vehicle type distributions are grouped by the highway functional classification in TMAS tiers 3 and 4
URB_RURAL	Urban/Rural	string	U	Y	Y	Vehicle type distributions are grouped by the urban and rural classification in TMAS tiers 3 and 4
COUNTY	FIPS County Code	int	13	Y	Y	Vehicle type distributions are grouped by the county in which the roadway resides in TMAS tier 2
REPCTY	Representative County for Emissions Lookup from National Emissions Inventory	int	13	Y	N	N/A
ROUTE_SIGN	Route Sign Name	int	0	Y	Y	Vehicle type distributions are grouped by the route sign level in the national highway system in TMAS tier 2
ROUTE_NUMBER	Route Number Identification	string	0000I290	Y	Y	Vehicle type distributions are grouped by the route number in the national highway system in the TMAS tier 2
LAT	Latitude of Count Station	float	33.360254	Y	N	N/A
LONG	Longitude of Count Station	float	-111.84701	Y	N	N/A
STATE_NAME	State Abbreviation	string	AZ	Y	Y	Vehicle type distributions are specific to the NPMRDS state in TMAS tiers 1, 2, and 3
COUNTY_NAME	County Name	string	Maricopa County	Y	N	N/A
LOCATION	Description of Location	string	SR 87	N	N/A	N/A

HPMS_TYPE10	Count of Vehicle Type 10 (motorcycles) vehicles	int	2	Y	Y	Indicates the vehicle type percentages for HPMS vehicle type 10
HPMS_TYPE25	Count of Vehicle Type 25 (passenger cars) vehicles	int	901	Y	Y	Indicates the vehicle type percentages for HPMS vehicle type 25
HPMS_TYPE40	Count of Vehicle Type 40 (busses) vehicles	int	13	Y	Y	Indicates the vehicle type percentages for HPMS vehicle type 40
HPMS_TYPE50	Count of Vehicle Type 50 (medium trucks) vehicles	int	55	Y	Y	Indicates the vehicle type percentages for HPMS vehicle type 50
HPMS_TYPE60	Count of Vehicle Type 60 (heavy trucks) vehicles	int	15	Y	Y	Indicates the vehicle type percentages for HPMS vehicle type 60
HPMS_ALL	Count of All Vehicles for HPMS	int	986	N	N/A	N/A
NOISE_AUTO	Count of NOISE Type Automobiles	int	901	Y	Y	Indicates the vehicle type percentages for noise vehicle type automobiles
NOISE_MED_TRUCK	Count of NOISE Type Medium Trucks	int	41	Y	Y	Indicates the vehicle type percentages for noise vehicle type medium trucks
NOISE_HVY_TRUCK	Count of NOISE Type Heavy Trucks	int	29	Y	Y	Indicates the vehicle type percentages for noise vehicle type heavy trucks
NOISE_BUS	Count of NOISE type Buses	int	13	Y	Y	Indicates the vehicle type percentages for noise vehicle type Buses
NOISE_MC	Count of NOISE type miscellaneous	int	2	Y	Y	Indicates the vehicle type percentages for noise vehicle type miscellaneous
NOISE_ALL	Count of All Vehicles for NOISE	int	986	N	N/A	N/A

Provided: User picks year

TMAS_Station_XXXX.csv, where XXXX represents the year						
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
STATE	FIPS State Code for Count Station	int	25	N	N/A	N/A
COUNTY	FIPS County Code for the Count Station	int	17	Y	N	N/A

STATION_ID	Station Identification for the Count Station	int	803	Y	Y	Matching with relevant station infomation from TMAS_CLASS_XXXX.csv
DIR	Direction of Travel Code	int	3	Y	N	N/A
URB_RURAL	Urban/Rural	string	U	Y	N	N/A
F_SYSTEM	Functional System Code	int	1	Y	N	N/A
ROUTE_SIGN	Route Sign Name	int	1	Y	N	N/A
ROUTE_NUMBER	Route Number Identification	string	0000I290	Y	N	N/A
LAT	Latitude of Count Station	float	42.361064	Y	Y	Geographic matching for TMAS tier 1
LONG	Longitude of Count Station	float	-71.597149	Y	Y	Geographic matching for TMAS tier 1
LOCATION	Description of Count Station Location	string	INTERSTATE 290	N	N/A	N/A
STATE_NAME	State Abbreviation	string	MA	Y	Y	Vehicle type distributions are specific to the NPMRDS state in TMAS tiers 1, 2, and 3
COUNTY_NAME	County Name	string	Middlesex County	Y	N	N/A
REPCTY	Representative County for Emissions Lookup from National Emissions Inventory	int	17	Y	N	N/A

NPMRDS speed data[46]						Externally Sourced
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
tmc_code	TMC Link Code	string	133-04099	Y	Y	Uniquely identifies all TMC links included in the input
measurement_tstamp	Timestamp in 5 minute intervals of the measurement	string	1/1/2018 7:00	Y	Y	Match to vehicle type distributions in TMAS data by speed measurement timestamp
speed	The harmonic average speed for all reporting vehicles on the segment.	int	62	Y	Y	Included as final output in link-level dataset

[46] Data for passenger vehicles, trucks, and all vehicles have the same format

average_speed	The historical average speed. Historical average speeds are calculated by the CATT Lab by taking the harmonic average of speeds on each segment for each hour of day and for each day of the week. For data from February 1, 2017 onward, this historical average speed is calculated over the period of February 1, 2017 - June 30th, 2017. For data prior to February 1, 2017, the average is calculated using the twelve-month period preceding November 2014.	int	56	N	N/A	N/A
reference_speed	An approximation of free-flow speed for the segment. This value is calculated by the CATT Lab using the 95th percentile of the speeds between 10 PM and 5 AM. The reference speed is calculated over a 6-month period starting April 1st, 2017 - September 30th, 2017.	int	65	N	N/A	N/A
travel_time_seconds	Travel time recorded in minutes or seconds. It is the ratio between the segment length and the harmonic average speed for all reporting vehicles on the segment.	float	155.83	Y	Y	Travel time included as output in final link-level dataset
data_density	Refers to one of three values: A) Fewer than five values, B) Five to nine values, C) More than nine values	string	A	N	N/A	N/A

Externally Sourced

TMC_Identification.csv						
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
tmc	the unique 9 digit value identifying the TMC segment.	string	129-04130	Y	Y	Uniquely identifies each NPMRDS roadway link
road	the route number or common name of the roadway	string	I-93	N	N/A	N/A
direction	the overall direction of the roadway	string	SOUTHBOUND	Y	Y	Links with direction in TMAS tier data
intersection	the cross street and/or interchange associated with the TMC segment	string	MA-28/FELLSWAY/EXIT 33	N	N/A	N/A
state	the postal abbreviation of the state to which the TMC Segment is assigned	string	MA	Y	Y	Links with state in TMAS tier data
county	county name	string	MIDDLESEX	Y	N	N/A
zip	zip code	int	2155	N	N/A	N/A
start_latitude	the latitude of the beginning of the TMC segment	float	42.457923	Y	Y	Geographically matches to

						TMAS stations in tier 1
start_longitude	the longitude of the beginning of the TMC segment	float	-71.102238	Y	Y	Geographically matches to TMAS stations in tier 1
end_latitude	the latitude of the end of the TMC segment	float	42.4367488	Y	Y	Geographically matches to TMAS stations in tier 1
end_longitude	the longitude of the end of the TMC segment	float	-71.1031972	Y	Y	Geographically matches to TMAS stations in tier 1
miles	the length of the TMC segment	float	1.49144	Y	Y	Used to calculate speed from travel time and emissions per mile in the summary output dataset
road_order	a numerical value indicating in what order the TMC segment would be encountered when traveling downstream relative to the other TMC segments on the same road	int	133	N	N	NA
timezone_name	name of timezone	string	America/New_York	N	N/A	N/A
type	the type of tmc code. "P1" is the typical TMC Code. "P3" indicates national, state, and county boundaries, rest areas, toll plazas, major bridges, etc. "P4" is for ramps.	string	P1	N	N/A	N/A
country	the country in which the TMC segment is located	string	USA	N	N/A	N/A
tmclinear	a reference to the "Linear TMC" that includes the TMC Segment. Typically, several TMC Segments are part of a Linear TMC, which usually represents a road corridor through a single county. The purpose of this column is to provide assistance for filtering and locating TMC Segments and simplifying the process of linking consecutive TMC Segments.	int	65	N	N/A	N/A
frc	the class or group of roads to which the road belongs	string	1	N	N/A	N/A
border_set	Code to indicate if the TMC is within a 5-mile radius of Canadian or Mexican Boarder	string	N	N	N/A	N/A

f_system	The FHWA-approved Functional Classification System code. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the value for the highest functional class (minimum code value) is assigned.	int	1	Y	Y	Matches to HPMS functional classification in TMAS tier data
urban_code	The U.S. Census Urban Area Code. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	9271	N	N/A	N/A
faciltype	The operational characteristic of the roadway. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	2	N	N	N/A
structype	Code for roadway section that is a bridge, tunnel or causeway. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	1	N	N/A	N/A
thrulanes	The number of lanes designated for through-traffic in BOTH TRAVEL DIRECTIONS. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	8	N	N	N/A
route_numb	The signed route number. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	93	Y	Y	Matches to national highway route number in TMAS tier data
route_sign	Code for the type of route signing. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	2	Y	Y	Matches to national highway route sign level in TMAS tier data
route_qual	Code for the route signing descriptive qualifier. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	1	N	N/A	N/A
altrtename	A familiar, non-numeric designation for a route. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	string	044A	N	N/A	N/A
aadt	Annual Average Daily Traffic. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the length-weighted average is assigned.	int	184104	Y	Y	Included as output in final link-level dataset

aadt_singl	Annual Average Daily Traffic for single-unit trucks and buses. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the length-weighted average is assigned.	int	4126	Y	Y	Included as output in final link-level dataset
addt_combi	Annual Average Daily Traffic for Combination Trucks. If multiple HPMS segments with different attribute values are assigned to a single TMC path, the length-weighted average is assigned.	int	4441	Y	Y	Included as output in final link-level dataset
nhs	Code for a roadway that is a component of the National Highway System (NHS). If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant "on-NHS" value (i.e., 1 through 9) by length is assigned.	int	1	N	N	N/A
nhs_pct	The percentage of the TMC path length that is designated as NHS by HPMS (applicable when multiple HPMS segments assigned to a single TMC path).	int	100	N	N/A	N/A
strhnt_typ	Code for a roadway section that is a component of the Strategic Highway Network (STRAHNET). If multiple HPMS segments with different attribute values are assigned to a single TMC path, the predominant value by length is assigned.	int	0	N	N/A	N/A

NEI2017_RepresentativeEmissionsRates.parquet						National Default
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
repcty	County ID This repcty id can be matched to the counties in the TMAS data	int	1073	Y	Y	Matches emission rates to counties in TMAS and NPMRDS data
season	Encoded season which communicates month infomation	int	120102	Y	Y	Matches emission rates to month in TMAS and NPMRDS data
hourid	hour of the day which the data represent	int	1	Y	Y	Matches emission rates to hour in TMAS and NPMRDS data
roadtypeid	MOVES road type id	int	4	Y	Y	Matches emission rates to road type in TMAS and NPMRDS data
hpmsvtypeid	HPMS vehicle type code	int	10	Y	Y	Matches emission rates to HPMS vehicle type in TMAS data

pollutantid	MOVES pollutant ID	int	2	Y	Y	Included as output indicating specific pollutant emission rates
avgspeedbinid	speed bin identifier for speed bins with midpoint speed. Convention is: 1=2.5, 2=5, 3=10, 4=15, …, 16=75	int	1	Y	Y	Matches emission rates to the speed of passenger and heavy duty vehicles in NPMRDS data
grams_per_mile	Emissions rate in grams-per-mile	float	76.89116	Y	Y	Included as output for specific vehicle type, speed and pollutant emissions rates

FIPS_County_Codes.csv

National Default

Column Name	Data Description	Format Type	Example	Read by DANA?	Used in Analysis?	Used for?
STATE_NAME	State Abbreviation	int	AL	Y	Y	Matches to state column in NPMRDS TMC configuration file
STATE_CODE	State Number	int	1	Y	Y	Matches to FIPS state-county codes in TMC configuration file and TMAS data
COUNTY_CODE	County Code	int	55	Y	Y	Matches to FIPS state-county codes in TMC configuration file and TMAS data
COUNTY_NAME	County Name	string	Etowah County	Y	Y	Matches to county name in NPMRDS TMC Configuration file
FIPS_TYPE	County Category	string	H1	Y	N	N/A

NEI2017_RepresentativeCounties.csv

National Default

Column Name	Data Description	Format Type	Example	Read by DANA?	Used in Analysis?	Used for?
stateid	FIPS State Code	int	1	Y	Y	Matches to state FIPS ID in NPMRDS TMC Configuration file
countyid	State-County FIPS Code	int	1061	Y	Y	Matches to county FIPS ID in NPMRDS TMC Configuration file

State_Name	Name of State	string	Alabama	N	N/A	N/A
County_Name	Name of County	string	Geneva County	N	N/A	N/A
repcty	County ID This can be matched to the counties in the TMAS data	int	1097	Y	Y	Matches to representative county in the NEI emissions rates file

4.2 PROCESS 2

Externally Sourced

HPMS

Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
objectid	shapefile object identifier	int	1	N	N/A	N/A
year_record	Year for which the data apply	int	2018	N	N/A	N/A
state_code	State FIPS code	int	25	N	N/A	N/A
route_id	Location reference ID for the linear feature	string	SR1A NB	N	N/A	N/A
begin_point	beginning milepoint	float	27.089	Y	Y	Calculating length of the roadway segment in miles
end_point	ending milepoint	float	27.1	Y	Y	Calculating length of the roadway segment in miles
aadt	average annual daily traffic	int	18257	Y	Y	Used in multiple calculations including VMT by vehicle type distributions and monthly and day type distributions when producing MOVES inputs.
aadt_combination	combination truck AADT	int	181	N	N/A	N/A
aadt_single_unit	Single Unit Truck and Bus AADT	int	750	N	N/A	N/A
access_control_	The degree of access control for a given section of road	int	3	N	N/A	N/A
county_code	FIPS County Code	int	21	Y	Y	Results grouped by county codes and output separately for each unique code in input data
f_system	Functional System	int	4	Y	Y	Used to assign MOVES road types when producing MOVES inputs
facility_type	The operational characteristic of the roadway	int	2	N	N/A	N/A

iri	International Roughness Index	int	155	N	N/A	N/A
nhs	National Highway System	int	1	N	N/A	N/A
ownership	public/private entity ownership code	int	1	N	N/A	N/A
psr	Present Serviceability Rating for pavement condition	int	3	N	N/A	N/A
route_number	signed route number	int	1	N	N/A	N/A
route_qualifier	route signing descriptor	int	1	N	N/A	N/A
route_signing	type of route signing	int	4	N	N/A	N/A
speed_limit	posted speed limit	int	40	N	N/A	N/A
strahnet_type	Roadway section that is a component of the Strategic Highway Network (STRAHNET)	int	2	N	N/A	N/A
structure_type	Roadway section that is a bridge, tunnel or causeway	int	3	N	N/A	N/A
surface_type	Surface Type of a given section	int	7	N	N/A	N/A
through_lanes	number of lanes designated for through-traffic	int	4	N	N/A	N/A
toll_charged	Identifies sections that are toll facilities regardless of whether or not a toll is charged	int	2	N	N/A	N/A
toll_type	Indicates the presence of special tolls (i.e., High Occupancy Toll (HOT) lane(s) or other managed lanes)	int	1	N	N/A	N/A
truck	National Truck Network	int	1	N	N/A	N/A
urban_code	U.S. Census Urban Area Code	int	9271	Y	Y	Used to assign MOVES road types when producing MOVES inputs

Provided: User picks year

State_VMT_by_Class_XXXX.csv, where XXXX represents the year						
Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
State	State Name	string	California	Y	Y	Matches to state column in link level dataset
Rural Interstate	Rural Interstate VMT	int	17,184	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Other Freeways and Expressways	Rural Freeway & Expressway VMT	int	5,977	Y	Y	Used to calculate total VMT by road type to

						produce MOVES input files
Rural Other Principal Arterial	Rural Principal Arterial VMT	int	11,150	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Minor Arterial	Rural Minor Arterial VMT	int	8,332	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Major Collector	Rural Major Collector VMT	int	9,485	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Minor Collector	Rural Minor Collector VMT	int	1,023	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Local	Rural Local VMT	int	3,329	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Rural Total	Total Rural VMT	int	56,480	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Interstate	Urban Interstate VMT	int	74,947	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Other Freeways and Expressways	Urban Freeway & Expressway VMT	int	61,513	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Other Principal Arterial	Urban Principal Arterial VMT	int	53,258	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Minor Arterial	Urban Minor Arterial VMT	int	48,280	Y	Y	Used to calculate total VMT by road type to

						produce MOVES input files
Urban Major Collector	Urban Major Collector VMT	int	24,231	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Minor Collector	Urban Minor Collector VMT	int	746	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Local	Urban Local VMT	int	21,380	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
Urban Total	Total Urban VMT	int	284,356	Y	Y	Used to calculate total VMT by road type to produce MOVES input files
All VMT	Total Urban & Rural VMT	int	340,836	Y	N	N/A

Provided: User picks year

County_Road_Mileage_XXXX.csv, where XXXX represents the year

Column Name	**Data Description**	**Format Type**	**Example**	**Read by DANA?**	**Used in Analysis?**	**Used for?**
Year_Record	Year for which the data apply	int	2018	N	N/A	N/A
State_Code	FIPS State Code	int	1	Y	Y	Matches to state FIPS code in link level dataset
County_Code	FIPS County Code	int	1	Y	Y	Matches to county FIPS code in link level dataset
F_System	Functional Class Code	int	6	Y	Y	Matches to county FIPS code in link level dataset
Urban_Code	The U.S. Census Urban Area Code	int	99999	Y	Y	Used to set urban or rural road types to produce MOVES inputs
Ownership	Public/private entity ownership code	int	2	N	N/A	N/A
RMC_L_System_Length	Rural minor collector and local system length	float	65.410	Y	Y	To calculate VMT for road classes 6 and 7
Last_Modified_By	Most Recent Editor	string	John L Formby Jr	N	N/A	N/A

Last_Modified_On	Date of Most Recent Edit	date time	2019-04-10 11:35:02.987	N	N/A	N/A

5 APPENDIX D. OUTPUT DATA DICTIONARIES BY PROCESS

5.1 PROCESS 1

XX_Composite_Emissions.parquet,[47] where XX represents the state abbreviation

Column Name	Data Description	Format Type	Example
measurement_tstamp	Timestamp in 5 minute intervals of the measurement	datetime	12/1/2017 0:00
travel_time_all	Travel time for all vehicles during measurement time	float	34.87
speed_all	speed for all vehicles during measurement time	float	61.975338
travel_time_pass	travel time for passenger vehicles during measurement time	float	35.16
speed_pass	average speed for passenger vehicles during measurement time	float	61.464165
travel_time_truck	travel time for trucks during measurement time	float	34.6
speed_truck	average speed for trucks during measurement time	float	62.458958
year	year of measurement	int	2017
month	month year	int	12
day	day of month	int	1
hour	hour of day	int	0
weekday	Day of the week, 1-8, where 8 is for holidays specifically.	int	5
dow	weekday or weekend (WD or WE)	string	WD
peaking	AM or PM, depending on if that direction peaks during the morning commute	string	PM
tmc	TMC Link identifier	string	129N04633
road	Name of the roadway, usually the interstate system designation	string	I-395
direction	NB, SB, EB, WB standing for north, south, east and west bound TMC links	string	SB
start_latitude	The starting latitude of the TMC link	float	42.064864
start_longitude	The starting longitude of the TMC link	float	-71.859803
end_latitude	The ending latitude of the TMC link	float	42.056371
end_longitude	The ending longitude of the TMC link	float	-71.861056
tmc_length	the length of the TMC segment	float	0.6003

47 XX_Composite_Emissions_SAMPLE.csv has the same format, containing a subset of the rows from XX_Composite_Emissions.parquet.

road_order	Numerical value indicating in what order the TMC segment would be encountered when traveling downstream relative to the other TMC segments on the same road	int	9
f_system	the FHWA-approved Functional Classification System code If multiple HPMS segments with different attribute values are assigned to a single TMC path, the value for the highest functional class (minimum code value) is assigned	int	1
faciltype	the operational characteristic of the roadway	int	2
thrulanes	the number of lanes designated for through-traffic in both travel directions	int	4
aadt	Annual Average Daily Traffic on the TMC link	float	15237.5
aadt_singl	AADT of single unit trucks	int	844
aadt_combi	AADT of combination trucks	int	1506
nhs	code for a roadway that is a component of the NHS	int	1
isprimary	defines overlapping (IsPrimary = 0) and non-overlapping (IsPrimary = 1) TMCs	int	1
active_start_date	active_start_date	datetime	2016-01-01 00:00:00-05:00
active_end_date	active_end_date	datetime	2018-01-01 00:00:00-05:00
urban_rural	U or R if TMC urban rural designation is Urban or Rural	string	U
state	FIPS State code	int	25
county	FIPS County code	int	27
repcty	representative county from National emissions inventory	int	25017
PCT_TYPE10	Vehicle Type 10 (motorcycles) percent of AADT (sums to 1 across 5 HPMS vehicle types and 24 hours)	float	3.20E-05
PCT_TYPE25	Vehicle Type 25 (passenger cars) percent of AADT (sums to 1 across 5 HPMS vehicle types and 24 hours)	float	0.007194146
PCT_TYPE40	Vehicle Type 40 (busses) percent of AADT (sums to 1 across 5 HPMS vehicle types and 24 hours)	float	0.00013238
PCT_TYPE50	Vehicle Type 50 (medium trucks) percent of AADT (sums to 1 across 5 HPMS vehicle types and 24 hours)	float	0.000515824
PCT_TYPE60	Vehicle Type 60 (heavy trucks) percent of AADT (sums to 1 across 5 HPMS vehicle types and 24 hours)	float	0.002350879
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise (sums to 1 across 5 noise vehicle types and 24 hours)	float	0.007194146
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise (sums to 1 across 5 noise vehicle types and 24 hours)	float	7.30E-05
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise (sums to 1 across 5 noise vehicle types and 24 hours)	float	0.002793666

PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise (sums to 1 across 5 noise vehicle types and 24 hours)	float	0.00013238
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise (sums to 1 across 5 noise vehicle types and 24 hours)	float	3.20E-05
tier	Tier 1-4, denoting how TMAS class volume data was matched to TMC link	int	1
dayid	MOVES Day ID	int	5
monthid	MOVES Month ID	int	12
hourid	MOVES Hour ID	int	1
roadtypeid	MOVES roadtype id	int	4
avgspeedbinid	MOVES Average Speed Bin	int	13
monthid3		int	120102
VMT	VMT total on the roadway	float	93.53047413
10_100[48,49]	Emission rate of pollutant 100 (PM_{10}) from Vehicle Type 10 (motorcycles)	float	0.016962
10_110	Emission rate of pollutant 110 ($PM_{2.5}$) from Vehicle Type 10 (motorcycles)	float	0.013992
10_2	Emission rate of pollutant 2 (CO) from Vehicle Type 10 (motorcycles)	float	5.2409
10_3	Emission rate of pollutant 3 (NO_x) from Vehicle Type 10 (motorcycles)	float	0.273095
10_5	Emission rate of pollutant 5 (CH_4) from Vehicle Type 10 (motorcycles)	float	0.006918
10_6	Emission rate of pollutant 6 (N_2O) from Vehicle Type 10 (motorcycles)	float	0.000412
10_87	Emission rate of pollutant 87 (VOC) from Vehicle Type 10 (motorcycles)	float	0.271542
10_90	Emission rate of pollutant 90 (CO_2) from Vehicle Type 10 (motorcycles)	float	134.1873
10_98	Emission rate of pollutant 98 (CO_2 equivalent) from Vehicle Type 10 (motorcycles)	float	134.4605
25_100	Emission rate of pollutant 100 (PM_{10}) from Vehicle Type 25 (passenger cars)	float	0.006899
25_110	Emission rate of pollutant 110 ($PM_{2.5}$) from Vehicle Type 25 (passenger cars)	float	0.004024
25_2	Emission rate of pollutant 2 (CO) from Vehicle Type 25 (passenger cars)	float	0.86889
25_3	Emission rate of pollutant 3 (NO_x) from Vehicle Type 25 (passenger cars)	float	0.09068
25_5	Emission rate of pollutant 5 (CH_4) from Vehicle Type 25 (passenger cars)	float	0.00167
25_6	Emission rate of pollutant 6 (N_2O) from Vehicle Type 25 (passenger cars)	float	0.000317

[48] Tables describing MOVES pollutant and vehicle type IDs used to name this and all remaining columns can be found here: https://github.com/USEPA/EPA_MOVES_Model/blob/master/docs/MOVES3CheatsheetOnroad.pdf

[49] Note that the emissions rate columns will not be included in the output file if DANA is run for Alaska or Hawaii due to missing emission rates and representative counties for those states in the 2017 NEI.

25_87	Emission rate of pollutant 87 (VOC) from Vehicle Type 25 (passenger cars)	float	0.016352
25_90	Emission rate of pollutant 90 (CO_2) from Vehicle Type 25 (passenger cars)	float	118.3027
25_98	Emission rate of pollutant 98 (CO_2 equivalent) from Vehicle Type 25 (passenger cars)	float	118.4362
40_100	Emission rate of pollutant 100 (PM_{10}) from Vehicle Type 40 (busses)	float	0.066876
40_110	Emission rate of pollutant 110 ($PM_{2.5}$) from Vehicle Type 40 (busses)	float	0.051667
40_2	Emission rate of pollutant 2 (CO) from Vehicle Type 40 (busses)	float	0.579334
40_3	Emission rate of pollutant 3 (NO_x) from Vehicle Type 40 (busses)	float	1.741097
40_5	Emission rate of pollutant 5 (CH_4) from Vehicle Type 40 (busses)	float	0.012671
40_6	Emission rate of pollutant 6 (N_2O) from Vehicle Type 40 (busses)	float	0.000507
40_87	Emission rate of pollutant 87 (VOC) from Vehicle Type 40 (busses)	float	0.089306
40_90	Emission rate of pollutant 90 (CO_2) from Vehicle Type 40 (busses)	float	453.4006
40_98	Emission rate of pollutant 98 (CO_2 equivalent) from Vehicle Type 40 (busses)	float	453.8238
50_100	Emission rate of pollutant 100 (PM_{10}) from Vehicle Type 50 (medium trucks)	float	0.062473
50_110	Emission rate of pollutant 110 ($PM_{2.5}$) from Vehicle Type 50 (medium trucks)	float	0.046663
50_2	Emission rate of pollutant 2 (CO) from Vehicle Type 50 (medium trucks)	float	1.272433
50_3	Emission rate of pollutant 3 (NO_x) from Vehicle Type 50 (medium trucks)	float	0.924419
50_5	Emission rate of pollutant 5 (CH_4) from Vehicle Type 50 (medium trucks)	float	0.006776
50_6	Emission rate of pollutant 6 (N_2O) from Vehicle Type 50 (medium trucks)	float	0.000711
50_87	Emission rate of pollutant 87 (VOC) from Vehicle Type 50 (medium trucks)	float	0.087783
50_90	Emission rate of pollutant 90 (CO_2) from Vehicle Type 50 (medium trucks)	float	294.8759
50_98	Emission rate of pollutant 98 (CO_2 equivalent) from Vehicle Type 50 (medium trucks)	float	295.2386
60_100	Emission rate of pollutant 100 (PM_{10}) from Vehicle Type 60 (heavy trucks)	float	0.057921
60_110	Emission rate of pollutant 110 ($PM_{2.5}$) from Vehicle Type 60 (heavy trucks)	float	0.042647
60_2	Emission rate of pollutant 2 (CO) from Vehicle Type 60 (heavy trucks)	float	0.2679
60_3	Emission rate of pollutant 3 (NO_x) from Vehicle Type 60 (heavy trucks)	float	1.302386
60_5	Emission rate of pollutant 5 (CH_4) from Vehicle Type 60 (heavy trucks)	float	0.00817
60_6	Emission rate of pollutant 6 (N_2O) from Vehicle Type 60 (heavy trucks)	float	0.000394
60_87	Emission rate of pollutant 87 (VOC) from Vehicle Type 60 (heavy trucks)	float	0.043465
60_90	Emission rate of pollutant 90 (CO_2) from Vehicle Type 60 (heavy trucks)	float	575.9307
60_98	Emission rate of pollutant 98 (CO_2 equivalent) from Vehicle Type 60 (heavy trucks)	float	576.2245

tier1_class.csv			
Column Name	**Data Description**	**Format Type**	**Example**
STATION_ID	Station Identification	int	101
tmc	The unique 9-digit value identifying the TMC segment	string	129+04375
DIR	Direction of Travel Code	int	1
MONTH	Month of Data	int	10
DAY_TYPE	Weekday/Weekend	string	WD
HOUR	Hour of Data	int	23
PCT_TYPE10	Vehicle Type 10 (motorcycles) percent of AADT	float	0.000005917
PCT_TYPE25	Vehicle Type 25 (passenger cars) percent of AADT	float	0.003722070
PCT_TYPE40	Vehicle Type 40 (busses) percent of AADT	float	0.000047340
PCT_TYPE50	Vehicle Type 50 (medium trucks) percent of AADT	float	0.000071009
PCT_TYPE60	Vehicle Type 60 (heavy trucks) percent of AADT	float	0.000112431
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise	float	0.003722070
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise	float	0.000065092
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise	float	0.000118349
PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise	float	0.000047340
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise	float	0.000005917

tier2_class.csv			
Column Name	**Data Description**	**Format Type**	**Example**
COUNTY	FIPS County Code	int	3
ROUTE_SIGN	Route Sign Name	int	8
ROUTE_NUMBER	Route Number Identification	string	000SR146
MONTH	Month of Data	int	10
DAY_TYPE	Weekday/Weekend	string	WD
PEAKING	Peak in Morning/Afternoon	string	AM
HOUR	Hour of Data	int	23
PCT_TYPE10	Vehicle Type 10 (motorcycles) percent of AADT	float	0.000002955

PCT_TYPE25	Vehicle Type 25 (passenger cars) percent of AADT	float	0.003180073
PCT_TYPE40	Vehicle Type 40 (busses) percent of AADT	float	0.000029555
PCT_TYPE50	Vehicle Type 50 (medium trucks) percent of AADT	float	0.000059109
PCT_TYPE60	Vehicle Type 60 (heavy trucks) percent of AADT	float	0.000150728
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise	float	0.003180073
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise	float	0.000056154
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise	float	0.000153684
PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise	float	0.000029555
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise	float	0.000002955

tier3_class.csv			
Column Name	**Data Description**	**Format Type**	**Example**
URB_RURAL	Urban/Rural	string	U
F_SYSTEM	Functional System Code	int	3
MONTH	Month of Data	int	10
DAY_TYPE	Weekday/Weekend	string	WD
PEAKING	Peak in Morning/Afternoon	string	AM
HOUR	Hour of Data	int	0
PCT_TYPE10	Vehicle Type 10 (motorcycles) percent of AADT	float	0.000164251
PCT_TYPE25	Vehicle Type 25 (passenger cars) percent of AADT	float	0.053997454
PCT_TYPE40	Vehicle Type 40 (busses) percent of AADT	float	0.003777769
PCT_TYPE50	Vehicle Type 50 (medium trucks) percent of AADT	float	0.001847822
PCT_TYPE60	Vehicle Type 60 (heavy trucks) percent of AADT	float	0.010224613
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise	float	0.053997454
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise	float	0.001314006
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise	float	0.010758428
PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise	float	0.003777769
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise	float	0.000164251

tier4_class.csv			
Column Name	**Data Description**	**Format Type**	**Example**
URB_RURAL	Urban/Rural	string	U
F_SYSTEM	Functional System Code	int	3
MONTH	Month of Data	int	10
DAY_TYPE	Weekday/Weekend	string	WD
PEAKING	Peak in Morning/Afternoon	string	AM
HOUR	Hour of Data	int	0
PCT_TYPE10	Vehicle Type 10 (motorcycles) percent of AADT	float	0.00032127
PCT_TYPE25	Vehicle Type 25 (passenger cars) percent of AADT	float	0.00602237
PCT_TYPE40	Vehicle Type 40 (busses) percent of AADT	float	0.00011841
PCT_TYPE50	Vehicle Type 50 (medium trucks) percent of AADT	float	0.00031556
PCT_TYPE60	Vehicle Type 60 (heavy trucks) percent of AADT	float	0.00384876
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise	float	0.00602237
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise	float	0.00022525
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise	float	0.00393907
PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise	float	0.00011841
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise	float	0.00032127

XX_Composite_Emissions_SUMMARY.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
tmc	TMC Link identifier	string	129N04633
road	Name of the roadway, usually the interstate system designation	string	I-395
tmc_length	the length of the TMC segment	float	0.6003
Average_Speed	Average speed over the full year	float	49.22254
Average AADT	Full year AADT	float	62048.5
TotEmissionsPerMile 2[50]	Total estimated emissions per roadway mile for pollutant 2 (CO)	float	1028475.19

[50] Note that the estimated emissions columns will appear as zeros in the output file if DANA is run for Alaska or Hawaii due to missing emission rates and representative counties for those states in the 2017 NEI.

TotEmissionsPerMile_3	Total estimated emissions per roadway mile for pollutant 3 (NO_x)	float	181619.7831
TotEmissionsPerMile_5	Total estimated emissions per roadway mile for pollutant 5 (CH_4)	float	2758.534895
TotEmissionsPerMile_6	Total estimated emissions per roadway mile for pollutant 6 (N_2O)	float	824.8526767
TotEmissionsPerMile_87	Total estimated emissions per roadway mile for pollutant 87 (VOC)	float	33847.90502
TotEmissionsPerMile_90	Total estimated emissions per roadway mile for pollutant 90 (CO_2)	float	189379548.4
TotEmissionsPerMile_98	Total estimated emissions per roadway mile for pollutant 98 (CO_2 equivalent)	float	189693182
TotEmissionsPerMile_100	Total estimated emissions per roadway mile for pollutant 100 (PM_{10})	float	26590.4183
TotEmissionsPerMile_110	Total estimated emissions per roadway mile for pollutant 110 ($PM_{2.5}$)	float	10137.31181
geometry	Geography of roadway link	string	LINESTRING (-71.07650579999995 42.38359160000005, -71.07653619999996 42.38384420000006, …, -71.08293899999995 42.39086300000002)

5.2 PROCESS 2

XX_MONTH_VMT.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
County	FIPS County Code	int	11
monthID	Month of Data	int	9
sourceTypeID	MOVES vehicle type identification	int	11
monthVMTFraction	Monthly VMT Fraction	float	0.020164844

XX_DAY_VMT.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
county	FIPS County Code	int	11
monthID	Month of Data	int	9
roadTypeID	MOVES functional class identification	int	2
dayID	Day of Data Code	int	2
sourceTypeID	MOVES vehicle type identification	int	11
dayVMTFraction	Daily VMT Fraction	float	0.146529563

XX_HOUR_VMT.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
county	FIPS County Code	int	11
roadTypeID	MOVES functional class identification	int	2
dayID	Day of Data Code	int	2
hourID	Hour of Data	int	1
sourceTypeID	MOVES vehicle type identification	int	11
hourVMTFraction	Hourly VMT Fraction	float	0.009950249

XX_REGIONAL_VMT.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
state	FIPS State Code	int	25
county	FIPS County Code	int	1
yearID	Year of Data	int	2015
baseYearOffNetVMT	always set to zero	int	0
HPMSBaseYearVMT	VMT for the given year	float	57880027.8996705
HPMStypeID	MOVES VMT Type Code	int	50

XX_ROADTYPE_VMT.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
county	FIPS State Code	int	1
roadTypeID	MOVES functional class identification	int	3
sourceTypeID	MOVES vehicle type identification	int	11
roadTypeVMTFraction	VMT Fraction by road type	float	0.007343567

XX_MONTH_VMT_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
sourceTypeID	MOVES vehicle type identification	int	11
monthID	Month of Data	int	9
monthVMTFraction	Monthly VMT Fraction	float	0.020164844

XX_DAY_VMT_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
sourceTypeID	MOVES vehicle type identification	int	11
monthID	Month of Data	int	9
roadTypeID	MOVES functional class identification	int	2
dayID	Day of Data Code	int	2
dayVMTFraction	Daily VMT Fraction	float	0.146529563

XX_HOUR_VMT_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
sourceTypeID	MOVES vehicle type identification	int	11
roadTypeID	MOVES functional class identification	int	2
dayID	Day of Data Code	int	2
hourID	Hour of Data	int	1
hourVMTFraction	Hourly VMT Fraction	float	0.009950249

XX_REGIONAL_VMT_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
HPMSVtypeID	MOVES VMT Type Code	int	50
yearID	Year of Data	int	2015
HPMSBaseYearVMT	VMT for the given year	float	57880027.8996705

XX_ROADTYPE_VMT_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
sourceTypeID	MOVES vehicle type identification	int	11
roadTypeID	MOVES functional class identification	int	3
roadTypeVMTFraction	VMT Fraction by road type	float	0.783252633

5.3 PROCESS 3

XX_SPEED_DISTRIBUTION.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
county	FIPS County Code	int	17
sourceTypeID	MOVES vehicle type identification	int	11
roadtypeid	MOVES functional class identification	int	3
hourdayID	MOVES hourly temporal identification	int	102
avgSpeedBinID	MOVES default speed bins	int	4
avgSpeedFraction	proportion of vehicles that travel at within each of the speed bins for a specific road typeID and hourDayID for the reference county	float	0.01254763

XX_SPEED_DISTRIBUTION_YY.csv, where XX represents the state abbreviation and YY represents the county code			
Column Name	**Data Description**	**Format Type**	**Example**
sourceTypeID	MOVES vehicle type identification	int	11
roadtypeid	MOVES functional class identification	int	3
hourdayID	MOVES hourly temporal identification	int	102
avgSpeedBinID	MOVES default speed bins	int	4
avgSpeedFraction	proportion of vehicles that travel at within each of the speed bins for a specific road typeID and hourDayID for the reference county	float	0.01254763

5.4 PROCESS 4

XX_Composite_Emissions_select.csv, where XX represents the state abbreviation			
Column Name	**Data Description**	**Format Type**	**Example**
tmc	TMC Link identifier	string	129+04375
date	Date of data	date	1/1/2018
hour	hour of day	int	0
road	Name of the roadway, usually the interstate system designation	string	I-95
direction	NB, SB, EB, WB standing for north, south, east and west bound TMC links	string	NB
state	State FIPS Code	int	25
county	County FIPS Code	int	17
start_latitude	The starting latitude of the TMC link	float	42.499986
start_longitude	The starting longitude of the TMC link	float	-71.152304
end_latitude	The ending latitude of the TMC link	float	42.502387
end_longitude	The ending longitude of the TMC link	float	-71.1325082
tmc_length	the length of the TMC segment	float	1.022317
road_order	Numerical value indicating in what order the TMC segment would be encountered when traveling downstream relative to the other TMC segments on the same road	int	1331
f_system	the FHWA-approved Functional Classification System code If multiple HPMS segments with different attribute values are assigned to a single TMC path, the value for the highest functional class (minimum code value) is assigned	int	1
thrulanes	the number of lanes designated for through-traffic in both travel directions	int	8
aadt	Annual Average Daily Traffic on the TMC link	float	87582.5
aadt_singl	AADT of single unit trucks	int	3858
aadt_combi	AADT of combination trucks	int	7188
travel_time_all	all-vehicle travel time	float	58.65
travel_time_pass	passenger car travel time	float	58.2
travel_time_truck	truck travel time	float	61.34
speed_all	all-vehicle speed	float	62.750916
speed_pass	passenger car speed	float	63.236103

speed_truck	truck speed	float	59.999043
PCT_NOISE_AUTO	Percent of AADT that is Auto vehicle type for noise	float	0.003722070
PCT_NOISE_MED_TRUCK	Percent of AADT that is Medium Trucks vehicle type for noise	float	0.000065092
PCT_NOISE_HVY_TRUCK	Percent of AADT that is Heavy Trucks vehicle type for noise	float	0.000118349
PCT_NOISE_BUS	Percent of AADT that is Bus vehicle type for noise	float	0.000047340
PCT_NOISE_MC	Percent of AADT that is Motorcycles vehicle type for noise	float	0.000005917

6 APPENDIX E. DEVELOPMENT OF DEFAULT EMISSIONS RATES FROM 2017 NEI DATA

Each row in the link level dataset output by Process 1 includes emissions rates for five HPMS vehicle types that are looked up based on the following parameters in that row:

- county,
- roadway type,
- average speed,
- month,
- weekday or weekend day type, and
- hour of the day.

The DANA Tool performs this lookup during Process 1 using the Emission Rates input to that process. The DANA Tool comes with a default emission rate input table that has rates by the above parameters. The rest of this appendix documents how the default emission rate table was created.

The default emission rate table uses a set of representative counties throughout the country. The representative counties are a subset of U.S. counties that are chosen based on the similarity of their meteorology to other counties for the year the National Emissions Inventory is performed. Using emissions rates from representative counties reduces the input database size by a factor of about 10 and makes the tool easier to use.

To obtain emissions rates for representative counties, data aggregation, combination, and manipulation was performed on public data from EPA's 2017 National Emissions Inventory (NEI). The process involves downloading large amounts of data from EPA's public file transfer protocol site and running queries and processing algorithms on that data using Python scripts. The following steps were completed:

1. The emissions factors are taken from emissions factors tables available at the following site: https://gaftp.epa.gov/Air/emismod/2017/2017emissions/moves_eftables/. These zip files contain rate-per-distance CSV databases for each representative county. The emission rates columns in these CSV files give the grams-per-mile-driven for a large number of pollutants at temperatures ranging from the lowest to the highest temperature experienced in that representative county in 5 degree Fahrenheit intervals.
2. These files were processed by dropping the unnecessary pollutants, assigning MOVES pollutant IDs to the remaining pollutants, and changing the format of the data from wide to long format where one emission rate for a single pollutant is stored in each row. Pollutants that are included in the final emissions rates dataset are "CO_INV", "NOX", "VOC_INV", "PM10", "BRAKEPM10", "TIREPM10", "PM2_5", "PM25BRAKE", "PM25TIRE", "CO2_INV", "CH4_INV", "N2O_INV."
3. Brakewear and tirewear PM emissions are aggregated into overall PM emissions rates during this process. This involved combining pollutant IDs 100, 106, and 107 into an overall PM10 emission rate, and IDs 110, 116, and 117 into an overall PM2.5 emission rate.
4. The rate-per-distance data is also separated by different fuel types. In order to create rates usable in the DANA Tool, the disaggregated emissions rates were weighted by fuel type

population fractions obtained from EPA (see Table 8 below), and then aggregated together to create an average emission rate for each vehicle type across all fuel types.

Table 8. Emissions and weightings by fuel type.

HPMSVtypeID	sourceTypeID	fuelTypeID	weight
10	11	1	1
25	21	1	0.478504464
25	21	2	0.003394789
25	21	5	0.000263097
25	21	9	0.000553189
25	31	1	0.448890746
25	31	2	0.015107857
25	31	5	0.000931073
25	31	9	1.06E-05
25	32	1	0.04861708
25	32	2	0.003566385
25	32	5	0.00015798
25	32	9	2.71E-06
40	41	2	0.326128987
40	42	1	0.039290949
40	42	2	0.150545521
40	42	3	0.009608678
40	43	1	0.037770108
40	43	2	0.436655756
50	51	1	0.00041291
50	51	2	0.022577004
50	52	1	0.131087302
50	52	2	0.438974035
50	53	1	0.090051887
50	53	2	0.289870407
50	54	1	0.019734772
50	54	2	0.007291682
60	61	1	0.001083251
60	61	2	0.353695351
60	62	2	0.645221399

5. Meteorology data was also obtained from EPA's file storage site. This dataset is available at the following link: https://gaftp.epa.gov/Air/emismod/2017/2017emissions/. The meteorology data contains the temperatures in all counties in the US for every hour of the day for every day of the year. Only the data for the representative counties were retained to reduce the size of the final dataset.
6. Finally, the processed and aggregated emissions rates data were merged onto the meteorology data based on the temperature of that hour of the day. In this way, hourly

emissions rates for the entire year were obtained for all representative counties in the U.S. This final table is used as input to the DANA tool.

7 APPENDIX F. COMPARISON OF SPEED DISTRIBUTIONS DERIVED FROM NATIONAL PERFORMANCE MANAGEMENT RESEARCH DATA SET (NPMRDS) VERSUS STREETLIGHT DATA

Note: Text in this appendix comes from the original research entitled "National Traffic Dataset Applications for Air Quality and Noise Analysis", which led to the creation of the DANA tool. This original research was performed by Cambridge Systematics and Eastern Research Group (ERG).

7.1 BACKGROUND

Previously, team member ERG developed speed distributions using data from the vendor StreetLight as part of CRC project A-100.[51] The goal was to develop speed distributions for the 2014 National Emissions Inventory (NEI). The StreetLight data is similar to the NPMRDS data, in that, it is based on probe vehicle or global positioning system (GPS) measurements from mobile devices, but it includes speed data on non-NHS roads in addition to NHS roads. These facts point out the problem with using the NPMRDS data for county-level speed distributions by covering only NHS roads, which are the higher order functionally classified roads; the resulting speed distribution is likely to be biased toward higher speeds. Table 9 shows the percentage of different functional class roads that are designated as on the NHS.

Table 9. Functional classification of National Highway System roadways.

Functional Class	Percent of Mileage that Are Designated as NHS
Interstate	100%
Freeways and Expressways	99%
Principal Arterial	95%
Minor Arterial	3%
Major Collector	<1%
Minor Collector	<1%
Local	0%

(Source: 2016 HPMS Data.)

As part of the current project, the team obtained these data for Colorado in order to compare the speed distributions that are produced from the two data sources. StreetLight data covered the 12-month period of September 2015 through August 2016. NPMRDS data, which were provided by HERE, were obtained for the same period (version 1 of the NPMRDS). The MOVES model uses the four road types listed below, differentiating based on whether the road is located within an urban or rural area, and whether it has restricted access points (i.e., by ramps), or unrestricted access, such as entry points at multiple intersections:

[51] Eastern Research Group, *Improvement of Default Inputs for MOVES and SMOKE-MOVES: CRC Project A-100*, prepared for the Coordinating Research Council, February 28, 2017, http://crcsite.wpengine.com/wp-content/uploads/2019/05/ERG_FinalReport_CRCA100_28Feb2017.pdf.

- MOVES Road type 2: Rural Restricted Access Roads.
- MOVES Road type 3: Rural Unrestricted Access Roads.
- MOVES Road type 4: Urban Restricted Access Roads.
- MOVES Road type 5: Urban Unrestricted Access Roads.

7.2 RESULTS

The analysis was conducted using the following parameters for counties in Colorado:

- Source type = 21 (passenger cars), 31 (passenger truck), and 62 (Combination Long-Haul Truck) combined. These source types were selected to match the data for "all vehicles" in the NPMRDS.
- Road type = 3 and 5 (rural and urban arterials).
- Day of week = 2 and 5 (weekday and weekend).
- Hour = 1 to 24.

About 56 of Colorado's 64 counties were matched between the datasets. Most counties only had rural arterials. In all, 1,343 combinations were studied. MOVES speed distributions were created for each of these combinations using the NPMRDS and StreetLight data. Because of the analytical complexity of comparing so many distributions visually, we employed the Komolgorov-Smirnov (K-S) two sample test to detect differences in the pairwise distributions.[52]

The SAS procedure NPAR1WAY was used for this purpose. The K-S test computes a statistic (K) based on the difference in the distribution and compares it to a critical value (D_{crit}). The probability of obtaining a value higher than the critical value (P) is then computed under the null hypothesis that the distributions are the same. When the probability is low, say 0.01, typically the null hypothesis is rejected, that is, the distributions are different. Conversely, when the probability is high, the null hypothesis cannot be rejected, and we conclude that the distributions are statistically the same. Therefore, for the purpose of this comparison, we want to obtain high probability numbers because it would be desirable if the distributions from the two datasets were the same. Figure 48 shows an example of two distributions that were tested to be statistically equivalent. Figure 49 shows an example of two distributions that were tested to be statistically different.

[52] https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test#Two-sample_Kolmogorov%E2%80%93Smirnov_test.

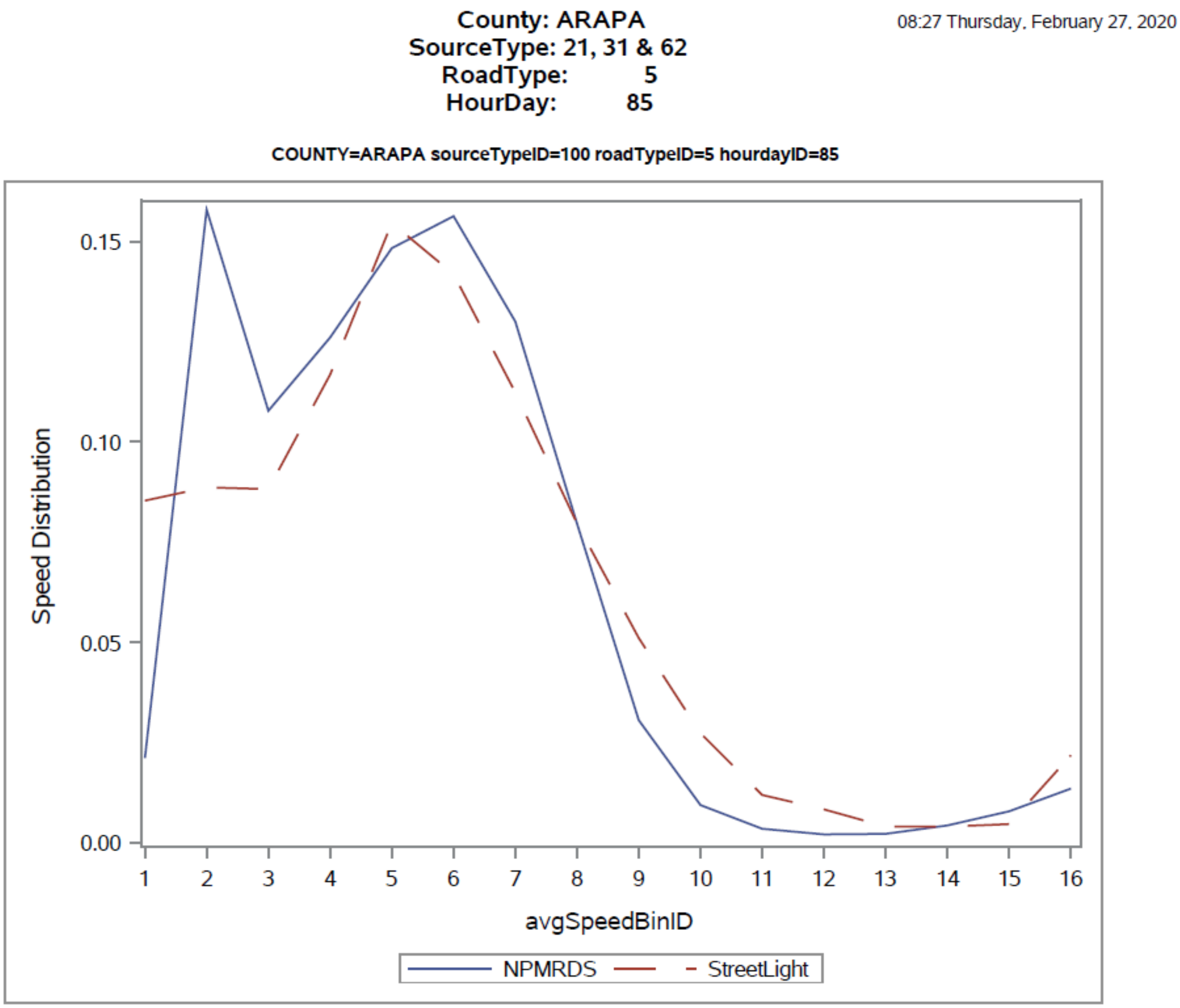


Figure 48. Chart. Example: statistically equivalent distributions.
(Source: Federal Highway Administration.)

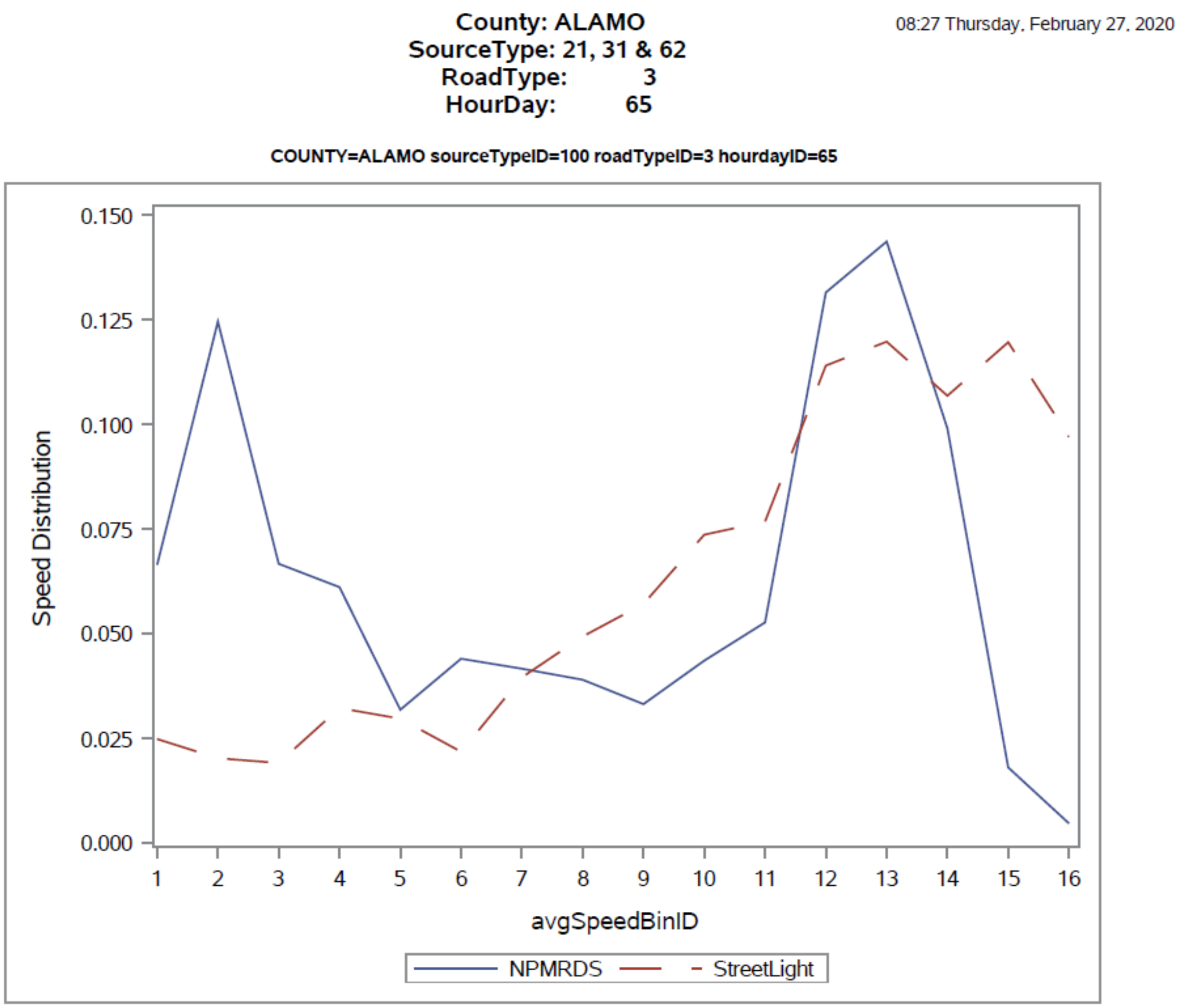


Figure 49. Chart. Example: statistically different distributions.
(Source: Federal Highway Administration.)

The 1,343 comparisons were summarized based on their K-S tests. We selected 0.1 as the cutoff alpha level. If the probability was less than or equal to 0.1, we conclude that the distributions are statistically different. If the probability was greater than 0.1, we conclude that the distributions are the same. The results were further broken down by urban versus rural and peak versus offpeak times. The results are shown in Table 10 through Table 14. Comparing Table 11 and Table 12 (urban versus rural), it can be seen that in rural areas, 62 percent of arterials have statistically different speed distributions, while in urban arterials all of the locations tested showed similar speed distributions.

Table 10. Komolgorov-Smirnov test results for all Colorado arterials.

Prob(K) > D_{crit}		**Number of Sections**	**Percentage**	**Cumulative (Percent)**
>=0	<0.1	746	56%	56%
>=0.1	<0.2	0	0%	56%
>=0.2	<0.3	137	10%	66%
>=0.3	<0.4	0	0%	66%
>=0.4	<0.5	126	9%	75%
>=0.5	<0.6	0	0%	75%
>=0.6	<0.7	123	9%	84%
>=0.7	<0.8	0	0%	84%
>=0.8	<0.9	0	0%	84%
>=0.9	<1	211	16%	100%
Total		**1,343**	**100%**	

(Note: Sections in the 0-0.1 probability range are deemed to be different.)

Table 11. Komolgorov-Smirnov tests for urban Colorado arterials.

Prob(K) > D_{crit}		**Number of Sections**	**Percentage**	**Cumulative (Percent)**
>=0	<0.1	0	0%	0%
>=0.1	<0.2	0	0%	0%
>=0.2	<0.3	0	0%	0%
>=0.3	<0.4	0	0%	0%
>=0.4	<0.5	5	4%	4%
>=0.5	<0.6	0	0%	4%
>=0.6	<0.7	16	11%	15%
>=0.7	<0.8	0	0%	15%
>=0.8	<0.9	0	0%	15%
>=0.9	<1	119	85%	100%
Total		**140**	**100%**	

Table 12. Komolgorov-Smirnov tests for rural Colorado arterials.

Prob(K) > D_{crit}		**Number of Sections**	**Percentage**	**Cumulative (Percent)**
>=0	<0.1	746	62%	62%
>=0.1	<0.2	0	0%	62%
>=0.2	<0.3	137	11%	73%
>=0.3	<0.4	0	0%	73%
>=0.4	<0.5	121	10%	83%
>=0.5	<0.6	0	0%	83%
>=0.6	<0.7	107	9%	92%
>=0.7	<0.8	0	0%	92%
>=0.8	<0.9	0	0%	92%
>=0.9	<1	92	8%	100%
Total		**1,203**	**100%**	

Table 13. Komolgorov-Smirnov tests for peak periods on Colorado arterials.

Prob(K) > D_{crit}		**Number of Sections**	**Percentage**	**Cumulative (Percent)**
>=0	<0.1	120	54%	54%
>=0.1	<0.2	0	0%	54%
>=0.2	<0.3	20	9%	63%
>=0.3	<0.4	0	0%	63%
>=0.4	<0.5	28	13%	75%
>=0.5	<0.6	0	0%	75%
>=0.6	<0.7	17	8%	83%
>=0.7	<0.8	0	0%	83%
>=0.8	<0.9	0	0%	83%
>=0.9	<1	39	17%	100%
Total		**224**	**100%**	

Table 14. Komolgorov-Smirnov tests for offpeak periods on Colorado arterials.

Prob(K) > D_{crit}		**Number of Sections**	**Percentage**	**Cumulative (Percent)**
>=0	<0.1	626	56%	56%
>=0.1	<0.2	0	0%	56%
>=0.2	<0.3	117	10%	66%
>=0.3	<0.4	0	0%	66%
>=0.4	<0.5	98	9%	75%
>=0.5	<0.6	0	0%	75%
>=0.6	<0.7	106	9%	85%
>=0.7	<0.8	0	0%	85%
>=0.8	<0.9	0	0%	85%
>=0.9	<1	172	15%	100%
Total		**1,119**	**100%**	

7.3 RECOMMENDATIONS

The analysis showed a difference in speed distributions developed from the 2 sources; overall 44 percent of the speed distributions for the arterial combinations were statistically the same. However, the results were dramatically different for urban versus rural highways. All of the 140 urban cases exhibited similar speed distributions, while only 38 percent of rural cases had similar speed distributions. The fact that a higher difference was detected in rural areas could be a function of three factors:

1. The NPMRDS includes only NHS highways, which are higher functionally classified roadways (e.g., arterials). The StreetLight data includes lower functionally classified highways (e.g., collectors).
2. Sample size: probes are less likely to be observed on low volume highways.

3. Rural arterials' speed distributions are skewed toward higher speeds, and the two data sources are showing the differences there. Analysts should thoroughly review speed distributions from either of these two sources when studying rural situations.

The data used in this comparison is now over three years old. We have observed that vehicle probe data available from vendors has consistently improved in quality over time, as more vehicle probes are recruited, and vendors improve their processing methods. For example, since 2017, the NPMRDS version 2 is provided by a different vendor (Inrix). Not only are the probe sources different from version 1, but the processing methods also are different. Further, it is not possible to determine which dataset more closely replicates reality; an independent and controlled data collection on speeds was not available.

However, the low similarity rate for the speed distributions found on rural arterials cannot be ignored. **It is recommended that users carefully review probe vehicle data in rural areas prior to developing speed distributions for MOVES inputs.** An easy check is to observe the percentage of speeds in the lower speed ranges. In rural areas, large percentages of speeds less than 10 to 15 mph over the course of a year would only be due to a long-term work zone or other lane closure. An indication of low sample sizes for probes is numerous time slots in the raw probe data with no speeds.