CRC Report No. A-115

DEVELOPING IMPROVED VEHICLE POPULATION INPUTS FOR THE 2017 NATIONAL EMISSIONS INVENTORY

APRIL 2019

COORDINATING RESEARCH COUNCIL, INC.
5755 NORTH POINT PARKWAY • SUITE 265 • ALPHARETTA, GA 30022

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The four main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance; heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); and light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars). CRC's function is to provide the mechanism for joint research conducted by the two industries that will help in determining the optimum combination of petroleum products and automotive equipment. CRC's work is limited to research that is mutually beneficial to the two industries involved. The final results of the research conducted by, or under the auspices of, CRC are available to the public.

Eastern Research Group, Inc.

Developing Improved Vehicle Population Inputs for the 2017 National Emissions Inventory: CRC Project A-115

Final Report

Prepared for:

The Coordinating Research Council

Prepared by:

Eastern Research Group, Inc.

April 11, 2019

ERG No. 4169.00.001

Developing Improved Vehicle Population Inputs for the 2017 National Emissions Inventory: CRC Project A-115

Final Report

Prepared for:

Coordinating Research Council, Inc.
3650 Mansell Road, Suite 140
Alpharetta, GA 30022

Prepared by:

Allison DenBleyker
Meredith Weatherby
Eastern Research Group, Inc.
3508 Far West Blvd., Suite210
Austin, TX 78731

April 11, 2019

Table of Contents

List of Tables

List of Figures

1.0 Introduction

The Coordinating Research Council (CRC) directed four recent projects to improve the on-road sector of the U.S. National Emissions Inventory (NEI). CRC contracted with Eastern Research Group, Inc. (ERG) to undertake each of these. The first, CRC project A-84, analyzed MOVES input data submitted by 30 states for over 1,400 counties in Version 1 of the 2011 NEI (ERG, 2013). The study found a high degree of variability in the state-supplied data, reported best practices for the input data and conducted an emissions sensitivity. The work highlighted the importance of local data for use in developing emission inventories with MOVES. The second study, CRC project A-88, built upon A-84 by developing improved default data that was applied at the local level for version 2 of the 2011 NEI (ERG, 2014). The A-88 study produced year 2011 light-duty vehicle age distributions and populations, as well as long-haul truck vehicle-miles travelled (VMT) allocations for each county in the US. The third study, CRC project A-100, further improved county level data for the NEI by transforming vehicle telematics data to generate local vehicle speed distributions and patterns of VMT for use in MOVES and SMOKE.

This work, CRC project A-115, prepares new MOVES inputs from 2017 vehicle registration data for use in the 2017 NEI and 2016 v1 modeling platforms. The registration data for this effort was provided by data firm IHS Markit (formerly Polk Automotive). In the past, IHS Markit provided their 2011 registration data to ERG for CRC project A-88 and their 2014 data to U.S. EPA for the 2014 NEI.

A major goal of the new 2017 registration data effort was to incorporate data-driven adjustments to reduce IHS Markit's possible overestimation of older populations of light duty vehicles. The products coming out of this study include three nationwide, county-level data tables formatted for use in MOVES and an additional summary that includes the gross vehicle weight rating (GVWR) category. These outputs were generated for an unadjusted baseline and several population-reduction scenarios described later in the report.

Work Product	Purpose
1. County-level MOVES 'SourceTypeYear' table	Total populations by vehicle class (source type)
2. County-level MOVES 'SourceTypeAgeDistribution' table	Fractions of source type population by age 0 years old through 30+ years old
3. County-level MOVES 'AVFT' table	Fractions of fuel type (e.g., gasoline vs. diesel) by source type and model year
4. Summary of vehicle counts by MOVES source type, model year, fuel type, GVWR	Population summary providing extra information on gross vehicle weight rating

The report begins with an overview of how vehicle registration data are used in on-road modeling and the results of the recent 2014 VIN analysis. After the overview, Section 3.0 describes the methodology to transform the 2017 data into MOVES inputs and estimate 2016 population and 2016 age distribution. Section 4.0 discusses the results and shows the impacts of the population reduction scenarios on average age and population totals. Section 5.0 finishes with a brief summary and recommendations.

2.0 Overview of Vehicle Registration Data

All vehicles in the U.S. must be registered as a condition for use on public roads. Each state has a department of motor vehicles (DMV) or equivalent agency to manage the registration process, and registration rules vary. For example, most states require annual registration renewal, but some have a biennial schedule or allow registrations for certain vehicle types to remain valid for 3 years or even indefinitely. These differences are important because a registration renewal signals that a vehicle is still active.

In addition, data handling practices vary among state DMVs. Differences in data handling include whether a DMV works primarily on paper records or electronically, as well as how quickly new vehicles are added to and removed (if removed at all) from the database. Often historical data remain permanently on the record, so judgment is required to assess whether a vehicle is currently operating vs. a late renewal.

Retrieving data from a centrally compiled source (IHS Markit) is convenient because of the large effort involved to contact 50 individual DMVs, work through privacy restrictions, and understand the nuances of each states' practices. In addition to convenience, a benefit of central compilation is that all VINs are decoded consistently. In the past national modeling efforts, EPA found that some states classify light trucks as cars, and vice versa, which can cause regional planning challenges because the emissions sources aren't consistent or fairly accounted for. A downside of the centrally compiled source is that the specific data processing steps and judgments are largely propriety, and the resulting assessment of vehicles in operation doesn't always align with similar work independently performed by state environmental agencies working from their DMV registration data. The 2014 VIN analysis highlights some of the differences seen in the three states of New Jersey, Maine, and Missouri.

2.1 Importance for the On-road Inventory

Vehicle registration data are invaluable for on-road mobile source emission inventories because they provide the number of operating vehicles, their age, and fuel type; all of which are important to classify the fleet and estimate emissions.

The recent NEIs and modeling platforms calculate the on-road emissions inventories by multiplying MOVES emission factors (grams of pollutant per unit activity) with activity. For some emission processes such as engine starts and fuel evaporation while parked, direct vehicle population is the activity combined with emission factors. For other emissions that occur over-the-road, such as exhaust from the tailpipe and crankcase, or fuel evaporation during driving, the activity is VMT. Registration data impacts both population and VMT, influencing the latter by changing how much VMT the MOVES model allocates to different vehicle ages and fuel types, which has a large impact on emissions.

2.2 The 2014 Data VIN Analysis

The 2014 data VIN analysis was outside the scope of CRC project A-115, but the results were provided to inform adjustments to the 2017 registration data in this work.

Agencies from three states convened in 2018 (hereafter called the "VIN workgroup") to compare their analysis of DMV registration data with the 2014 data IHS Markit provided for the 2014 NEI version 2. For each of the three states, the IHS Markit population estimates were higher than the state, with increasing discrepancies with older vehicle age. The VIN workgroup looked at differences in the estimates separately for passenger cars and light trucks in each state and selected several samples of passenger cars where the discrepancies were particularly high and requested that IHS Markit provide the following 100 individual VINs listed below, so that the VIN workgroup could investigate what happened to these vehicles in 2014.

Subset No.	Model Year, Make, Model, County, and State	2014 IHS Count	2014 State Count
1	1990 Ford Escorts in Hudson County, NJ	24	4
2	1987 Chevrolet Caprice/Classic Somerset County, NJ	23	6
3	2000 Subaru Legacy in Aroostook County, ME	21	2
4	1992 Ford Thunderbird Jackson County, MO	19	6
5	1987 GMC Jimmy in Jackson County, MO	4	1
6	1994 Dodge Dakota in Androscoggin, ME	9	2
Total		**100**	**21**

The VIN workgroup purchased Carfax reports for the subset of vehicles that were in one set (IHS or State) but not the other. The VIN workgroup considered the Carfax reports as the truth standard for whether the vehicle was active. The Carfax reports note not only when a vehicle owner renews its registration, but also any inspections, title changes (to "junk" or "salvage" for example), and routine oil changes. The conclusion of the VIN level checks was that although some states needed correction to their data, overall the IHS Markit set included more vehicles than the state and Carfax report indicated should be still in operation. While the VIN workgroup analyzed all three state trends, their final recommendation was to use a New Jersey statewide adjustment by model year, separately for Passenger Car vs. Passenger Truck/Light Commercial Truck MOVES source types[1]. The recommended model year reductions are shown later in the Methodology section of this report, Table 3-4. The national impact on 2017 populations by model year is also shown for Passenger Cars (Figure 3-1), Passenger Trucks (Figure 3-2), and Light Commercial Truck (Figure 3-3).

[1] Mark Janssen, LADCO, provided the adjustment factors in the file *Summary_Final_proposal_v1.0_12062018.xlsx* via email on December 6, 2018.

3.0 Methods

3.1 Source Data

IHS Markit provided three files to the CRC containing the 2017 vehicle populations for this work: one each for motorcycles (MC), light duty (LD) vehicles, and heavy duty (HD) vehicles. The MC file contains only motorcycle population, and it contained 4.1 million records with a total population of 13,001,738 motorcycles operating in the US in 2017. Unlike the LD and HD deliveries, the MC delivery file represents a January 2017 data draw of registration data. The LD and HD data are a snapshot from July 2017. Table 3-1 lists each of the data field names along with a description of the data.

Table 3-1. Description of the IHS Data Delivery for Motorcycles

IHS Data Field Name	Description of the Data Values
STATE NAME	51 states (includes District of Columbia)
STATE/COUNTY CODE	3,141 5-digit state-county codes
COUNTY NAME	3,141 counties
YEAR MODEL	51 model years (1968 – 2018)
MAKE NAME	149 makes (e.g., HARLEY-DAVIDSON, HONDA, YAMAHA, SUZUKI, POLARIS etc.)
MODEL NAME	1,787 models (e.g., CL350, DS, YZF1000, etc.)
CYLINDER	6 values (0,1,2,3,4,6)
ENG DSPLCMNT	484 values (e.g., 63, 260, 740, etc.)
STROKE	5 values (0,1,2,4,5)
VIO	Vehicles in Operation, the population for each of the 4.1 million combinations

The LD dataset contained 35.8 million records of data, capturing all cars, light-duty trucks, and "light-heavy" duty (Class 3) trucks operating in the U.S. in 2017, for a total population of 270,902,844 vehicles. The 35.8 million records categorize the vehicle populations by state and county of registration, vehicle type, owner type, model year, make, model, gross vehicle weight rating (GVWR), body style, cab configuration, fuel type, number of engine cylinders, engine displacement, 2-wheel vs. 4-wheel drive, IHS's assessment of MOVES source type ID, 12 different fuel types, antique designation, and the population. Table 3-2 lists each of the field names and a description of the content.

Table 3-2. Description of the IHS Data Delivery for Light Duty Vehicles

IHS Data Field Name	Description of the Data Values
STATE NAME	51 states (includes District of Columbia)
COUNTY NAME	3,142 counties
IHS VEH TYPE	2 values: “CAR”, “LIGHT TRUCK”
OWNER TYPE	“COMMERCIAL” or “PERSONAL”
YEAR MODEL	58 model years (1961 – 2018)
MAKE NAME	177 makes (e.g., TOYOTA, DODGE, BUICK, CHEVROLET, GMC, JEEP, etc.)
MODEL NAME	1,787 models (e.g., PRIUS, AVENGER, LACROSS, EQUINOX, SIERRA, CHEROKEE, etc.)
GVW CDE	9 values (blank, 1, 1a, 1b, 2, 2a, 2b, 3, 3a, 3b)
BODY STYLE NAME	34 values (e.g., HATCHBACK SEDAN, SPORT UTILITY VEHICLE, etc.)
CAB CONFIG DESC	15 values (e.g., UNKNOWN, VAN, SPORT UTILITY, EXTENDED CAB, etc.)
MOVES FUEL CODE	12 values (blank, C, D, E, G, H, L, M N P X Y)
MOVES FUEL DESC	12 values (blank, CONVERTIBL, DIESEL, ELECTRIC, GAS, GAS HYBRID, ETHANOL, METHANOL, CNG, PROPANE, GAS, HYDROGEN)
CYLINDER	10 values (0,2,3,4,5,6,8,10,12,16)
ENG DSPLCMNT CI	293 values (e.g, 85, 158, 284, 379, etc.)
VOCATION	All blank.
DRIVE WHEEL	3 values (blank, 4X2, and 4X4)
MOVES VEHICLE TYPE CODE	3 values (21, 31, 32)
MOVES VEHICLE TYPE DESCRIPTION	3 values (Passenger Car, Passenger Truck, Light Commercial Truck)
IHS FUEL TYPE	12 values (blank, CONVERTIBLE, DIESEL, ELECTRIC, GAS ELECTRIC AND GAS HYBRID, ETHANOL, METHANOL, COMPRESSED NATURAL GAS, PROPANE, FLEXIBLE, HYDROGEN FUEL CELL)
ANTIQUE	2 values (Y, N)
VIO	Population for each of the 35.8 million combinations

The HD dataset contained 4.4 million records and a total population of 8,184,144 vehicles. These populations were categorized by the data fields listed below in Table 3-3.

Table 3-3. Description of the IHS Data Delivery for Heavy Duty Vehicles

IHS Data Field Name	Description of the Data Values
STATE NAME	51 (includes District of Columbia)
COUNTY NAME	3,142 counties
IHS VEH TYPE	17 values (e.g., TRACTOR TRUCK, STRAIGHT TRUCK, BUS SCHOOL, FIRE TRUCK, etc.)
OWNER TYPE	1 value "COMMERCIAL"
YEAR MODEL	53 model years (1966 – 2018)
MAKE NAME	88 makes (e.g., INTERNATIONAL, KENWORTH, MACK, PETERBILT, VOLVO, etc.)
MODEL NAME	1598 models (e.g., 9900I, T2000, T300, CHN600, etc.)
GVW CODE	5 values (4,5,6,7,8)
BODY STYLE NAME	18 values (e.g., TRACTOR TRUCK, STRAIGHT TRUCK, GLIDERS, CAB CHASIS, etc.)
CAB CONFG DESC	17 values (e.g., LONG CONVENTIONAL, LOW TILT, SPECIAL, CAB FORWARD, etc.)
MOVES FUEL CODE	11 values (D, G, N, P, U, X, C, I, H, E, M)
MOVES FUEL DESC	11 values (DIESEL, GAS, CNG, PROPANE, UNKNOWN, GAS, CONVERTIBL, DIESEL HYB, GAS HYBRID, ELECTRIC, METHANOL)
CYLINDER	8 values (0,2,4,5,6,8,10,12)
ENG DSPLCMNT CI	189 values (e.g.,110, 1098, etc.)
VOCATION	34 values (e.g., BE, MI, MO, MS, etc.)
DRIVE WHEEL	18 values (e.g., 6X4, 4X2, blank, 6X6, 6X2, 8X4, etc.)
MOVES VEHICLE TYPE CODE	8 source types (41,42,43,51,52,54,61,62) *Note that there aren't any 53s (Single Unit Long-haul Trucks).
MOVES VEHICLE TYPE DESCRIPTION	8 source type names corresponding to above (i.e., Intercity Bus, etc.)
IHS FUEL TYPE	11 values (DIESEL, GAS, COMPRESSED NATURAL GAS, PROPANE, FLEXIBLE, CONVERTIBLE, ELECTRIC AND DIESEL HYBRID, ELECTRIC AND GAS HYBRID, ELECTRIC, METHANOL)
ANTIQUE	2 values (Y, N)
VIO	Population for each of the 4.4 million combinations

3.2 Four Scenarios

To reflect uncertainty in the number of older LD vehicles operating in the U.S. in 2017, this study provides four scenarios of output files. The only difference among them are the LD populations, which are represented by the three MOVES source type ID codes 21 (Passenger Car), 31 (Passenger Truck), and 32 (Light Commercial Truck). Populations of motorcycles, buses, and heavy trucks do not vary. The four scenarios are as follows:

1. No adjustments
2. Remove all LD antique vehicles
3. Apply the LD 2014 NJ-based reductions nationally
4. Apply the LD 2014 NJ-based reductions nationally and remove all LD antique vehicles

Scenario 1 reflects the direct use of the IHS Markit data without any reduction in population. Scenario 2 removes the relatively small number of LD vehicles that self-report to their state DMV as 'antique,' while Scenario 3 applies larger reductions to LD populations by model year shown in Table 3-4. The NJ-based adjustments are largest for older vehicles. For example, the adjustments remove 72.4% of the 1961-1985 passenger cars reported by IHS Markit. The reductions become progressively smaller with newer model years and there is zero adjustment to car populations in model years 2006-2018. Scenario 4 reflects all the reductions of Scenario 3 plus the removal of any remaining antique vehicles.

Table 3-4. The New Jersey 2014-based Adjustments to Population[2]

Model Year	Passenger Car Fraction of Removed Vehicles	Passenger Truck, Light Commercial Truck Fraction of Removed Vehicles
1985 & earlier	0.724	0.393
1986	0.662	0.419
1987	0.597	0.373
1988	0.585	0.391
1989	0.507	0.386
1990	0.418	0.394
1991	0.359	0.357
1992	0.310	0.353
1993	0.260	0.334
1994	0.215	0.333
1995	0.189	0.292
1996	0.159	0.254
1997	0.131	0.242
1998	0.104	0.207
1999	0.096	0.189
2000	0.077	0.173
2001	0.059	0.160
2002	0.037	0.124
2003	0.036	0.105
2004	0.019	0.090
2005	0.019	0.059
2006	0	0.057
2007	0	0.024
2008 & later	0	0

[2] Ibid.

By comparison to the New Jersey based reductions (Table 3-4 above), the effect of Scenario 2's antique removal is smaller at a national level (Table 3-5). Although the table shows the national total impact by model year, the prevalence of 'antique' LDs varies substantially by state, likely due to local incentives for vehicle owners to apply for antique status. The reason to remove these vehicles is that many states have low mileage limitations or a requirement that these vehicles are for display only at car shows. Antique vehicles make up a small portion of the population at just 0.8% nationally, but are more prevalent in some states (e.g., Mississippi) than others (e.g., California).

Table 3-5. The National Impact of Removing Antique Vehicles

Model Year	Passenger Car Antique Fraction	Passenger Truck Antique Fraction	Light Commercial Truck Antique Fraction
1985 & earlier	0.332	0.217	0.028
1986	0.095	0.070	0.009
1987	0.088	0.064	0.008
1988	0.069	0.049	0.007
1989	0.053	0.037	0.006
1990	0.040	0.033	0.005
1991	0.032	0.024	0.004
1992	0.021	0.018	0.003
1993	0.014	0.012	0.002
1994	0.007	0.005	0.001
1995	0.005	0.004	0.001
1996	0.003	0.002	0.001
1997	0.001	0.001	0.001
1998	0.001	0.001	0.000
1999	0.000	0.000	0.000
2000	0.000	0.000	0.000
2001	0.000	0.000	0.000
2002	0.000	0.000	0.000
2003	0.000	0.000	0.000
2004	0.000	0.000	0.000
2005	0.000	0.000	0.000
2006	0.000	0.000	0.000
2007	0.000	0.000	0.000
2008 & later	0.000	0.000	0.000

The next three figures show the national population by vehicle model year, the number of antique vehicles removed by Scenario 2, and the number of vehicles removed by Scenario 3, the New Jersey based adjustment. Figure 3-1 (Passenger Cars), Figure 3-2 (Passenger Trucks), and Figure 3-3 (Light Commercial Trucks) all show that the most visible impact of vehicle

removal in the scenarios is to reduce the "tail" of the age distribution. The impact reduces the total populations in the MOVES table 'Source Type Year' as well as the influence shape of the fleet age curve in the MOVES table 'Source Type Age Distribution'.

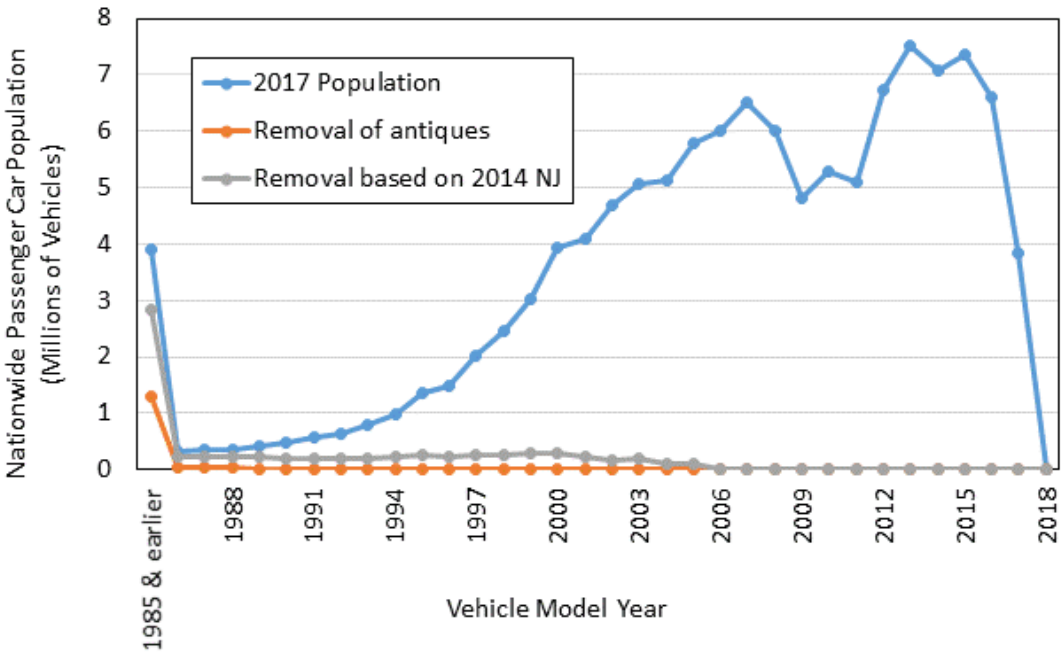

Figure 3-1. Passenger Car Population and Vehicles Removed by Scenarios 2 and 3

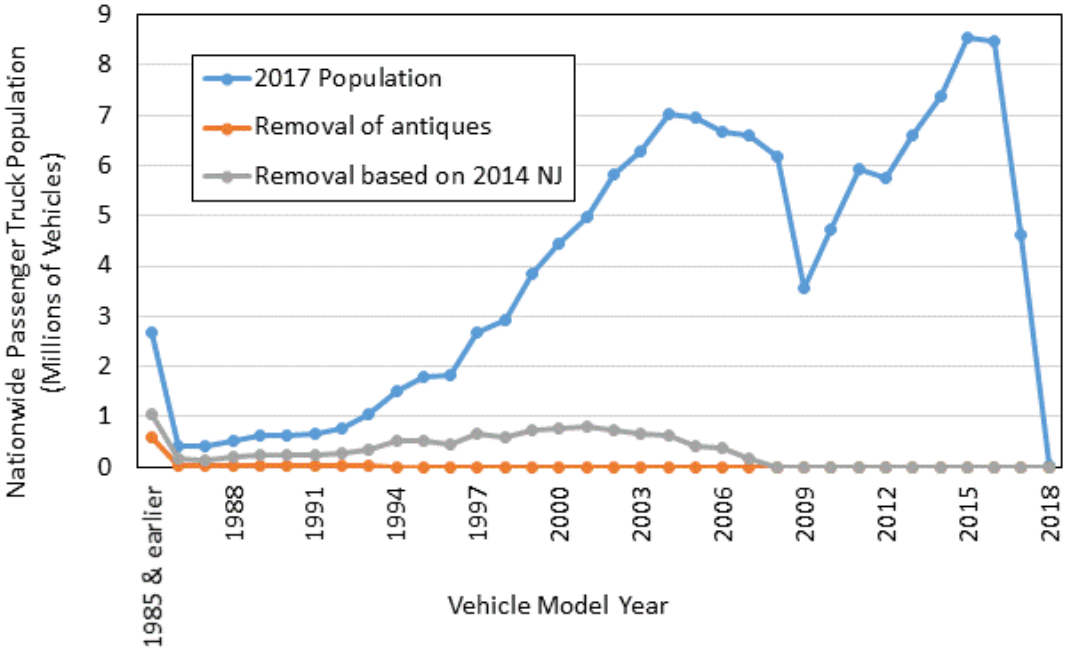

Figure 3-2. Passenger Truck Population and Vehicles Removed by Scenarios 2 and 3

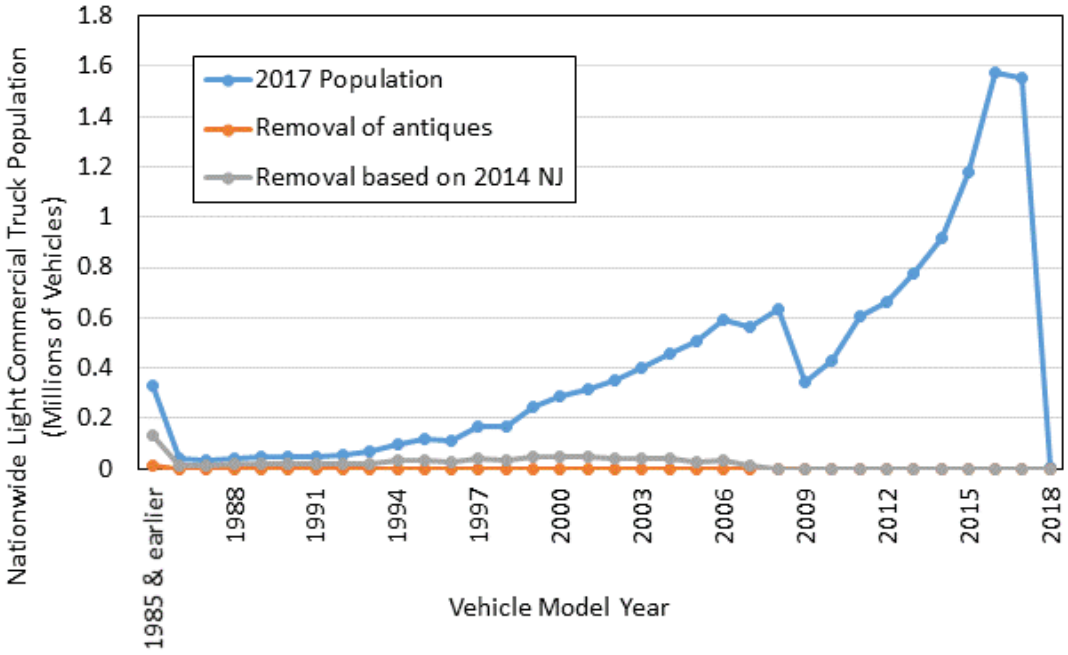

Figure 3-3. Light Commercial Truck Population and Vehicles Removed by Scenarios 2 and 3

3.3 Calculations and Final Data Products

ERG created a program using Statistical Analysis Software [SAS] and MySQL to read in the source data, clean it up, incorporate several assumptions from past NEIs, and calculate the three MOVES output files: population (`SourceTypeYear` table), age distribution (`SourceTypeAgeDistribution` table), and fuel type fractions by model year (`AVFT` table) as well as the summary file of populations by GVWR. This section describes the data processing steps as well as the format and content of the final data products.

The first clean-up step altered the car/light truck designation of several makes and models. Based on EPA emissions certification lists, nearly 2.5 million vehicles listed in Table 3-6 were changed from "Light Truck" (source types 31 and 32) to "Car" (source type 21). Only one (1) vehicle was found to need recategorization in the other direction, from "Car" to "Light Truck" – a single Chevrolet El Camino. There were more than 100,000 other Chevrolet El Caminos, all already categorized as "Light Truck."

Table 3-6. Population Changed from "Light Truck" to "Car" Vehicle Type

Make	Model	Population of "Light Truck" changed to "Car"
ACURA	ZDX	4,489
BUICK	ENCORE	257,792
CHRYSLER	PT CRUISER	690,228
HONDA	ELEMENT	290,755
KIA	SOUL	869,614
NISSAN	JUKE	199,740
SATURN	OUTLOOK	70,084
TOYOTA	SCION XB	76,775
Total		***2,459,477***

The second clean-up step relabeled over 3.4 million vehicles originally designated source types 31 and 32 (Passenger Truck and Light Commercial Truck) to heavy-duty source types ID 52/53 (Single Unit Trucks) because of their Class 3 weight rating (defined as 10,000 to 14,000 lbs. GVW). Class 3 trucks are considered single unit trucks by DOT and therefore are modeled exclusively as heavy-duty source types in the MOVES model (EPA, 2016).

A third clean-up step grouped long-haul trucks together with short-haul because these are indistinguishable from parameters provided in the HD source data (Table 3-3). The source data did not attempt to distinguish single unit trucks that were long-haul from short-haul – there were no source type 53 vehicles. However, the HD source data does label some combination unit trucks as long-haul (source type 62) separately from the short-haul (source type 61). The combination unit trucks were grouped together as 61/62 temporarily. After these three steps,

the program sums population using Equation 1 into categories of state and county of registration, the corrected MOVES source type, GVWR, vehicle model year, and the uncorrected fuel type, and it writes the first output file of four – the GVWR summary. Table 3-7 lists the format this output file and provides additional details on the uncorrected fuel type codes. They are "uncorrected" in that there are fuels and combinations of fuel & vehicle in the summary for which that MOVES does not have emission rates (e.g., cars operating on CNG).

$$population = \sum_{SC,S,G,M,F} VIO \qquad \text{Eqn 1.}$$

Where: ***population*** is the number of vehicles for each combination of county *SC*, source type *S*, gross vehicle weight rating *G*, model year *M*, and fuel type *F*
VIO is the number of vehicles in operation in 2017
SC is the U.S. state and county
G is the gross vehicle weight rating (GVWR) category code
S is the MOVES source type
M is the vehicle model year
F is the uncorrected MOVES fuel type ID

Table 3-7. Output File GVWR Summary – Data Fields and Description

Field Name	Description of the Data
STATE_NAME	Name of state
COUNTY_NAME	Name of county
sourceTypeID	MOVES vehicle class 11, 21, 31, 32, 41, 42, 43, 51, 52/53, 54, 61/62
gvw_code	The GVWR code from IHS, values include: blank, 1, 1a, 2, 2a, 2b, 3, 3a, 3b, 4, 5, 6, 7, and 8.
YEAR_MODEL	Vehicle model Year
fuelTypeID	Uncorrected MOVES fuel type codes assigned to IHS fuel descriptions: 0 Convertible, Hydrogen Fuel Cell, Methanol, Unknown, blank 1 Gasoline, Electric and Gas Hybrid 2 Diesel, Electric and Diesel Hybrid 3 Compressed Natural Gas 4 Propane 5 Flexible Fuel, Ethanol 9 Electric
population	The number of vehicles operating

Next, the program applies fractions to split the combined source type groups 52/53 and 61/62 into short-haul and long-haul activity vehicles (Table 3-8). The fractions were originally developed for the CRC A-88 project, separately by MOVES road type for the purpose of dividing VMT into the activity from short-haul vs. long-haul trucks. U.S. EPA VMT-weighted the CRC A-

88 fractions over road type to develop fractions suitable as a surrogate for splitting truck populations into short- and long-haul source types in the 2014 NEI, resulting in the Table 3-8 fractions. The factors vary geographically by U.S. Census Regions which map to states, one-to-many, according to Figure 3-4.

Table 3-8. Fractions to Split Heavy-Duty Truck Populations into Short- and Long-Haul

Truck Type	Census Region	Fraction Short-haul	Fraction Long-haul
Single Unit 52/53	Midwest	0.807	0.193
	Northeast	0.919	0.081
	South	0.860	0.140
	West	0.882	0.118
Combination Unit 61/62	Midwest	0.442	0.558
	Northeast	0.448	0.552
	South	0.535	0.465
	West	0.468	0.532

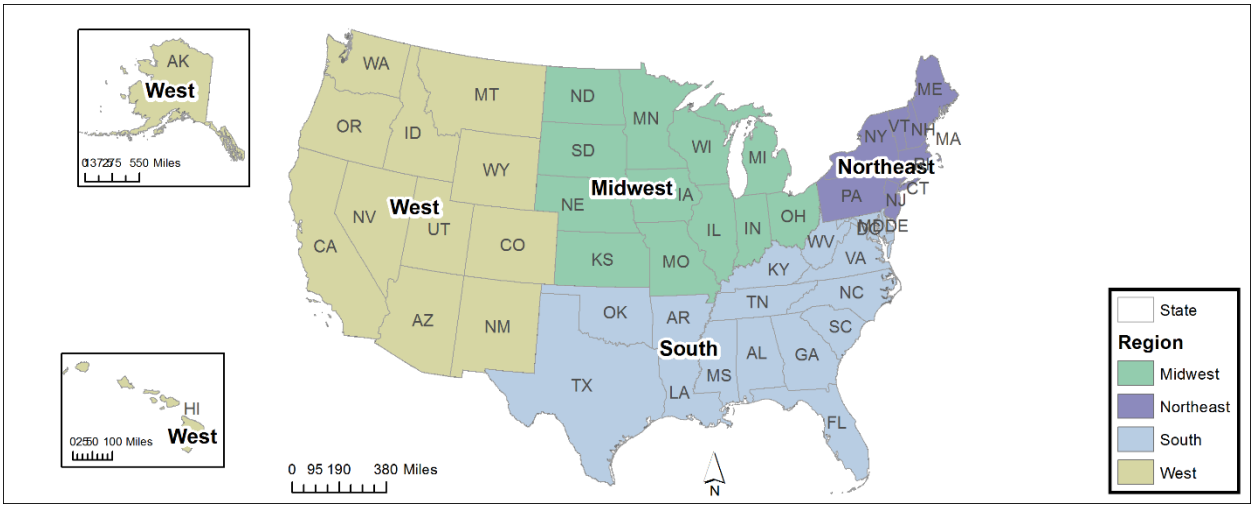

Figure 3-4. Map of the U.S. Census Regions[3]

After dividing trucks into short- and long-haul, the program sums population by county ID and source type ID using Equation 2 below and outputs the second of four output files, the MOVES-formatted population input table 'SourceTypeYear'. Table 3-9 lists the data fields and content.

[3] Accessed January 2019. https://www.census.gov/geo/reference/webatlas/regions.html

$$sourceTypePopulation = \sum_{C,S} VIO$$ Eqn 2.

Where: ***sourceTypePopulation*** is the MOVES-ready county totaled for the county *C* and source type *S*
VIO is the number of vehicles in operation in 2017
C is the U.S. state and county
S is the MOVES source type

Table 3-9. Output File MOVES 'SourceTypeYear' Table – Data Fields and Description

Field Name	Description of the Data
countyID	All state and county FIPS codes
yearID	value=2017
sourceTypeID	MOVES vehicle class 11, 21, 31, 32, 41, 42, 43, 51, 52, 53, 54, 61, 62
sourceTypePopulation	The number of vehicles operating

Next, the program determines the allowable combinations of vehicle class and fuel type prior to generating the last two output file types, the MOVES tables 'SourceTypeAgeDistribution' and 'AVFT.' First, a regulatory class ID code is assigned following the definitions in Table 3-10. All source type ID 21s (Passenger Cars) were assigned the regulatory class 20. Source type 31s and 32s were assigned to regulatory class 40 if their GVWR code was Class 2b, and otherwise were assigned class 30. The remaining HD trucks were categorized into regulatory classes solely based on GVWR codes in Table 3-10; all class 3 trucks are regulatory class 41, class 4 and 5 trucks are regulatory class 42, class 6 and 7 are regulatory class 46, and class 8 trucks are regulatory class 47. The combination of regulatory assignment and source type ID of each vehicle class and allowable fuel types in MOVES are shown in Table 3-11. Any combination in the 2017 data that was outside the allowable fuel types was either deleted or assigned a new fuel type ID prior to producing the 'AVFT' table. The program performs the following actions to correct fuels for AVFT:

- Delete all electric (fuel type 9) heavy-duty vehicles
- Rename all CNG (fuel type 3) to Gasoline (fuel type 1) for source types that are not transit bus (source type 42)
- Rename all LPG (fuel type 4) to Gasoline (fuel type 1)
- Rename E-85 (fuel type 5) to Gasoline (fuel type 1) for heavy-duty vehicles

Table 3-10. Regulatory Class Definitions in MOVES2014

regClassID	Regulatory Class Name	Description
0	Doesn't Matter	Doesn't Matter
10	MC	Motorcycles
20	LDV	Light-Duty Vehicles
30	LDT	Light-Duty Trucks
40	LHD<=10k	Class 2b Trucks with 2 Axles and 4 Tires (8,500 lbs < GVWR <= 10,000 lbs)
41	LHD<=14k	Class 2b Trucks with 2 Axles and at least 6 Tires or Class 3 Trucks (8,500 lbs < GVWR <= 14,000 lbs)
42	LHD45	Class 4 and 5 Trucks (14,00 lbs < GVWR <= 19,500 lbs)
46	MHD	Class 6 and 7 Trucks (19,500 lbs < GVWR < =33,000 lbs)
47	HHD	Class 8a and 8b Trucks (GVWR > 33,000 lbs)
48	Urban Bus	Urban Bus (see CFR Sec. 86.091_2)

Table 3-11. Combinations of Source Type, Fuel, and Regulatory Class in MOVES2014

sourceTypeID	fuelTypeID	regClassID
11	1	10
21	1, 2, 5, 9	20
31	1, 2, 5, 9	30, 40
32	1, 2, 5, 9	30, 40
41	2	41, 42, 46, 47
42	1	42, 46, 47
	2, 3	48
43	1, 2	41, 42, 46, 47
51	1, 2	41, 42, 46, 47
52	1, 2	41, 42, 46, 47
53	1, 2	41, 42, 46, 47
54	1, 2	41, 42, 46, 47
61	1, 2	46, 47
62	2	46, 47

Table 3-12. MOVES Fuel Type ID Codes

MOVES Fuel Type ID	Description
1	Gasoline
2	Diesel Fuel
3	Compressed Natural Gas (CNG)
5	Ethanol (E-85)
9	Electricity

Next, the program calculates the age distributions using Equation 3 and outputs the third of four output file types, the MOVES table `SourceTypeAgeDistribution.' MOVES only uses 31 age ID codes numbered 0 to 30; therefore, for calendar year 2017 purposes the vehicle model years 2017 and 2018 are age 0, and the 1987 and older models are all age 30. Table 3-13 lists the data fields and content of the age distribution deliverables. There is an exception to this approach for 53 (Single Unit Long-haul Truck) and 62 (Combination Long-haul Truck). The final age distributions of long-haul trucks reflect the national data rather than individual counties because those vehicle types operate over long distances.

$$ageFraction = \frac{\sum_{C,S,A} VIO}{\sum_{C,S} VIO} \qquad \text{Eqn 3.}$$

Where: ***ageFraction*** is the fraction of the population by age for source type *S* in county *C*
VIO is the number of vehicles in operation as of July 1, 2017
C is the U.S. state and county
S is the MOVES source type
A is the MOVES age bin 0 through 30

Table 3-13. Output File MOVES 'SourceTypeAgeDistribution' Table – Data Fields and Description

Field Name	Description of the Data
countyID	All state and county FIPS codes
sourceTypeID	MOVES vehicle class 11, 21, 31, 32, 41, 42, 43, 51, 52, 53, 54, 61, 62
yearID	value=2017
ageID	MOVES vehicle age bin 0 through 30
ageFraction	fraction of the population in age bin for each vehicle type and county

After writing the age distribution output file, the program deletes the unknown fuel types prior to preparing the MOVES AVFT table. These include "CONVERTIBLE," "HYDROGEN FUEL CELL," "METHANOL," "UNKNOWN," and blanks. The MOVES `AVFT` table has the fields shown in Table 3-14 except for the leading countyID field. The important data field in this table is named fuelEngFraction, and it is calculated by Equation 4. The values of this fraction sum to one (1) over each fuelTypeID, for every combination of county, source type, and model year. The exception to this approach is for source types 41 (Intercity Bus), 53 (Single Unit Long-haul Truck), and 62 (Combination Long-haul Truck). MOVES has only diesel emission rates for source type 41, so the fractions associated with this source type are all value of 1 and only listed for the diesel fuel type. The long-haul trucks, 53 and 62, do not reflect the county data, but instead the national total data because these trucks operate over long distances.

$$fuelEngFraction = \frac{\sum_{C,S,M,F} VIO}{\sum_{C,S,M} VIO}$$ Eqn 4.

Where: ***fuelEngFraction*** is the fraction of the population with each fuel type *F* for each combination of source type *S* with model year *M* registered in county *C*
VIO is the number of vehicles in operation as of July 1, 2017
C is the U.S. state and county
S is the MOVES source type
M is the vehicle model year
F is the MOVES fuel type ID

Table 3-14. Output File MOVES 'AVFT' Table – Data Fields and Description

Field Name	Description of the Data
countyID	All state and county FIPS codes
sourceTypeID	MOVES vehicle class 11, 21, 31, 32, 41, 42, 43, 51, 52, 53, 54, 61, 62
modelYearID	Vehicle model year
fuelTypeID	Corrected MOVES allowable fuel type codes 1, 2, 3, 5, 9 for gasoline, diesel, CNG, E-85 capable, and electric, respectively.
engTechID	MOVES engine technology code 1=Conventional Internal Combustion or 30=Electric
fuelEngFraction	fraction of the population in each fuel type, by source type, model year, and county

2016 Deliverables

In addition to the 2017 files described above, ERG prepared 2016 estimates of population (MOVES 'SourceTypeYear' table) and age distribution (MOVES 'SourceTypeAgeDistribution' table) based primarily on the 2017 data.

The 2016 population data were prepared by a straightforward interpolation between the 2014 and 2017 populations by source type and county. For age distribution, the 2016 data were estimated from the 2017 age distributions, primarily by shifting the age IDs by one year. For example, the age ID 2 in year 2017 is a model year 2015 and that is age 1 in year 2016. In this way, ages 1 through 28 were shifted by one year. The head (age 0) and tail (ages 29 and 30) of the distribution received different treatment. The age 0 in 2016 needed to reflect a mid-year draw date, so the age 0 fraction was calculated using the ratio of population in age 0 to age 1 in the 2017 data to calculate a new fraction for age 0 in 2016. A similar ratio approach from age 28 was taken for the tail, so that age 29 did not reflect a large tail in 2016, and age 30 did reflect a tail in 2016 similar in magnitude to the one in 2017. The 2016 age distributions were then normalized so that each summed to one (1) for all counties and source types.

Listing of Deliverables

This work provides 24 files inside an archive with a "zip" extension. The files are organized under two directory levels, first the year, then scenario, as outlined below.

- 2016
  Each scenario subdirectory contains a `SourceTypeAgeDistribution` and `SourceTypeYear` file, for 8 files total corresponding to calendar year 2016.
  - Scenario1_Unadjusted
  - Scenario2_OnlyRemoveAntiques
  - Scenario3_ApplyNJAdjs_and_KeepAntiques
  - Scenario4_ApplyNJAdjs_and_RemoveAntiques

- 2017
  Each scenario subdirectory contains four files (`AVFT,` `SourceTypeAgeDistribution,` `SourceTypeYear` , and the extra population summary with GVWR categorization), for a total of 16 files corresponding to 2017.
  - Scenario1_Unadjusted
  - Scenario2_OnlyRemoveAntiques
  - Scenario3_ApplyNJAdjs_and_KeepAntiques
  - Scenario4_ApplyNJAdjs_and_RemoveAntiques

Note that `AVFT` files were not generated separately for 2016. The same AVFT may be used for both 2016 and 2017 MOVES modeling. Unlike the age distribution file, the `AVFT` specifies the individual vehicle model years, so the distributions do not require a shift to account for age ID differences by calendar year.

4.0 Results

The 2017 data show that light duty vehicles ages vary geographically across the country. One way to parametrize the distribution in each county is to calculate an average age. Figure 4-1 shows the average age in each U.S. county for Passenger Cars, Passenger Trucks, and Light Commercial Trucks, color-coded by the legend enlarged in the lower right quadrant. The purple and blue counties represent the youngest fleets, while the reds indicate the oldest. For both passenger cars and passenger trucks, most of California, Texas, Florida, the Upper Midwest, and Northeast have the youngest fleet age, while the South, Central states, and Northwest have the oldest vehicles. This pattern in the passenger car and trucks is stronger than for Light Commercial Trucks, which are also newer overall.

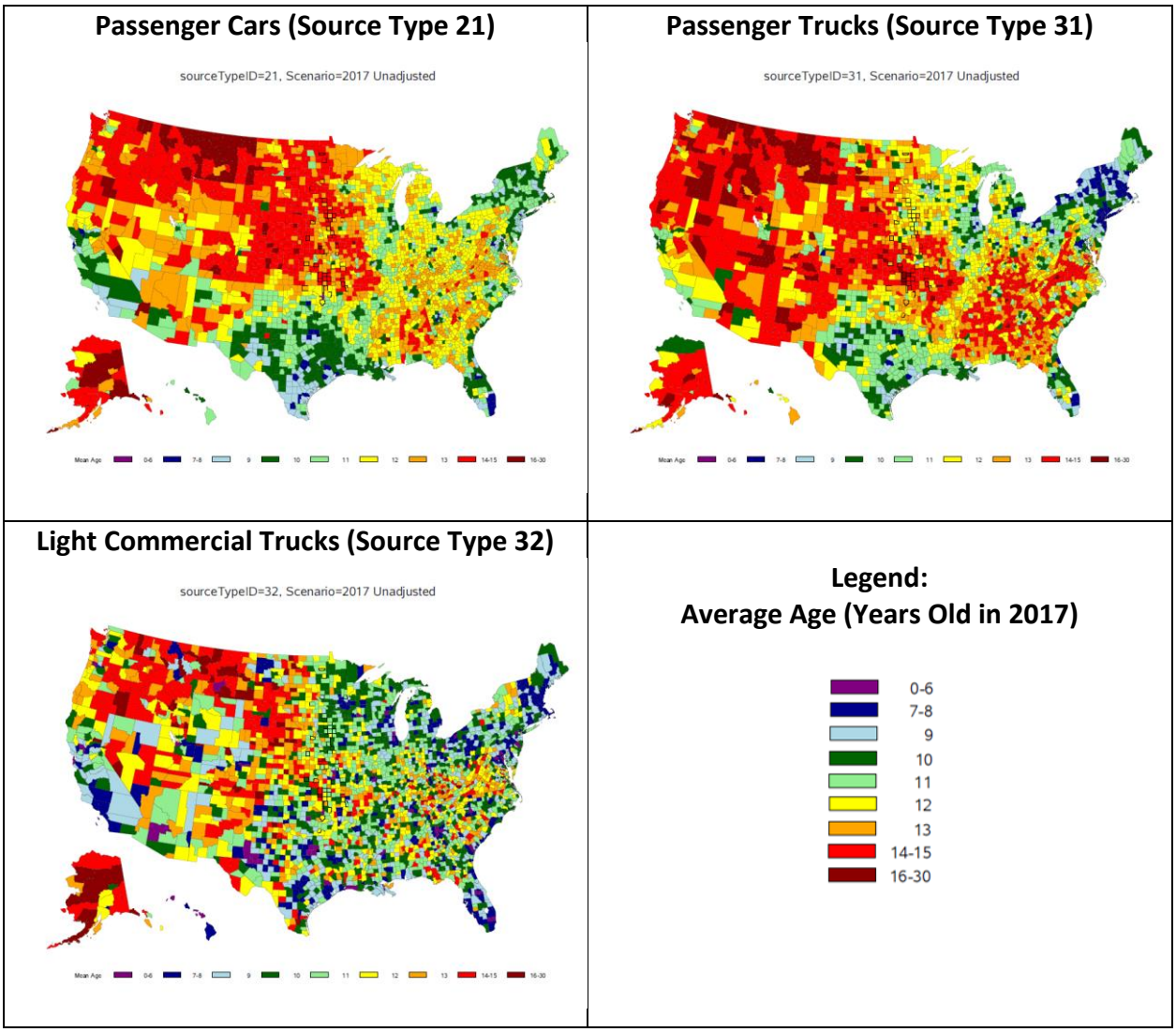

Figure 4-1. Light Duty Vehicle Average Age in 2017

4.1 Changes from 2014 to 2017

Light duty vehicles in 2017 have an older average age than in 2014. Figure 4-2 shows the percent change in average age of LD vehicles over the 3-year period. The red hues indicate an increase in age (older vehicles in 2017) and blues indicate a decrease (younger vehicles in 2017).

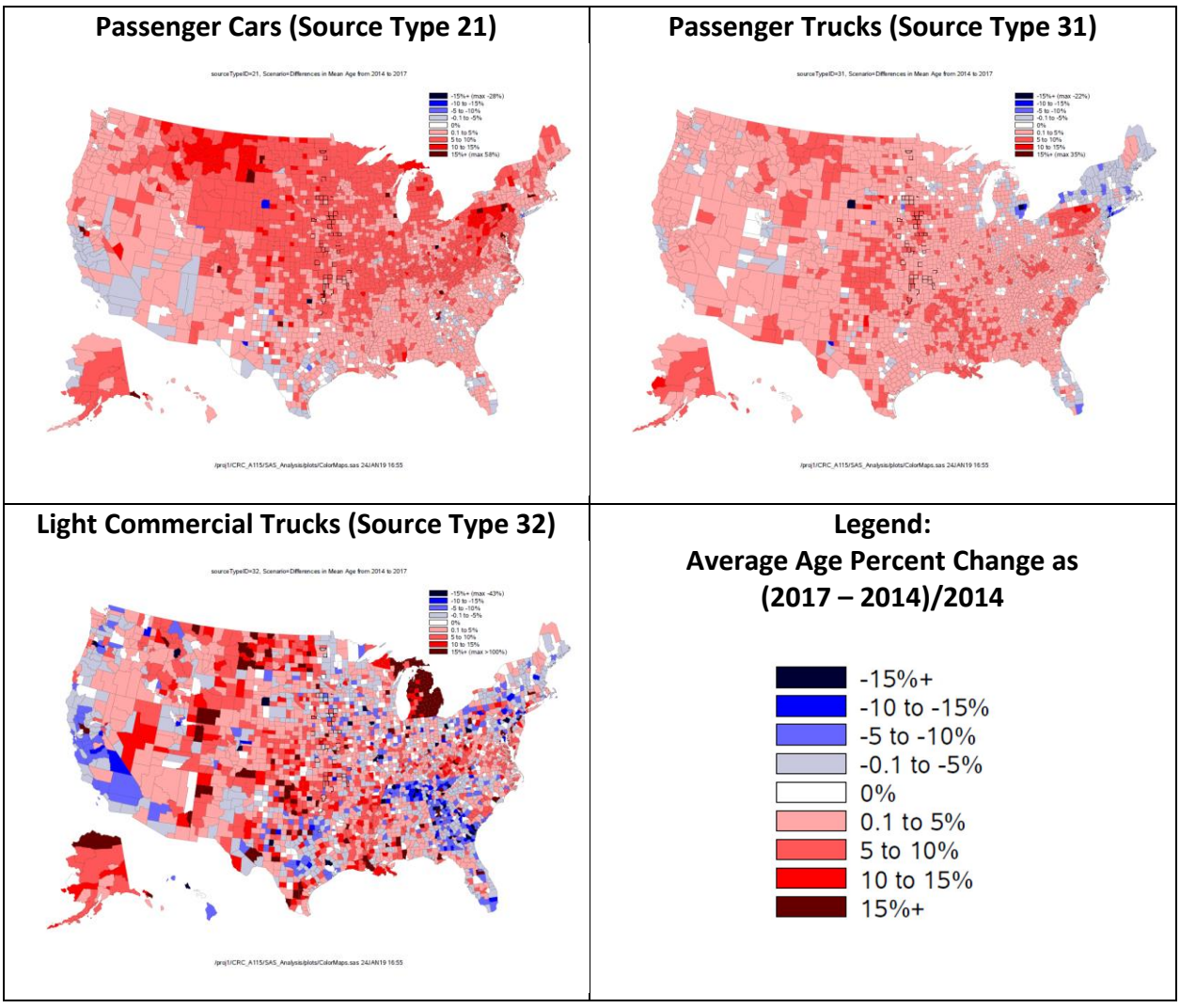

Figure 4-2. Differences in Average Age from 2014 to 2017

The national total light duty population grew from 2014 by 6%, driven largely by Passenger Truck population increase, which grew 12%. By contrast, passenger car populations grew by just 0.4% over three years. Figure 4-3 shows which counties had increases (red) in population and decreases (blue). The top 5 light duty truck make and models driving the truck growth were new model years (2014-2018) of Chevrolet Silverado, Honda CR-V, Ford Escape, Toyota Rav4 and RAM 1500. Passenger car populations decreased in Alaska, Maine, and across the central US. Tennessee and Michigan show the largest percent decreases in population of light

commercial trucks. Because these patterns are so clearly defined along two states boundaries, perhaps the change in 2017 is due to a state change in data reporting.

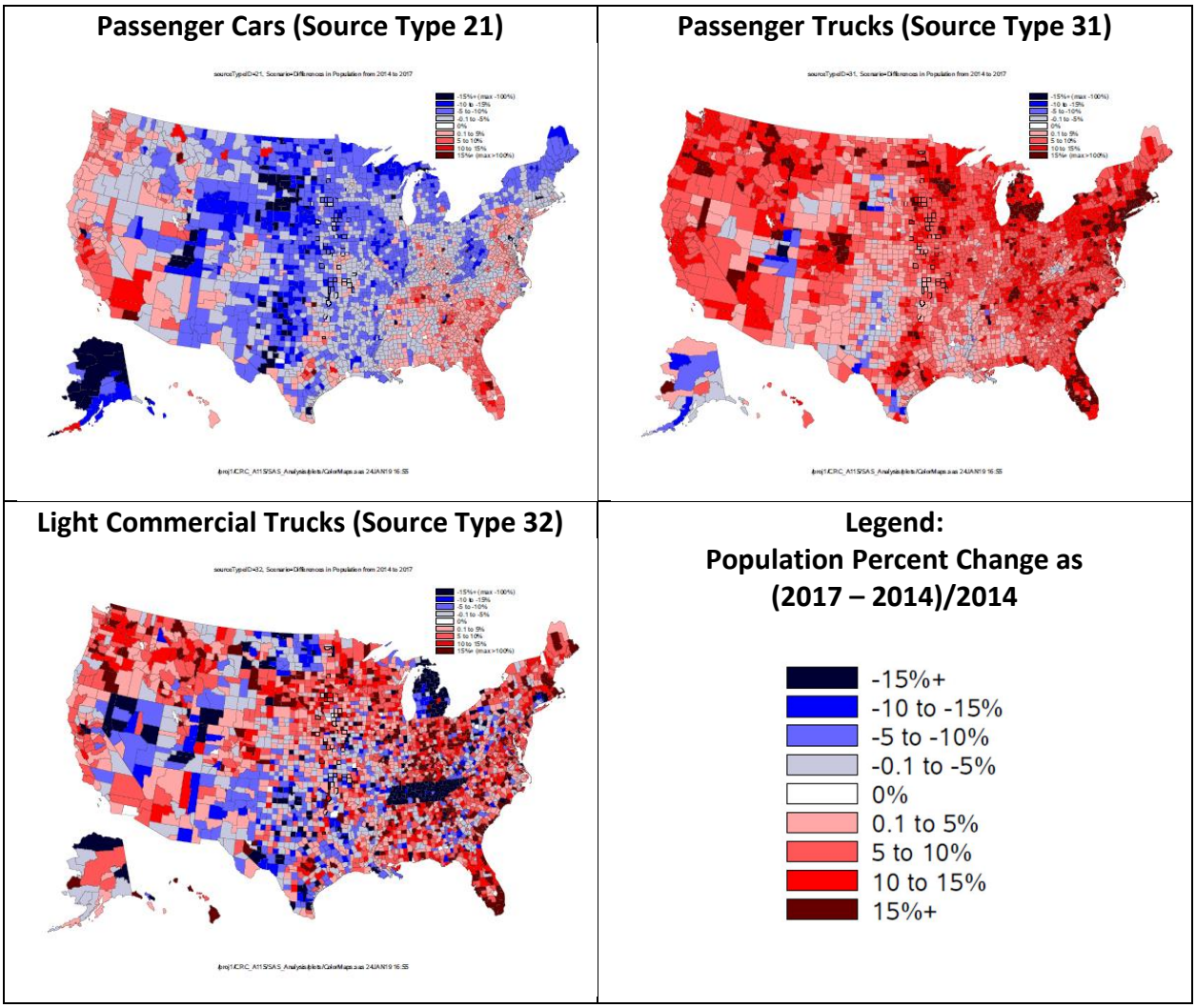

Figure 4-3. Differences in Population from 2014 to 2017

The US maps above all show the 2017 data without any reductions to the populations of older vehicles. The next two sections of results show the decreases in average age and population that occur from applying the New Jersey based adjustments nationally (Section 4.2) and removing antiques (Section 4.3).

4.2 Impact of NJ-based Adjustments

The scenario NJ-based adjustments to reduce vehicle populations causes decreases in both the average age (Figure 4-4) and total population (Figure 4-5) of light-duty vehicles. The average age decreased the most for Passenger Cars (upper left of Figure 4-4) and have the least impact on Light Commercial Trucks, which tend to be younger than Passenger Cars and Passenger Trucks (recall Figure 4-1).

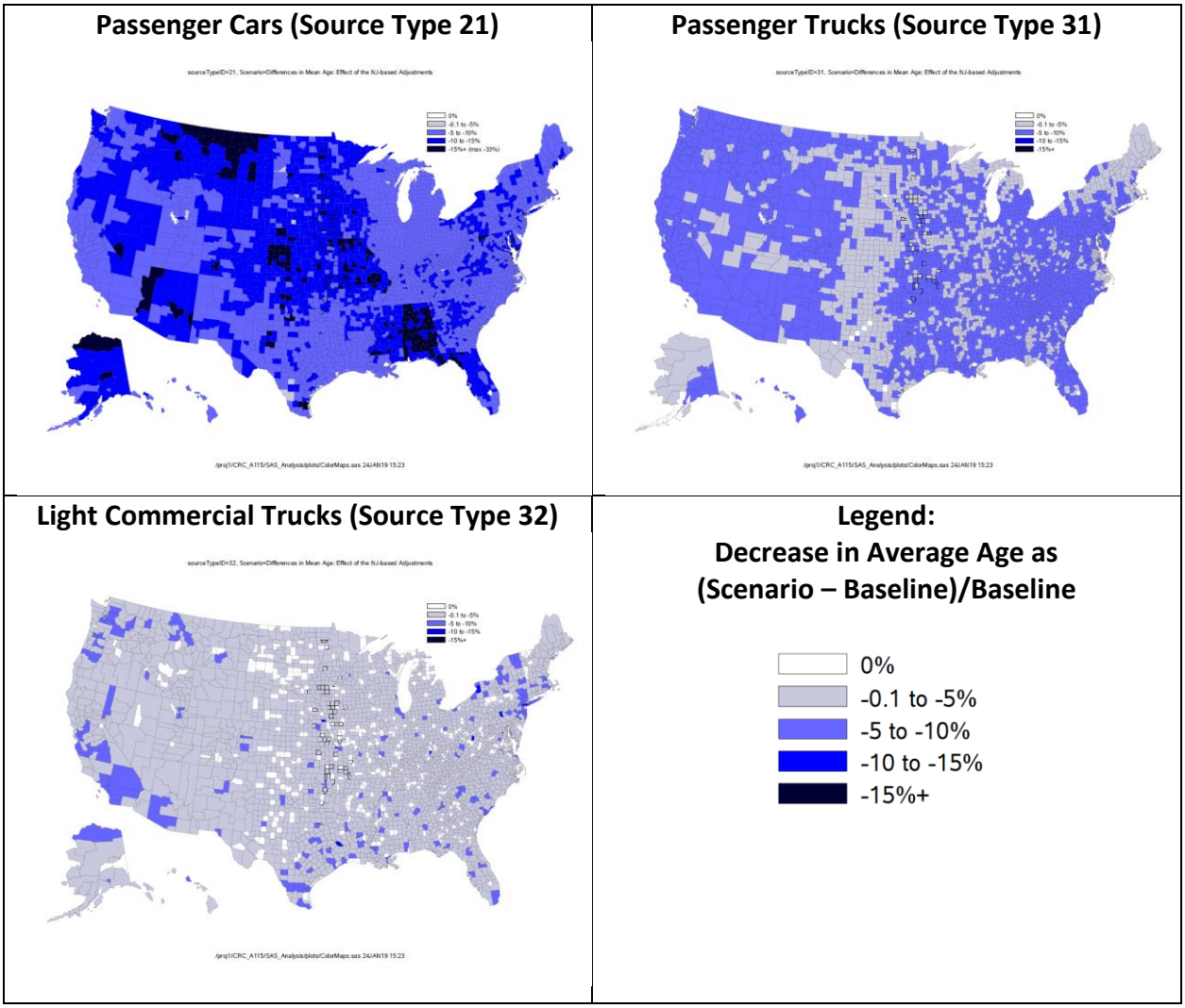

Figure 4-4. Differences in Average Age in 2017: Baseline vs. NJ-Adjustments

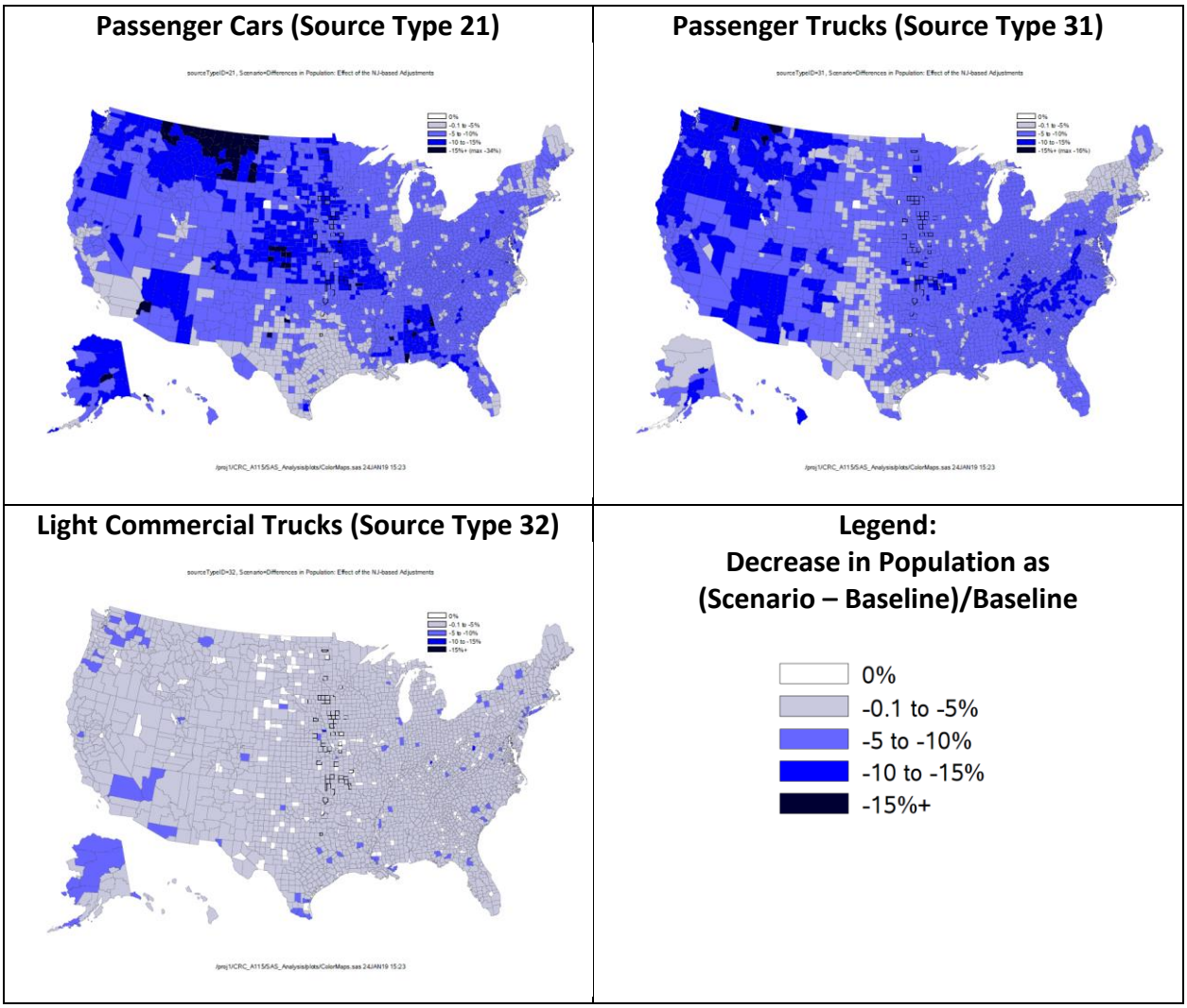

Figure 4-5. Differences in Population in 2017: Baseline vs. NJ-Adjustments

4.3 Impact of Removing Antique Vehicles

The effect of removing antiques from the unadjusted 2017 data is milder than the NJ-based adjustments. Figure 4-6 shows several states with near-zero change in mean age; these are areas that do not have very many vehicles registered as "antique." Mississippi and Maryland have the highest impact, with portions of both states where the average age of passenger cars becomes more than 15% younger and the population totals (Figure 4-7) drop by up to 15%.

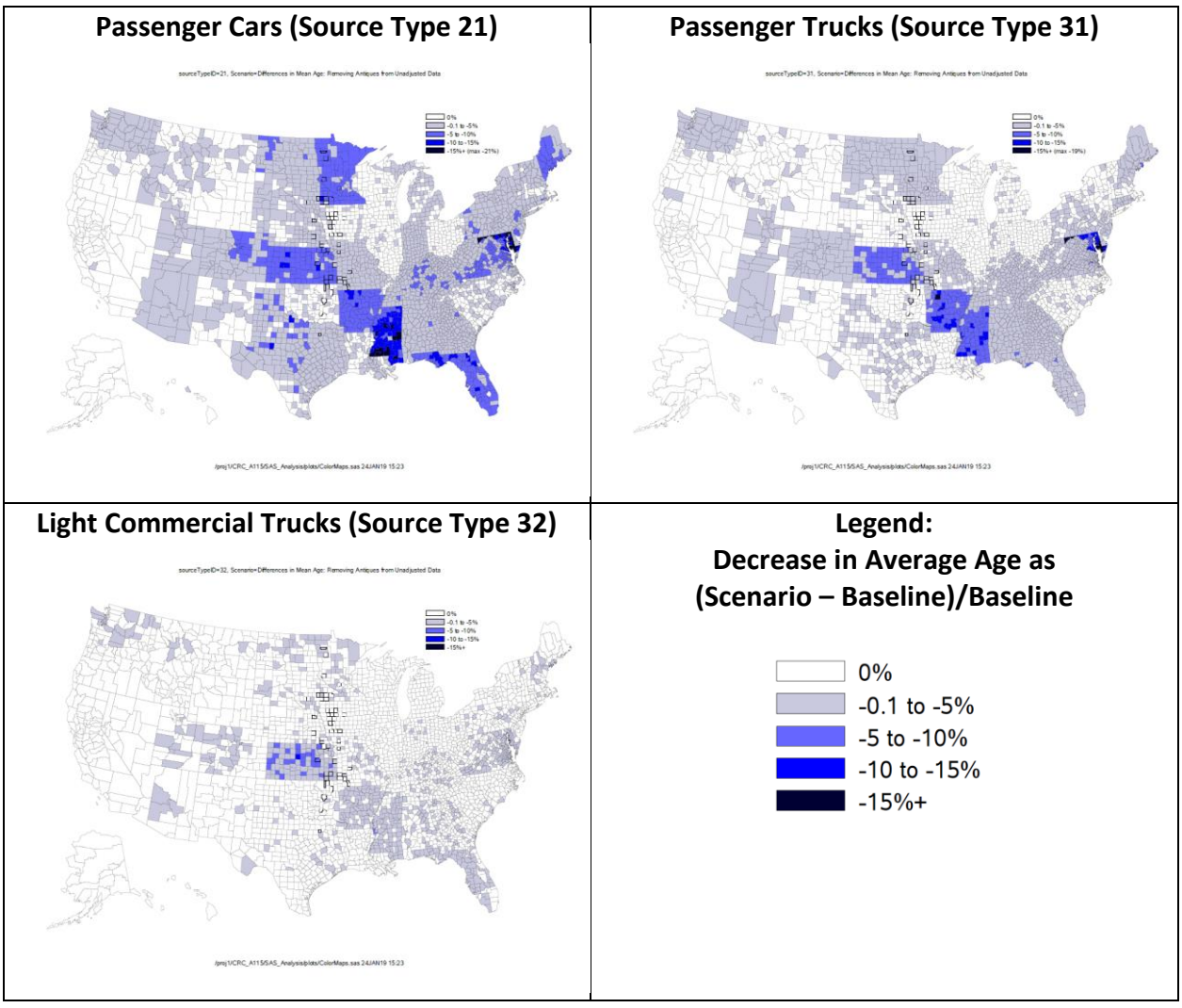

Figure 4-6. Differences in Average Age in 2017: Baseline vs. No-Antiques

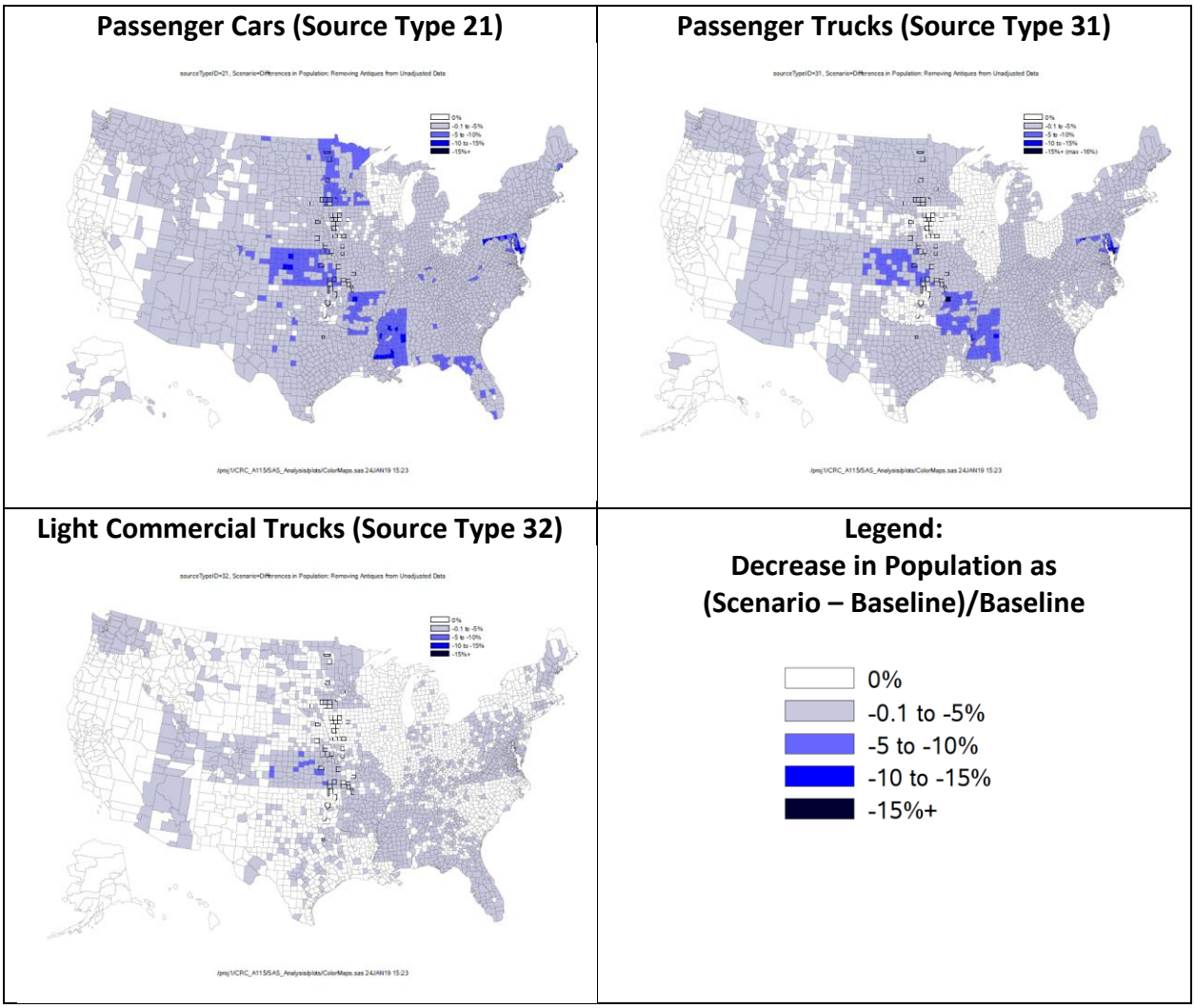

Figure 4-7. Differences in Population in 2017: Baseline vs. No-Antiques

5.0 Summary and Recommendations

CRC sponsored the A-115 project to provide NEI inputs of population and age distribution for the 2017 NEI and 2016 v1 modeling platforms, including several scenarios of adjustments that remove older vehicles as a correction to the data. One limitation of the adjustments based on the NJ data is that the data are limited to one state and the year 2014. Many states are expected to submit 2017 vehicle population estimates and some will also include age distribution inputs; these submittals can be compared to the 2017 IHS Markit data. Repeating the VIN workgroup's analysis for 2017 data leaves open the possibility for extending the analysis to a broader group of states and develop adjustments using more state data and for the current emission inventory year. The results of such a comparison could help evaluate whether the differences observed in the 2014 NJ comparisons by model year hold true for model years in 2017 and whether the profiles of over-estimation by model year have shapes that are consistent by state.

6.0 Acknowledgements

The authors acknowledge the CRC Atmospheric Impacts Committee for devising project objectives, purchasing the data, and project leaders Mark Janssen (LADCO) and Susan Collet (Toyota), who provided technical guidance. Mark Janssen (LADCO), Tom Dvorak (New Jersey DEP), Denise Cormier (Maine), and Stacy Allen (Missouri DNR) also provided the 2014 VIN analysis results. Alison Eyth and Megan Beardsley (U.S. EPA) provided valuable comments and input on this work.

7.0 References

1. Eastern Research Group, Inc., 2013. "Study of MOVES Information for the National Emission Inventory: CRC Project A-84", CRC Final Report, October.

2. Eastern Research Group, Inc., 2014. "MOVES Input Improvements for the 2011 NEI: CRC Project A-88", CRC Final Report, October.

3. U.S. EPA, 2016. "Population and Activity of On-road Vehicles in MOVES2014." Assessment and Standards Division, Office of Transportation and Air Quality. EPA-420-R-16-003, p.10. January.

4. U.S. EPA, 2015. "Technical Support Document, Preparation of Emissions Inventories for the Version 6.2, 2011 Emissions Modeling Platform," Office of Air Quality Planning and Standards, August.