



3508 Far West Blvd., Suite 210
Austin, TX 78731

MEMORANDUM

TO: David Brzezinski, EPA

FROM: Meredith Weatherby, Doug Jackson, John Koupal, Allison DenBleyker, ERG

DATE: July 21, 2017

SUBJECT: Analysis of IHS Registration Data and Preparation of WA 5-08 Task 1 Deliverables

This memo summarizes ERG's analysis of the IHS Markit (IHS) registration database, quality control (QC) checks of the incoming source type ID assignments, and preparation of the deliverables under WA 5-08 Task 1. The deliverables for this work assignment are the following two tables, formatted for the MOVES county database (CDB) with an additional leading field to identify the county FIPs code. ERG processed these data tables two ways, with and without grouping over the 2014 v2 NEI representative county areas.

- **Source Type Age Distribution** – disaggregates source type population to 31 model years
- **Alternative Fuel and Vehicle Technology (AVFT)** – disaggregates source type and model year population into engine types of gasoline, diesel, CNG, flex-fuel, and electric.

The IHS Registration Database

ERG purchased a mid-year 2014 vehicle registration database from IHS. The registration database is a county-level summary of registered vehicles that IHS retrieves from each state DMV. IHS compiled the data in-house, decoded vehicle identification numbers (VINs), and assigned each record a MOVES source type identification (ID) code. IHS provided many descriptive fields in their delivery to ERG, but did not provide all the information they used in making their initial MOVES source type ID assignments. The database doesn't provide the VINs or identify individual vehicles, but rather provides a summary count of the populations in each county by parameters including make, model, model year, gross vehicle weight (GVW) class, and other fields. For example, a hypothetical row in the IHS database might indicate there are 32 vehicles that were a 2009 Toyota Prius passenger car in Wayne County, Michigan. In total, there were over 44 million records provided that identified 277 million vehicles registered in the US. Tables 1 and 2 list all fields of the light-duty (LD) and heavy-duty (HD) datasets that IHS delivered to ERG. The delivered registration data were separated LD and HD, and while fields were the same in the two, they are presented below in separate tables to give a sample of the unique values in each set.

Table 1. 2014 IHS Light Duty (LD) Data Summary

Filename: MOVES_EPA_CAR_LT_MCY.csv
File size: 5.2 Gb uncompressed
N records: 39.5 Million (39,564,751 rows of data)
N Population: 268 Million

IHS Data Field Name	Unique Values
STATE NAME	51 (includes District of Columbia)
COUNTY NAME	3,143 counties
VEH TYPE	3 unique values: "CAR", "LIGHT TRUCK", or "MOTORCYCLE"
OWNER TYPE	"COMMERCIAL" or "PERSONAL"
YEAR MODEL	55 MDYs (1961 – 2015)
MAKE NAME	318 unique makes (e.g., PONTIAC, JEEP, HARLEY-DAVIDSON, TOYOTA, CHRYSLER, etc.)
MODEL NAME	2,802 unique models (e.g., SUNBIRD, WAGONEER, FXDWG, CELICA, TOWN & COUNTRY, etc.)
GVW CDE	6 unique values (null, 1, 2, 2a, 2b, 3)
BODY STYLE NAME	64 unique values (e.g., CONVERTIBLE, SPORT UTILITY VEHICLE, SCOOTER, SEDAN, etc.)
TRK CAB CONFG DESC	16 unique values (e.g., UNKNOWN, VAN, SPORT UTILITY, EXTENDED CAB, etc.)
MOVES FUEL CODE	10 unique values (e.g., G, D, null, N, E, X, P, E)
MOVES FUEL DESC	10 unique values (e.g., GAS, DIESEL, null, CNG, ELECTRIC, UNKOWN, PROPANE, ETHANOL)
CYLINDER	5 unique values (0,4,6,8,10)
ENG DSPLCMNT CI	6 unique values (0, 98, 121, 330, 403, 415)
VOCATION	All null.
DRIVE WHEEL	3 unique values (null, 4X2, and 4X4)
VEHICLE TYPE CODE	4 unique values (11, 21, 31, 32)
VEHICLE TYPE DESCRIPTION	4 unique values corresponding to above (i.e., MOTORCYCLE, PASSENGER CAR, PASSENGER TRUCK, LIGHT COMMERCIAL TRUCK)
REGISTRATION TYPE	6 unique values (e.g., DEALER/MANUFACTURER, LARGE FLEET, RENTAL, PERSONAL, etc)
IHS FUEL TYPE	10 unique values (e.g., null, COMPRESSED NATURAL GAS, CONVERTIBLE, FLEXIBLE, GAS, DIESEL, etc.)
VIO	Population for each of the 39.5 million combinations

Table 2. 2014 IHS Heavy Duty (HD) Data Summary

Filename: tipnet_epa.csv.csv
File size: 637 Mb uncompressed
N records: 4.6 Million (4,567,436 rows of data)
N Population: 8.6 Million (8,627,130 sum of VIO)

IHS Data Field Name	Unique Values
STATE NAME	51 (includes District of Columbia)
COUNTY NAME	3,139 counties
VEH TYPE	All “HD”
OWNER TYPE	“COMMERCIAL” or “PERSONAL”
YEAR MODEL	55 MDYs (1961 – 2015)
MAKE NAME	88 makes (e.g., INTERNATIONAL, KENWORTH, MACK, PETERBILT, VOLVO, etc.)
MODEL NAME	1598 models (e.g., 9900I, T2000, T300, CHN600, etc.)
GVW CDE	5 unique values (4,5,6,7,8)
BODY STYLE NAME	14 unique values (e.g., TRACTOR TRUCK, STRAIGHT TRUCK, GLIDERS, CAB CHASIS, etc.)
TRK CAB CONFG DESC	17 unique values (e.g., LONG CONVENTIONAL, LOW TILT, SPECIAL, CAB FORWARD, etc.)
MOVES FUEL CODE	7 unique values (e.g., G, D, E, N, X, P, U)
MOVES FUEL DESC	7 unique values (e.g., GAS, DIESEL, ELECTRIC, CNG, UNKOWN, PROPANE, CONVERTIBLE)
CYLINDER	7 unique values (0,4,5,6,8,10,12)
ENG DSPLCMNT CI	171 unique values (e.g., 915, 893, 441, 763, etc.)
VOCATION	33 unique values (e.g., PETROLEUM, GENERAL FREIGHT, CONSTRUCTION, etc.)
DRIVE WHEEL	18 unique values (e.g., 6X4, 4X2, UNK, 6X6, 6X2, 8X4, etc.)
VEHICLE TYPE CODE	8 unique source types (41,42,43,51,52,54,61,62)
VEHICLE TYPE DESCRIPTION	8 source type names corresponding to above (i.e., Intercity Bus, etc.)
REGISTRATION TYPE	5 unique values (e.g., DEALER/MANUFACTURER, FLEET, RENTAL, PERSONAL, etc)
IHS FUEL TYPE	11 unique values (e.g., null, COMPRESSED NATURAL GAS, CONVERTIBLE, FLEXIBLE, GAS, DIESEL, etc.)
VIO	Population for each of the 4.6 million combinations

Updates to Vehicle and Fuel Categories

ERG reviewed the data delivery including IHS's initial assessment of source type ID codes and made five key changes to the data for consistency and compatibility with MOVES. ERG assigned also MOVES regulatory class ID codes to each record. After completing all necessary updates to the vehicle classifications, a SAS program read in the data and calculated the age distribution and AVFT distributions for each county in the US. All changes ERG made to the IHS registration data are listed below.

- 1. Updated several vehicle make/models from source type ID 31 (Light Commercial Truck) to 21 (Passenger Car).** ERG updated the source type IDs of eight vehicle make/models because the registration database identified them as a light truck, but the EPA emission certification list shows they are regulated as a car. The makes and models changed to source type 21 were the following: Acura ZDX, Buick Encore, PT Cruiser, Honda Element, Kia Soul, Nissan Juke, Scion xB, and Suzuki XL7. The full list of 172 makes and models ERG checked against the EPA certification list are in Appendix A.
- 2. Moved all GVW class 3 trucks from their initial assignment of the 30's source type ID (light trucks) to the 50's source type ID (single unit trucks).** According to EPA, class 3 trucks are considered single unit trucks by DOT and therefore are modeled exclusively as heavy-duty source types in MOVES¹.
- 3. Assigned MOVES regulatory class IDs based on the GVW class and vehicle type field.** Table 3 shows the definitions of regulatory class IDs. The original purpose of assigning regulatory class was to generate a new 'sampleVehiclePopulation' table for MOVES, but EPA later cut this task from the scope of the work assignment. However, the regulatory classes were still useful in the next step.
- 4. Deleted or remapped fuel types that are not allowed in MOVES for certain combinations of regulatory class and fuel.** The allowable fuel types (and IDs) modeled for on-road vehicles in MOVES are gasoline (1), diesel (2), CNG (3), E85 capable (5), and electricity (9). Furthermore, the model database contains emission rates that only cover of source type, regulatory class, and fuel shown Table 4. All others vehicle-fuel types that may exist (for example, CNG passenger cars), do not have MOVES emission rates and therefore ERG updated the categories to a similar fuel type so that MOVES can still estimate their emissions. In consultation with EPA, ERG deleted electric populations of heavy-duty regulatory classes and renamed E85 capable vehicle populations to gasoline. For light-duty, natural gas fuel types were renamed to gasoline.
- 5. Grouped source type IDs together that could not reliably be distinguished based on available data.** ERG, EPA, and IHS agreed that the following pairs of source type IDs are not distinguishable from one another based on the registration and title information alone.

¹ 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.

- Non-school buses – intercity vs. transit (source types 41 and 42)
- Single Unit Trucks – short haul vs. Long haul (source types 52 and 53)
- Combination Trucks – short haul vs. long haul (source types 61 and 62)

For purposes of this work assignment, the populations of 41 and 42 were summed together, and likewise 52 and 53 as well as 61 and 62 prior to calculating age distributions and AVFT. Furthermore, long-haul source types 53 and 62, the values were calculated nationwide rather than individual county registration due to the nature of their travel patterns. Therefore, each county in the deliverables shows that source type 41 and 42 have the same age distribution. They do not have the same AVFT only because MOVES consider intercity buses (source type 41) to be 100% diesel and transit buses (source type 42) can be gasoline, diesel, or CNG. The age distributions and AVFT are not identical between 52 and 53 or 61 and 62 due to the national grouping of data for long-hauls.

Table 3. Regulatory Class Definitions in MOVES2014²

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)

² *ibid.*, p. 10.

Table 4. Summary of Source Type, Fuel Type, and Regulatory Class Combinations in MOVES2014³

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

Spatial Grouping of the County Data

The spatial grouping of the age distribution and AVFT data reflect the representative county groups for the 2014 v2 NEI for most source types. For example, in a hypothetical 5-county group, all the age distributions and AVFT fuel type fractions are identical by county for the same source type. The distributions were a population-weighted average of the member counties in each group, and so the grouped data looks most similar to the county with the highest population. The spatial grouping applied to all source types except long-haul trucks (source type IDs 11, 21, 31, 32, 41, 42, 43, 51, 52, 54, and 61). For the long-haul source type IDs 53 and 62, as already mentioned above the data reflect the nationwide IHS population distribution of the appropriate trucks.

The age distribution and AVFT were also delivered by individual county, rather than representative county groups, with the exception of long-hauls which are national. Note that the sample sizes of certain source types (e.g., refuse trucks) are very low in some areas of the country.

Gap Filling

Even when using population data from multiple-county groups, some areas did not have any vehicles in some source type categories (specifically, refuse trucks and non-school buses). In these cases, ERG substituted the nationwide IHS age distribution for refuse trucks or non-school buses, as needed. No gap filling was performed on any AVFT deliverable. The AVFT

³ *ibid.*, p. 16.

table is left with “holes” where certain model years did not have coverage in the registration data. MOVES will revert to its default method to split source type population into fuel types by model year in these cases.

Output Files Delivered to EPA

The results were reported both in MySQL database and a comma separated (CSV) format. These files contained information for all counties. Tables 5 and 6 show the fields in the deliverables for age distribution and AVFT. The filenames posted on FTP for EPA are the following, and contain two tables each.

1. MySQLDatabase_ihs_avft_agedist_30jun17b_grouped.zip
2. CSVs_grouped.zip
3. MySQLDatabase_ihs_avft_agedist_17jul17_individual.zip
4. CSVs_individual.zip

Table 5. Table Contents in the Nationwide County Level Source Type Age Distribution Deliverables

Field Name
countyID
sourceTypeID
yearID (2014)
ageID
ageFraction
fuelEngFraction

Table 6. Table Contents in the Nationwide County Level AVFT Deliverables

Field Name
countyID
sourceTypeID
modelYearID
fuelTypeID
engTechID
fuelEngFraction

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
ACURA			
	MDX	TRK	Light Truck
	RDX	TRK	Light Truck
	SLX	TRK	Light Truck
	ZDX	CAR	Light Truck
BMW			
	X3	TRK	Light Truck
	X5	TRK	Light Truck
	X6	TRK	Light Truck
BUICK			
	Enclave	TRK	Light Truck
	Encore	CAR	Light Truck
	Rainier	TRK	Light Truck
	Rendezvous	TRK	Light Truck
	Terraza	TRK	Light Truck
CADILLAC			
	Suburban	TRK	Light Truck
	Escalade	TRK	Light Truck
	SRX	TRK	Light Truck
	Sierra	TRK	Light Truck
CHEVROLET			
	Astro	TRK	Light Truck
	Blazer	TRK	Light Truck
	Captiva	TRK	Light Truck
	Colorado	TRK	Light Truck
	El Camino	TRK	Light Truck
	Equinox	TRK	Light Truck
	Express G1500	TRK	Light Truck
	Express G2500	TRK	Light Truck
	Express G3500	TRK	Light Truck
	Geo Tracker	TRK	Light Truck
	HHR	TRK	Light Truck
	Jimmy	TRK	Light Truck
	Lumina	TRK	Light Truck
	Safari	TRK	Light Truck
	Savanna	TRK	Light Truck
	Sonoma	TRK	Light Truck

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
	Tahoe	TRK	Light Truck
	Tracker	TRK	Light Truck
	Traverse	TRK	Light Truck
	Uplander	TRK	Light Truck
	Venture	TRK	Light Truck
CHRYSLER			
	Aspen	TRK	Light Truck
	PT Cruiser	CAR	Light Truck
	Pacifica	TRK	Light Truck
	Town & Country	TRK	Light Truck
	Voyager	TRK	Light Truck
DODGE			
	Caravan	TRK	Light Truck
	Durango	TRK	Light Truck
	Nitro	TRK	Light Truck
FORD			
	Aerostar	TRK	Light Truck
	Bronco	TRK	Light Truck
	E150	TRK	Light Truck
	E250	TRK	Light Truck
	E350	TRK	Light Truck
	Edge	TRK	Light Truck
	Escape	TRK	Light Truck
	Expedition	TRK	Light Truck
	Explorer	TRK	Light Truck
	F100	TRK	Light Truck
	F150	TRK	Light Truck
	F250	TRK	Light Truck
	F350	TRK	Light Truck
	Freestar	TRK	Light Truck
	Freestyle	TRK	Light Truck
	Taurus X	TRK	Light Truck
	Windstar	TRK	Light Truck
GMC			
	Envoy	TRK	Light Truck
	Denali	TRK	Light Truck
	Safari	TRK	Light Truck
	Savanna	TRK	Light Truck

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
	Yukon	TRK	Light Truck
HONDA			
	CR-V	TRK	Light Truck
	Element	CAR	Light Truck
	Odyssey	TRK	Light Truck
	Passport	TRK	Light Truck
	Pilot	TRK	Light Truck
	Ridgeline	TRK	Light Truck
HUMMER			
	H1	TRK	Light Truck
HYUNDAI			
	Entourage	TRK	Light Truck
	Santa Fe	TRK	Light Truck
	Tuscon	TRK	Light Truck
	Veracruz	TRK	Light Truck
ISUZU			
	Amigo	TRK	Light Truck
	Axiom	TRK	Light Truck
	Oasis	TRK	Light Truck
	Rodeo	TRK	Light Truck
	Trooper	TRK	Light Truck
JEEP			
	CJ5	TRK	Light Truck
	CJ6	TRK	Light Truck
	CJ7	TRK	Light Truck
	CJ8	TRK	Light Truck
	Cherokee	TRK	Light Truck
	Liberty	TRK	Light Truck
	Wagoneer	TRK	Light Truck
KIA			
	Borrego	TRK	Light Truck
	Sedona	TRK	Light Truck
	Sorento	TRK	Light Truck
	Soul	CAR	Light Truck
	Sportage	TRK	Light Truck
LAND ROVER			
	Defender	TRK	Light Truck
	Discovery	TRK	Light Truck

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
	Freelander	TRK	Light Truck
	Range Rover	TRK	Light Truck
LEXUS			
	GX470	TRK	Light Truck
	LX470	TRK	Light Truck
	LX570	TRK	Light Truck
	RX300	TRK	Light Truck
	RX330	TRK	Light Truck
	RX350	TRK	Light Truck
LINCOLN			
	Aviator	TRK	Light Truck
	Navigator	TRK	Light Truck
MAZDA			
	CX-5	TRK	Light Truck
	CX-7	TRK	Light Truck
	CX-9	TRK	Light Truck
	MPV	TRK	Light Truck
	Navajo	TRK	Light Truck
	Tribute	TRK	Light Truck
MERCEDES			
	G500	TRK	Light Truck
	G55	TRK	Light Truck
	GL320	TRK	Light Truck
	GL450	TRK	Light Truck
	GL550	TRK	Light Truck
	ML320	TRK	Light Truck
	ML350	TRK	Light Truck
	ML430	TRK	Light Truck
	ML500	TRK	Light Truck
	ML55	TRK	Light Truck
	ML550	TRK	Light Truck
	ML63	TRK	Light Truck
	R320	TRK	Light Truck
	R350	TRK	Light Truck
	R500	TRK	Light Truck
MERCURY			
	Mariner	TRK	Light Truck
	Mountaineer	TRK	Light Truck
	Villager	TRK	Light Truck

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
MITSUBISHI			
	Endeavor	TRK	Light Truck
	Montero	TRK	Light Truck
	Outlander	TRK	Light Truck
NISSAN			
	Juke	CAR	Light Truck
	Murano	TRK	Light Truck
	Pathfinder	TRK	Light Truck
	Quest	TRK	Light Truck
	Rogue	TRK	Light Truck
	Xterra	TRK	Light Truck
OLDSMOBILE			
	Bravada	TRK	Light Truck
	Silhouette	TRK	Light Truck
PLYMOUTH			
	Voyager	TRK	Light Truck
PONTIAC			
	Aztek	TRK	Light Truck
	Montana	TRK	Light Truck
	Torrent	TRK	Light Truck
	Transport	TRK	Light Truck
PORSCHE			
	Cayenne	TRK	Light Truck
	Macan	TRK	Light Truck
SATURN			
	Acadia	TRK	Light Truck
	Outlook	TRK	Light Truck
	Relay	TRK	Light Truck
	Vue	TRK	Light Truck
SCION			
	xB	CAR	Light Truck
SUZUKI			
	Vitara	TRK	Light Truck
	Samurai	TRK	Light Truck
	Sidekick	TRK	Light Truck
	XL7	CAR	Light Truck
	X90	TRK	Light Truck
TOYOTA			
	4Runner	TRK	Light Truck

Appendix A

Make	Model	EPA Classification for MOBILE Modeling	IHS data classification
	FJ Cruiser	TRK	Light Truck
	Highlander	TRK	Light Truck
	Land Cruiser	TRK	Light Truck
	Previa	TRK	Light Truck
	RAV4	TRK	Light Truck
	Sequoia	TRK	Light Truck
	Sienna	TRK	Light Truck
	Venza	TRK	Light Truck
VOLKSWAGEN			
	Routan	TRK	Light Truck
	Tiguan	TRK	Light Truck
	Touareg	TRK	Light Truck

Shading indicates a model where the IHS classification disagrees with the one from EPA.