A Framework for Characterizing the Ambient Conditions Experienced by Light Duty Vehicles in the United States

Matthew Moniot , Jason Lustbader , Eric Wood , Byungho Lee , Justin Fink & Scott Agnew

Key Words: light-duty vehicle, ambient conditions, representative, temperature, solar irradiation, humidity, air density, energy consumption, fuel economy, real-world

Abstract

The relationship between ambient conditions and light-duty vehicle energy consumption has been widely researched. Relatively little effort, however, has been dedicated to understanding representative ambient conditions a light-duty vehicle may experience. The framework introduced in this paper provides a means of quantifying ambient conditions specific to light-duty vehicle operation by incorporating both when and where vehicles are driven. The analysis presented expands the literature beyond solely focusing on temperature; distributions for humidity, solar irradiance, and air density are also included. A procedure is presented that calculates the ambient condition distributions for each metric by relating open-source data sets describing representative vehicle utilization and representative ambient conditions. While this study explores ambient conditions related to light-duty vehicle utilization, the framework may also be applied to separate vocations. Finally, the paper concludes with an example use case of the ambient condition weighting process. A binning methodology is introduced that facilitates insight into vehicle energy consumption in response to ambient conditions at the national and local levels while minimizing the number of tests or simulations required.

Introduction, Literature Review

While greenhouse gas emissions (GHG) from various industries continue to fall, the United States transportation sector is distinct in its steady or rising share of emissions over time. In 2017, 29% of GHG emitted within the United States were attributed to the transportation sector, of which the majority were credited to light-duty vehicles (LDVs) [1]. Beyond contributing to GHG emissions, the transportation sector also reduces local air quality; in 2014 transportation was identified as the leading source of domestic nitrogen oxide emissions [2]. These environmental realities have motivated efforts to improve the environmental impact of LDVs and corresponding efforts by automotive manufacturers to improve the energy efficiency of their vehicle offerings. Efforts by automakers include introducing fuel-saving technologies and an increased share of plug-in vehicles [3].

Many studies have documented the influence of ambient conditions on vehicle energy efficiency. Colder temperatures are associated with reduced energy efficiency for conventional vehicles due to a variety of factors, including reduced transmission efficiency [4], non-linear viscosity of engine lubricant [5], and other effects. Colder temperatures also negatively influence the performance of battery electric vehicles, primarily due to reduced performance of the vehicle battery [6] and increased auxiliary loads to maintain passenger comfort within the cabin [7]. Warm ambient conditions have also been shown to influence LDV energy consumption, largely due to increased loads associated with air conditioning operation [8] Acknowledging the impact of ambient temperature on vehicle energy consumption, the U.S. Environmental Protection Agency (EPA) expanded the federal test procedure in 2006 to include additional cycles capturing impacts of cold ambient conditions (-6.7°C) and warm ambient conditions (+35°C) starting in 2008 [9]. Incorporation of ambient temperatures into the estimation of real world energy consumption is a useful improvement, but such efforts do not include the full impact of ambient conditions. Studies have also shown humidity, solar radiation, and air density influence vehicle energy consumption in addition to temperature [10] [11] [12] .

Despite the importance of ambient conditions on vehicle energy consumption understood by automakers and regulators, there is limited information regarding the real-world ambient conditions experienced by LDVs within the United States. For example, there is no information readily available describing the prevalence of given conditions relevant to energy consumption, such as extremely low temperatures, high solar irradiance, etc. Such an analysis is needed to inform the prioritization of energy saving technologies and calculation of representative real-world vehicle energy efficiency.

The most specific information the authors are aware of is an estimation of vehicle miles traveled (VMT) versus ambient temperature by LDVs calculated from the EPA model MOVES [13]. The results are informative, but solely at the national level with no information regarding local or regional ambient condition distributions. Yurksel et al. [14] introduced a framework for incorporating the ambient temperature experienced by LDVs by relating plausible Typical Meteorological Year 3 (TMY3) weather data [15] against driving profiles sampled from the National Highway Transportation Survey (NHTS). While an effective approach for the application being considered, the study does not incorporate the geographic density of vehicle adoption, which is necessary to produce representative ambient conditions at a regional or national level.

Other studies [16] [17] perform a similar routine for appending plausible ambient temperature information to NHTS drive cycles with a scope specific to individual drive cycles rather than a nationally informative distribution of experienced ambient conditions. Additional studies, primarily focused on real-world technology benefits, also attempt to model representative thermal states but in a framework that is technology-specific and limited solely to ambient temperature [18] [19]. The authors are not aware of studies that attempt to quantify the representative prevalence of multiple ambient conditions simultaneously (temperature, solar irradiance, humidity, and air density) while considering both temporal and spatial factors associated with typical LDV operation.

The ambient condition assessment introduced in this paper is intentionally vehicle, technology, and drive cycle agnostic. As such, the developed framework is intended solely to quantify representative ambient conditions that a vehicle may experience. Results are intended to be application-neutral and compatible with real-world assessments of any relevant vehicle technology. In addition, the methodology is highly flexible, enabling the development of nationally representative ambient conditions that may also be calculated at the local level. Results can be weighted to describe local representative conditions for 504 urbanized areas and rural regions throughout the United States, including Alaska and Hawaii. This capability is demonstrated in an example application of the calculated temperature profiles, involving a binning methodology that facilitates the calculation of representative energy consumptions in response to ambient conditions at the national and local levels while minimizing the simulation or testing resources required.

This paper first introduces the underlying weather data set utilized for the study. Next, methodologies are introduced for performing temporal and geographical adjustments to account for when and where LDVs are typically driven. The temporal and geographical adjustments are aggregated at the national level and a binning procedure is outlined that describes how to produce nationally representative energy consumption values that may be localized to various regions throughout the United States. The paper concludes by discussing limitations associated with the framework and plans for future research.

Finally, the authors acknowledge that just as an LDV will experience a range of ambient conditions, it will also experience a range of different driving cycles (city, highway, etc.). The scope of this analysis is limited exclusively to characterizing representative ambient conditions; a companion study explores nationally representative drive cycles in a similar fashion [20].

Introduction to TMY3 Data

A comprehensive, representative, and geographically diverse data set is needed for understanding the full range of ambient conditions experienced by LDVs in the United States. Temporal coverage is needed for characterizing the variation in ambient conditions over a given day and between seasons. In addition, geographic coverage is required to describe ambient conditions throughout the entirety of the United States. The TMY3 data set, published by the National Renewable Energy Laboratory, was identified as meeting all the necessary criteria [15]. TMY3 data comprise real-world measured values across several years. For each location and for each month of the TMY3 year, the most representative corresponding month for all years of data at that location is identified. Each of these representative months, from various years, is then aggregated into a single, representative year by location. An algorithm is used to smooth the data at the transition between months. Ambient conditions are described for each location at an hourly level with values for ambient temperature, solar irradiance, relative humidity, and many other parameters. In total, there are 8,760 hourly weather values for each of the 1,020 stations for which full yearly weather coverage is available within the United States. Figure 1 illustrates the TMY3 stations from which data were used within this study. Note that a total of 10 weather stations in Guam, Puerto Rico, and the Virgin Islands are excluded from this analysis.

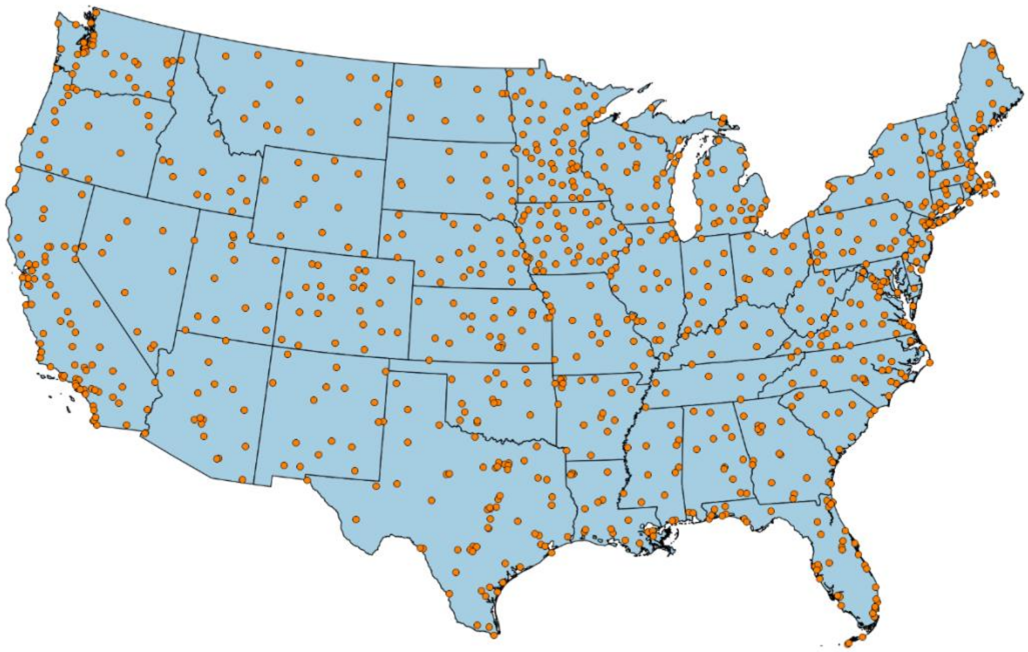

Figure 1. Geographical location of TMY3 stations across the United States. Although not shown, stations located in Alaska and Hawaii are included in this analysis.

TMY3 Data Set Adjustments

The raw TMY3 data introduced in the previous section provide representative ambient conditions for each location across a full year. Representative conditions experienced during vehicle operation, however, are different than the overall ambient conditions experienced at a given location. For example, LDVs are disproportionately operated during daytime hours and driven near or within population centers. Thus, the raw TMY3 data must be adjusted to account for temporal and geographical considerations to better describe representative ambient conditions present during vehicle operation.

Temporal Analysis – "When are Vehicles Driven?"

Vehicle utilization is highly correlated with the time of day (TOD). Typical LDV travel behavior consists of small amounts of driving overnight with significantly more activity during daytime hours. LDV use by TOD is quantified through analysis of the publicly available 2017 NHTS data [21]. The survey is considered to be nationally representative because it contains vehicle trip information across a variety of regions throughout the United States and significant sample sizes. Survey results were filtered to include only entries describing LDV travel. Fields of interest within the data set include LDV trip information such as local start time, local end time, and miles traveled. Vehicle activity by TOD was understood by relating the VMT across all trips in the data sets against the TOD the trips occurred. Figure 2 shows the resulting VMT-TOD relationship with percentage values describing the share of VMT taking place in each hour. The trace relating VMT by TOD shows small amounts of overnight driving, with peaks during the day likely arising from morning and afternoon commuting behavior.

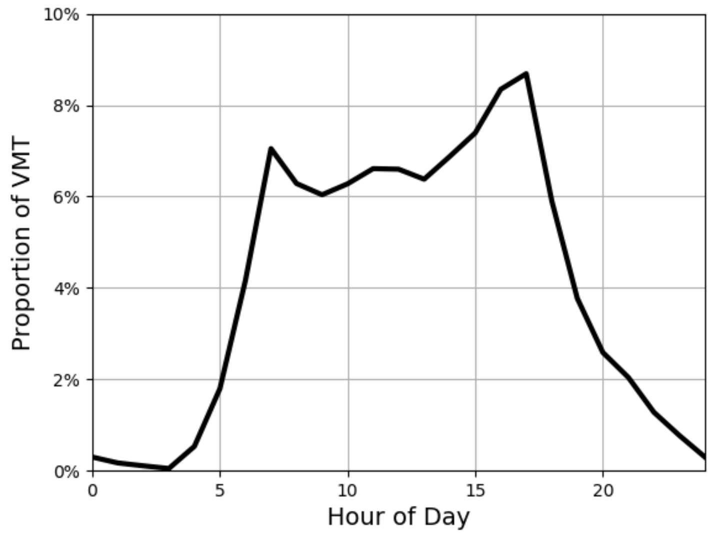

Figure 2. LDV travel by TOD.

Knowledge of vehicle use by TOD is used to inform the characterization of representative weather conditions during vehicle operation. Ambient conditions during daytime hours coincide with a disproportionately large amount of vehicle activity. Thus, daytime ambient conditions are more likely to be experienced by LDVs compared to conditions occurring during overnight hours when there is minimal driving. The VMT-TOD relationship shown in Figure 2 is used to emphasize ambient conditions that take place during periods of peak vehicle use and deemphasize conditions occurring during periods of limited vehicle use. Temporal adjustments to the TMY3 data set using the established VMT-TOD relationship are performed on an hourly basis by applying weights relative to the amount of VMT taking place. An example demonstration of the temporal weighting process follows for a single station representing conditions in Phoenix, Arizona.

As stated in the previous section, TMY3 data contain hourly information across a typical full year for a multitude of different parameters, including ambient temperature. A histogram of hourly temperatures from the raw TMY3 data set for Phoenix is shown in Figure 3. While the raw TMY3 data displayed provide valuable insight into conditions a vehicle may experience in Phoenix, the relative frequencies of different temperature conditions may not yet be inferred. For instance, although Figure 3 suggests that temperatures between +25°C and +30°C represent 15% of annual hours, there is no understanding of how often these temperatures occur *during vehicle operation.* Accurately capturing experienced ambient conditions during vehicle use requires incorporating the VMT-TOD relationship illustrated in Figure 2.

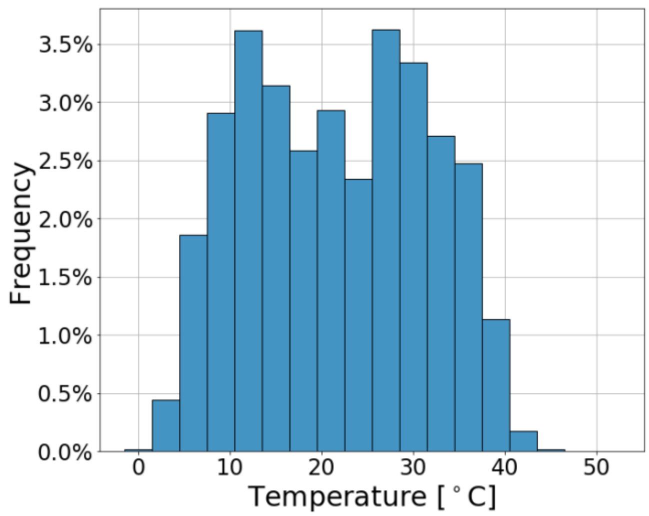

Figure 3. Raw TMY3 annual temperature distributions for a weather station near Phoenix.

The same raw weather data illustrated in Figure 3 may be better understood by categorizing ambient conditions by TOD. The 8,760 hours of TM3 weather data for Phoenix describe ambient conditions for each hour (24) of each day (365). Individual histograms describing temperature distributions are created for each hour of the day and depicted as a ridgeline plot in Figure 4. Each individual histogram may be understood as the variance of all ambient conditions occurring at a given hour over the course of a year. The sample size of each histogram shown in Figure 4 is 365, as there is one occurrence of each hour per day of the year.

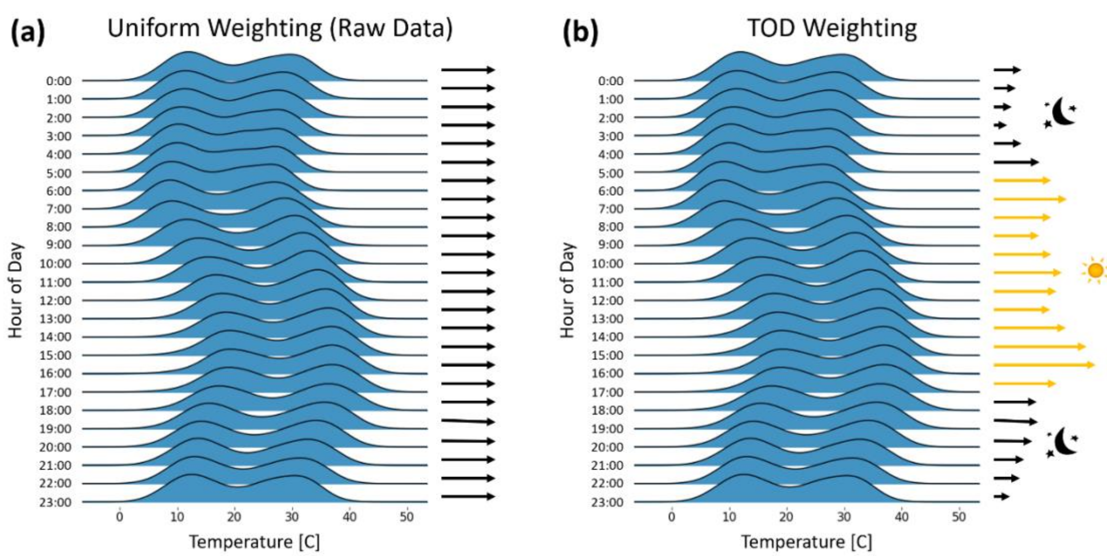

Figure 4. TMY3 annual temperature distributions for a weather station near Phoenix broken out by hour of day.

Note that aggregating all 24 hourly histograms uniformly into a single distribution would reproduce the plot shown in Figure 3. The underlying data visualized in both Figure 3 and Figure 4 are identical; the two plots are simply different views of the same, unweighted information. Temporal considerations accounting for when vehicles are driven are incorporated by weighting the histograms *non*-uniformly. Rather than equally aggregating the 24 hourly histograms in Figure 4, the VMT-TOD relationship is utilized to emphasize constituent hourly distributions in proportion to the corresponding share of VMT taking place during a given hour. For instance, VMT-TOD analysis suggests that 0.3% of driving takes place at 12 AM versus 6.6% of driving at 12 PM. These disparities in miles traveled by time of day underpin hourly weights which are applied to each hour within the TMY3 dataset on the basis of relative VMT (Equation 1). When applying this equation, the values occurring at 12 AM are accordingly weighted by 0.3% versus 6.6% for the values as 12 PM.

$$TOD_h = \frac{VMT_{TOD\,h}}{\sum_{i=1}^{24} VMT_{TOD\,i}} \qquad \text{Equation 1}$$

Where TOD_h = the time of day weight applied to ambient weather data occurring on hour h of the day and $VMT_{TOD\,h}$, $VMT_{TOD\,i}$ = the total VMT occurring during a given hour as calculated from the NHTS

The resulting TOD-weighted distribution of weather is considered representative for the ambient conditions experienced for LDV travel near the Phoenix TMY3 weather station. Figure 5 compares unweighted and temporally weighted distributions for a full year of weather data for Phoenix. The blue curve, representing unweighted weather values, visualizes the same data previously introduced in Figure 3 and Figure 4. Weather from each hour of the day is assumed to represent an equal amount of VMT. Meanwhile, the orange curve represents the same weather data from the Phoenix station after hourly-weights have been applied using Equation 1. As expected, the temporally weighted distribution skews toward warmer temperatures; warmer temperatures typically coincide with daytime hours and periods of more frequent vehicle travel. Note that this analysis assumes the VMT-TOD relationship to be similar across all regions throughout the United States; the adjustment process is performed for all locations in the TMY3 data set utilizing the established VMT-TOD relationship in an identical manner. The analysis is also repeated for the separate ambient condition parameters of interest—humidity, solar irradiance, and air density. Similar histograms showing weighted and unweighted traces for these additional metrics can be found in the National Ambient Condition Aggregation section, which incorporates both temporal and geographical weighting.

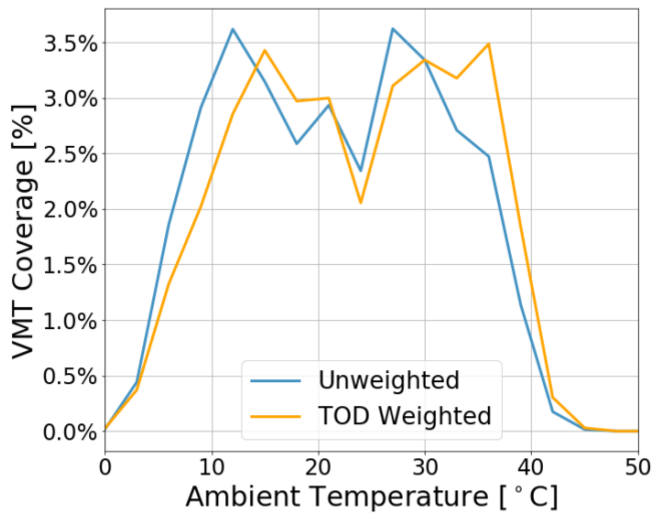

Figure 5. Unweighted and temporally weighted ambient temperature distributions using TMY3 data for a single weather station in Phoenix.

Finally, the temporal analysis introduced in this section is not limited to characterizing LDV travel behavior. The temporal weighting analysis of TMY3 weather data could have been similarly performed assuming a different utilization profile for different vocations such as heavy-duty vehicle operation or even applications outside of transportation entirely. The framework can be generalized, relating typical operation patterns with publicly available and representative ambient condition data.

Geographic Analysis – "Where are Vehicles Driven?"

Station-Level Weighting

Data from the 1,010 TMY3 stations used in this analysis include remote weather stations in rural regions and stations located in dense urbanized areas. Just as ambient conditions occurring at 12 AM must be emphasized differently than weather occurring at 12 PM, so too must ambient conditions reported by stations associated with small amounts of VMT versus ambient conditions reported by stations associated with large amounts of VMT. For example, consider how differently stations representing ambient conditions for metropolitan regions must be emphasized more than stations representing conditions for rural regions when characterizing ambient conditions at the state or national level. A methodology, similar to the VMT-TOD weighting from the last section, is needed that incorporates the amount of nearby VMT associated with each TMY3 weather station. These station weights account for the geospatial distribution of driving and will be referred to as VMT-Geographic (GEO) weights.

First, region-level counts are obtained using values reported by the Highway Performance Monitoring System (HPMS). These region-level counts provide a spatial understanding of where VMT is taking place throughout the United States. Additional spatial resolution surrounding where VMT is taking place is obtained by using vehicle registrations to map VMT at the ZIP code level. Finally, weights are calculated for each station by identifying the nearest station to each ZIP code and summing all corresponding VMT values. The attributed station weights are proportional to the amount of nearby driving calculated. The remainder of this section will introduce data sources and how they are used to produce station-level weights for aggregating ambient conditions across regions containing multiple stations.

Data provided by the HPMS are used to understand *where* LDVs are predominately driven. Specifically, values used are from Table HM-71 reported by the HPMS [22]. Data from the HPMS are used in lieu of NHTS values for understanding VMT by geography given the greater amount of data underpinning reported values; HPMS values are derived from a high quantity of traffic-counts while NHTS values result from household surveys. While the NHTS is useful for describing representative trends (such as when vehicles are driven by time of day), the HPMS is useful for describing aggregate values (such as total VMT in a region). Table HM-71 is specifically used for this application, as it describes the annualized VMT within all 504 urbanized areas – defined as geographies with more than 50,000 inhabitants – and 50 rural areas within the United States.

The total VMT for a given state is considered to be the sum of VMT values for all encompassed urbanized areas and rural areas. In other words, full VMT coverage is assumed from the values present in the HPMS table. Within the weighting process expanded below, regions and areas are distinct; the word "area" shall refer to a defined urbanized area or rural area and the word "region" shall refer to an ambiguously sized geography containing one or more areas (such as a state, collection of states, or the entire United States).

HPMS VMT values broken out by area are useful for understanding where LDVs are operated. The values, however, are only provided at the urban or rural area. Additional analysis is needed to relate the amount of driving in an area (urban or rural) and the corresponding amount of spatial weighting a nearby station should receive. A fictional rectangular state is depicted in Figure 6 with a rural area, three urbanized areas, and three unique TMY3 stations. A procedure is needed to map the area-level VMT values known for each of the four areas onto the three stations in accordance to geographical proximity.

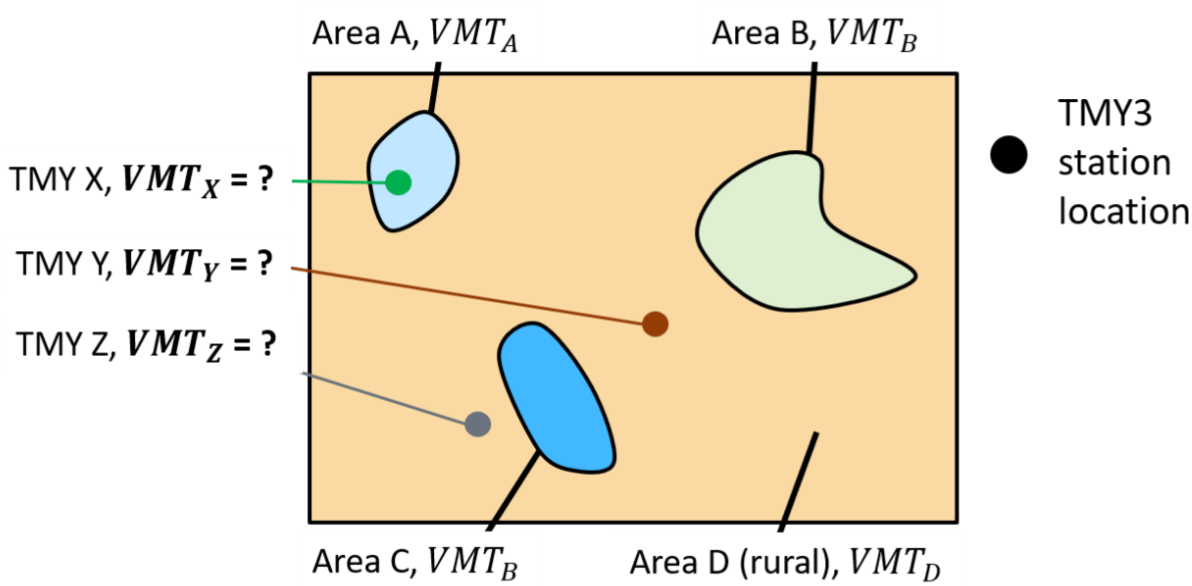

Figure 6. Example depiction of a rectangular state that includes three urban areas, one rural area, and three TMY3 stations.

LDV registrations, sourced by IHS Markit [23] at the ZIP code level, are used to infer the spatial distribution of driving at a more granular level than the area-level VMT that are available. Utilizing vehicle registration information as a proxy for VMT relies on an assumption that the amount of driving in a location is related to the amount of registered LDVs. This assumption is supported by a comparison of aggregated LDV vehicle registrations at the area level compared to the amount of VMT reported by the HPMS. Figure 7 contains a comparison of these results, illustrating a strong linear relationship between vehicle registrations and VMT. This relationship, established at the area level, is assumed to also be present at the ZIP code-level.

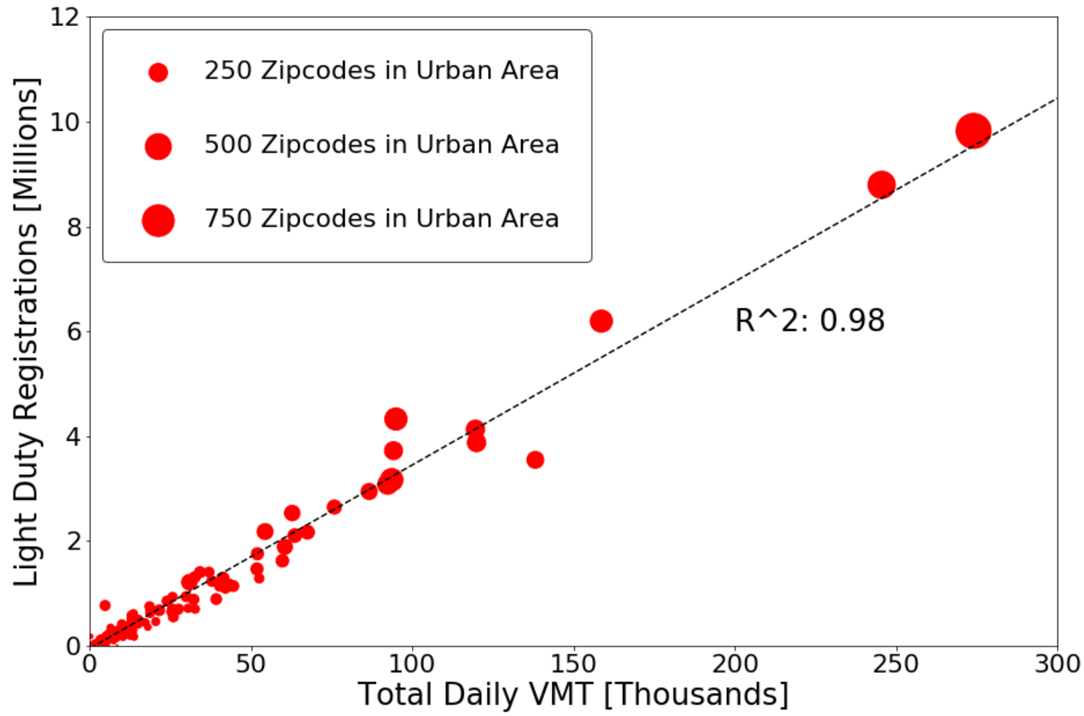

Figure 7. Area-level comparison of VMT (sourced from HPMS) and aggregate vehicle registrations from constituent ZIP codes within the corresponding area (sourced from IHS Markit).

The region depicted in Figure 6 is revisited in Figure 8, now with the location of ZIP code centroids included. ZIP codes, being utilized as a crosswalk between HPMS areas and discrete TMY3 station locations, are affiliated with both TMY3 stations and urbanized areas; a particular ZIP code is affiliated with a TMY3 station on the basis of the nearest neighbor and with a HPMS area on the basis of which HPMS area the ZIP code resides within. TMY3 station affiliations for each ZIP code centroid are displayed through the symbol color.

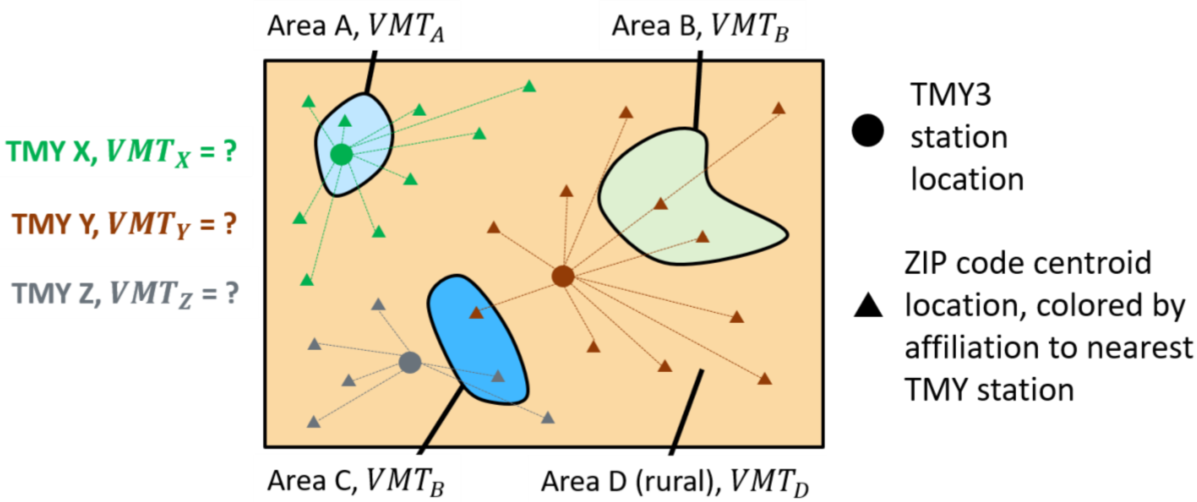

Figure 8. The same region depicted in Figure 6. Example depiction of a rectangular state that includes three urban areas, one rural area, and three TMY3 stations.

The proportional relationship between vehicle registrations and VMT established in Figure 7, combined with the spatial resolution of ZIP code centroids, is used to map area-level HPMS VMT values onto each TMY3 station. First, VMT values are assigned from HPMS areas onto its associated ZIP code centroids by percentage of vehicle registrations (Equation 2 and Equation 3). For example, consider an area containing 1 million miles of HPMS VMT and three ZIP code centroids. Five hundred vehicles are registered in the first ZIP code, 200 in the second, and 300 in the third. Application of Equation 2 and Equation 3 produces VMT values of 500k miles, 200k miles, and 300k miles attributed to the three ZIP codes, respectively.

$$VR_A = \sum_{i=1}^{|a|} VR_i \qquad \text{Equation 2}$$

$$VMT_i = \frac{VR_i}{VR_A} * VMT_A \qquad \text{Equation 3}$$

Where $|a|$ = set of all ZIP code centroids contained within an area A
VR_i = vehicle registrations in ZIP i
VR_A = total registrations within area A
VMT_A = total VMT for area A
and VMT_i = VMT in ZIP code i

The attributed VMT for all ZIP codes are grouped and summed by their TMY3 affiliation, producing station-level VMT counts (Equation 4). These counts, informed by underlying HPMS values, describe the amount of driving taking place near each station and are used to understand the relative weight (Equation 5) to apply to each station when aggregating ambient conditions to regional levels. Figure 9 illustrates the attributed VMT for each of the 1,010 TMY3 stations for which this procedure was performed.

$$VMT_{TMY_N} = \sum_{i=1}^{|A|} VMT_i$$ Equation 4

Equation 5
$$GEO_{TMY_N} = \frac{VMT_{TMY_N}}{VMT_A}$$

Where $|A|$ = set of all ZIP code centroids for which TMY3 station N is the closest
VMT_{TMY_N} = VMT attributed to TMY3 station N
and GEO_{TMY_N} = GEO weight associated with all ambient conditions reported by TMY3 station N

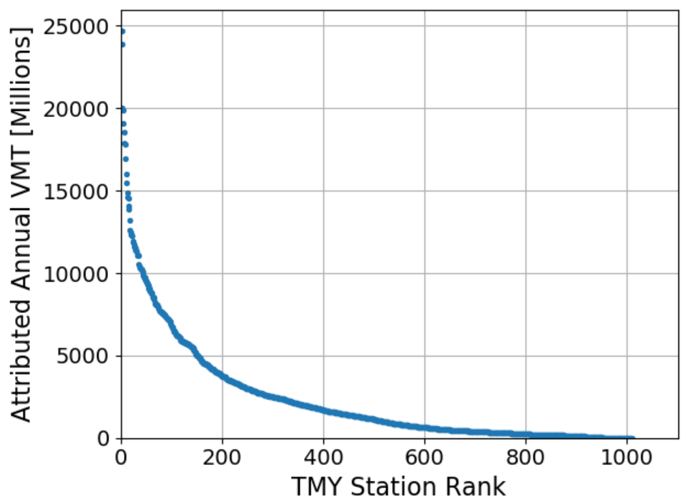

Figure 9. Attributed VMT counts by TMY3 station, sorted by rank. The VMT-GEO weights applied are proportional to the counts shown.

Finally, note that the process of attributing a finite supply of HPMS VMT to TMY3 stations addresses a potential biasing of weather conditions near areas with multiple TMY3 stations. A high density of TMY3 stations in a single geographic area produces moderate weights because the nearby VMT is shared by all relevant stations. In Figure 9, stations with a large amount of attributed VMT are typically located near areas with a large amount of HPMS VMT and with few nearby competing stations. Stations with low amounts of attributed VMT are typically in remote locations or near several other TMY3 stations that siphon off available HPMS VMT.

Multi-Station Aggregation

Calculation of station-level weights enables aggregating weather results across multiple stations at the region, state, and national level. Aggregation at the region level across multiple stations is performed by applying the VMT-TOD and VMT-GEO weights discussed above to all hours and across all relevant stations (Equation 6).

Equation 6
$$w_{j,N} = TOD_j * GEO_{TMY_N}$$

Where $w_{j,N}$ = the hour-level weight associated with an ambient condition at hour of day *j* and reported by TMY3 station *N*

and TOD_j = VMT-TOD weighting, or the percentage of VMT occurring at hour of day j

For example, summarizing the ambient conditions experienced by the region depicted in Figure 6 and Figure 8 would consist of first individually weighting the weather conditions by the hour of day they occur in accordance to the amount of VMT by TOD. The TOD weights would be applied identically for each station under the assumption that driving activity by TOD does not vary significantly with respect to geography. Next, additional station-specific weights would be applied to the hourly ambient conditions in accordance to the amount of nearby VMT. Finally, the weighted hourly results from each station are aggregated in accordance with hourly weight. The prevalence of specified ambient conditions at the region level, such as hot conditions with temperatures above +35°C, may be quantified by summing the weights associated with qualifying hours relative to the total weight across the region. This procedure is expanded upon further in the Representative Evaluation Points: Ambient Condition Binning, Weighting, and Localization section.

National Ambient Condition Aggregation

A quantitative distribution of the ambient conditions present during LDV operation at the national level is obtained by applying the temporal and spatial adjustments introduced in the previous section. First, ambient conditions from each of the 1,010 stations are weighted by TOD to account for when vehicles are driven. Next, ambient conditions are aggregated to the national level using the TMY3 station weighting procedure introduced to account for where vehicles are driven. Figure 10a – Figure 10d compare the unweighted distribution of raw TMY3 values (blue trace) to the LDV temporally and spatially weighted TMY3 values (orange trace) for all ambient conditions of interest. The weighted probability distributions for all four ambient condition metrics may be found in the supplemental files accompanying this report.

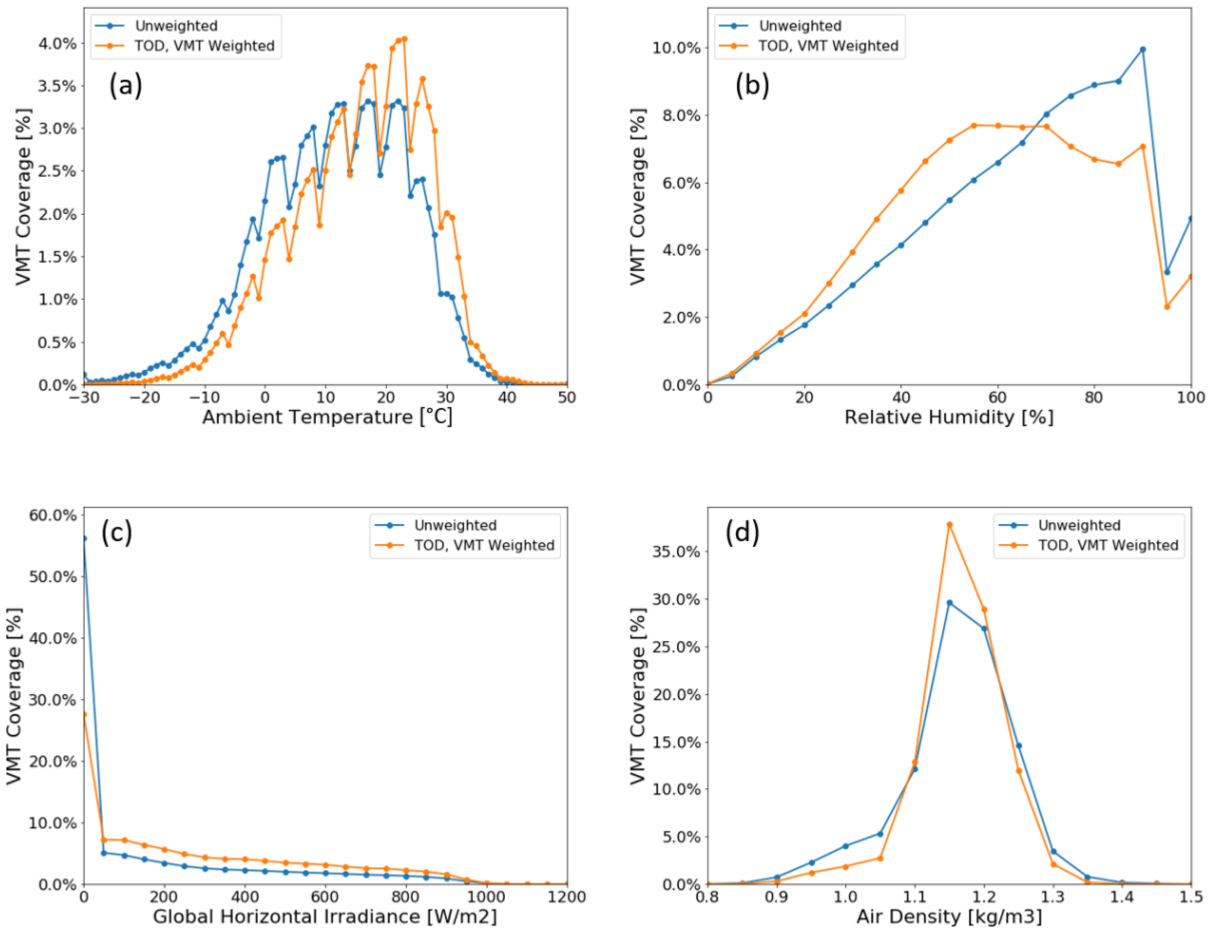

Figure 10. Nationally aggregated ambient condition distributions for temperature (a), relative humidity (b), solar irradiance (c), and air density (d).

A considerable shift is shown in the distribution of ambient conditions for each metric of interest between the unweighted and weighted cases. Although the metrics are shown individually, note that many are inherently related (such as the positive relationship between solar irradiance and ambient temperature). Each of the distribution shifts may be explained by a combination of the temporal and spatial factors applied as well as the relationship between metrics:

- **Temperature:** Warmer temperatures are more prominent in the weighted distribution largely due to the temporal adjustments. Temperature values during nighttime hours are typically cooler than values during daytime hours. As a result, the emphasis on daytime conditions caused by VMT-TOD weighting increases the prevalence of warmer temperature values. Spatial weighting also influences warmer temperatures to a lesser degree, as VMT disproportionately driven at lower elevations near moderate latitudes. The jagged behavior in both profiles is an artifact of the underlying TMY3 data, which was measured in Fahrenheit, rounded, then converted to Celsius. The unrounded data were not available, and the raw data were utilized as-is. While smoothing the data may be desirable depending on application, the authors elected to illustrate the raw and weighted data sets as-is for repeatability.

- **Relative Humidity:** An increased frequency of lower relative humidity values is also driven primarily by temporal adjustments. Cooler air can hold a smaller amount of water vapor; as the air warms, the relative humidity is lower for the same amount of absolute vapor present. Thus, emphasizing the daytime conditions when VMT takes place increases the share of lower relative humidity conditions.
- **Solar Irradiance:** Larger solar irradiance values are more prevalent in the data set due to the temporal weighting of daytime hours. The share of irradiance values of zero (associated with overnight conditions) is nearly halved after TOD weights are applied. Spatial weighting reduces large, outlier values occurring near minimal amounts of VMT (such as in rural Arizona and Nevada), but this impact is overwhelmed by the temporal adjustments.
- **Air Density**: The share of moderate air density values is increased largely due to the spatial weighting applied and the elevation of typical cities. The raw data contain many weather stations at higher elevations, producing a large amount of lower air density values. When spatial weighting is applied in accordance to nearby VMT, high-altitude stations with low air densities are deemphasized in comparison to stations located near coastal cities. Note that the share of high air density is also reduced. This is due to the TOD weighting, which emphasizes high temperatures that occur during daytime hours.

Inspection of the VMT-TOD and VMT-GEO weighted ambient condition distributions reveals the importance of considering how vehicles are used both temporally and spatially when understanding the ambient conditions experienced. Using raw ambient condition data without weighting is shown to produce disproportionately large shares of cold temperatures, high relative humidity conditions, low solar irradiance values, and extreme air densities. The following section will use the ambient condition distributions developed to inform representative energy consumption calculations at the local and national levels.

Representative Evaluation Points: Ambient Condition Binning, Weighting, and Localization

The national ambient condition distributions reflect the prevalence of ambient conditions that influence vehicle performance, such as increased energy consumption for electric vehicles during cold temperatures due to increased heating load or increased cabin cooling due to high incident solar irradiance . This section introduces a framework for distilling the national-level ambient condition distributions into representative operating points that may be localized to any urban or rural area of interest while minimizing the number of tests or simulations that would be required to understand vehicle performance. Key steps include binning ambient conditions (such as cold, moderate, hot), identifying representative values for each bin, and finally quantifying the relative prevalence of these conditions. The methodology will be introduced first for a single ambient condition variable of interest. This will then be expanded for multiple variables. This section will conclude with a description of how results may be localized without requiring additional simulations and a procedure for quantifying local coverage specific to each area.

Single Variable Binning

The ambient condition binning and weighting methodology for estimating local operating conditions and corresponding weights can be applied to any of the ambient conditions but will first be introduced assuming a single variable of interest: temperature. Recall the national ambient temperature distribution shown in Figure 10, which incorporates TOD and GEO-VMT weighting (when and where) LDVs are driven.

The distribution describes the full range of temperature values an LDV may experience throughout the United States, from extremely cold conditions to extremely hot conditions, by share of VMT. An understanding of how frequently vehicles experience certain conditions may be inferred at the fleet level through inspection of the temperature distribution. For instance, the percentage of driving that takes place during certain conditions may be calculated using Equation 7, which is a simple numerical summation between threshold values. Ambient condition ranges, defined by a minimum and maximum value, are referred to as bins.

$$w_n = \sum_{T_l}^{T_u} f(T) \qquad \text{Equation 7}$$

Where w_n = weight associated with bin n
$f(T)$ = weighted frequency distribution for ambient temperature
and T_l, T_u = lower and upper temperature values defining bin n

Equation 7 may be applied to multiple bins to obtain an understanding of how frequently certain ambient conditions occur during LDV operation. For instance, application of Equation 7 reveals that extremely cold conditions, defined as between -30°C and -10°C, are present during 1.4% of VMT while extremely hot conditions, defined as between +30°C and +50°C, are present during 8.4% of VMT. (Note that the number of bins and specific bin limits are flexible. Equation 7 will output accurate weights for any bin limits describing ambient conditions of interest.) Moreover, the underlying frequency distribution, $f(T)$, enables identifying exemplar values to represent each bin; the 50th percentile of each bin is identified as the best value and is referred to as the bin evaluation point. The bin weights (Equation 7) in combination with the bin evaluation points may be used for calculating aggregate vehicle performance across a wide variety of ambient conditions. The aggregate vehicle performance is calculated by summing the vehicle performance at each evaluation point multiplied by the corresponding bin weight (Equation 8).

$$\eta_r = \sum_{i=1}^{|T_{bins}|} \eta_i * w_i \qquad \text{Equation 8}$$

Where η_r = representative vehicle performance across all temperatures
η_i = vehicle performance for bin i
and $|T_{bins}|$ = set of all temperature bins

An example of this process follows, which uses the binning and weighting methodology introduced for calculation of representative vehicle energy consumption. The example demonstrates how a single aggregate value may be obtained using an example selection set of bins.

Representative operating points at which energy consumption can be evaluated are obtained by first discretizing the national ambient condition distribution for a variable into bins. In this example, temperature is divided into four bins to describe conditions for extremely cold (-30°C to -10°C), moderately cold (-10°C to 10°C), moderately warm (10°C to 30°C), and extremely warm conditions (+30°C to +50°C). For each bin, the 50th percentile value is identified to represent the conditions contained within the bin boundaries. These 50th percentile values are used as evaluation points, influencing simulated or measured energy consumption at that condition. Figure 11 shows the national weighted temperature distribution with sample bins and the corresponding 50th percentile evaluation points labeled for each bin.

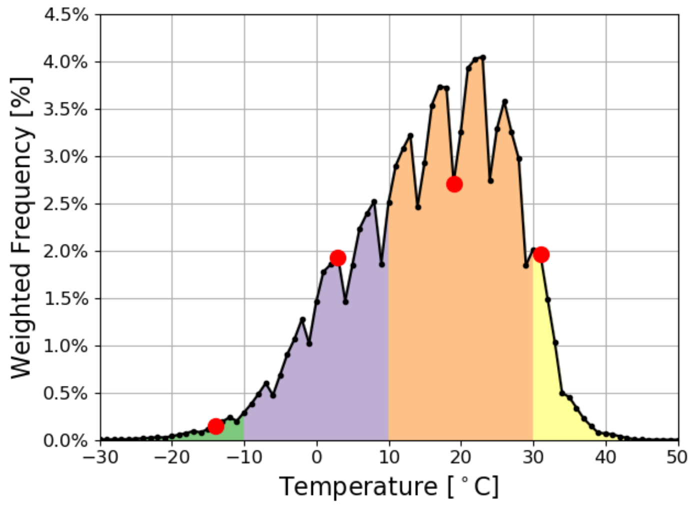

Figure 11. National ambient temperature distribution with representative evaluation points: discretization of the nationally aggregated temperature distribution using four equidistant bins. Red points indicate the 50th percentile temperature value, or evaluation point, for each bin.

In addition to identifying representative operating points, weights are calculated for each bin by numerically summing all qualifying frequencies in the national distribution (Equation 7). These bin weights are the relative share of total LDV VMT within each bin boundary. To evaluate aggregate energy consumption, energy use at each operating point is determined through simulation or testing. An aggregate energy consumption value is then obtained by calculating the energy consumption associated with each bin evaluation point and applying the corresponding bin weight (Equation 8).

In this example, the aggregate representative energy consumption is largely driven by the moderately cold and moderately hot bins, which contain shares of 26.5% and 63.7% of LDV VMT, respectively, at the national level. The Ambient Condition Localization section below will expand upon this example using four temperature bins, exploring how bin weights may change considerably from location to location.

Finally, note that the selection of four evenly spaced bins is for example purposes. The quantity of bins may be increased for technologies that are highly sensitive to temperature or reduced for technologies with minimal sensitivity to temperature. At the extreme, evaluation points for every single temperature value between -30°C and +50°C may be prescribed and the corresponding weights calculated accordingly. While a higher quantity of bins is likely to produce more accurate representative energy consumption, this would be at the expense of more numerous tests or simulations. In addition, variably sized bins may also be created for representative evaluation of technologies with highly variable performance, such as a technology with a step-change control strategy response to temperature. Engineering judgement and familiarity with relevant ambient conditions for a given technology are necessary for determining an appropriate number of ambient condition bins and whether bins should be variably sized.

Multivariable Binning

The binning and weighting methodology introduced in the previous section can also be expanded to consider multiple ambient conditions simultaneously, as the TMY3 data set reports values for ambient temperature, solar irradiance, relative humidity, and air density for each hour. Equation 9, expanded from Equation 7, provides weights that correspond with multiple ambient conditions simultaneously. Lower and upper bounds must be defined for each ambient condition metric being considered when calculating bin weights. Calculation of the bin weight in Equation 9 requires lower and upper values for both temperature values and solar irradiance values. In addition, evaluation points now consist of multiple ambient conditions simultaneously, such as a hot temperature value and a large evaluation solar irradiance value. Just as before, the 50th percentile is used for each metric when identifying evaluation points and vehicle performance is aggregated using bin weights. While only two metrics are shown in the equation—solar irradiance and ambient temperature—the expression can be expanded to consider all four ambient condition metrics simultaneously. The energy consumption example from the Single Variable Binning section is continued below to further illustrate how to use binning and weighting process for determining representative vehicle performance when considering multiple ambient conditions simultaneously.

$$w_i = \sum_{S_l}^{S_u} \sum_{T_l}^{T_u} f(T, S) \qquad \text{Equation 9}$$

Where w_i = weight associated with bin i

$f(T, S)$ = weighted frequency distribution considering ambient temperature and solar irradiance

and S_l, S_u = lower and upper solar irradiance values defining bin n

The energy consumption associated with a given technology may be sensitive to multiple ambient conditions simultaneously. Consider a technology that is highly sensitive to both temperature and solar irradiance. Prescribed operating points for temperature and solar irradiance are needed simultaneously with corresponding weights. Like before, the temperature and solar domains are discretized into bins with 50th percentile operating points identified. The resulting bins are now defined with bounds on both temperature and solar irradiance (16 total, depicted in Figure 12). Figure 12 also contains the weights for each bin.

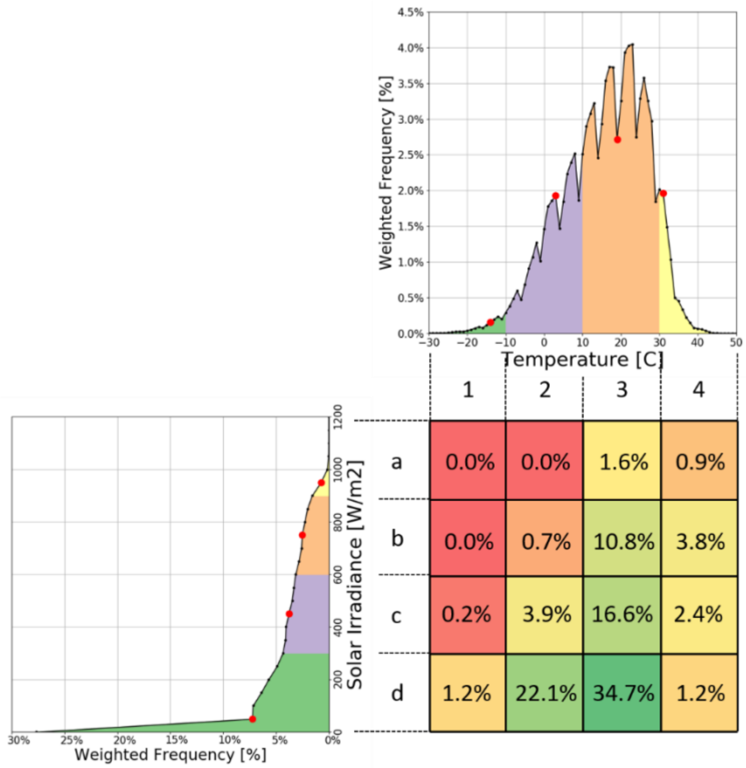

Figure 12. Example discretization of the nationally aggregated ambient condition distribution using four equidistant domains for temperature and solar irradiance simultaneously. The resulting permutation consists of 16 bins with corresponding weights.

Aggregate vehicle energy consumption in response to both temperature and solar irradiance is evaluated similarly by calculating the energy consumption for each of the 16 bins and applying bin weights. Note the prevalence of bins with very small weights (or in some cases, a weight of zero). These bins correspond with temperature and solar conditions that rarely occur simultaneously, such as high solar load and low temperature or high temperature and low solar load. If resources are available for only a limited number of simulations or tests, bins describing a high percentage of experienced conditions may be prioritized while bins describing minimal shares of experienced conditions may be excluded. The VMT coverage section below will elaborate on bin prioritization and the consequences of excluding certain bins when calculating aggregate vehicle performance.

Ambient Condition Localization

This process for nationally representative ambient condition bins and weights can be extended further to provide insight at the local level where results may vary considerably from area to area. The national ambient condition aggregations are a function of 1,010 stations; however, a typical vehicle will be predominately driven within a single region. For example, while the national ambient condition distributions are influenced by both Anchorage, Alaska, and Miami, Florida, the aggregate results do not describe typical conditions experienced by vehicles driven in either of these cities. Thus, additional analysis is introduced to enable localizing vehicle performance results without performing additional simulations or tests. Recall that Equation 7 and Equation 9 utilize the *national* ambient condition distributions. Localizing results is performed by using the *local* ambient condition distributions for regions of interest. Bin weights specific to each area are calculated using only TMY3 stations nearby. In the cases of Miami and Anchorage, bin weights may be calculated using only weather data that correspond to these two cities. Finally, it is important to note that this process only involves the calculation of area-specific weights. The bin limits and evaluation points previously defined at the national level are maintained. Thus, aggregate results calculated at the national level may be localized to any region by simply re-weighting vehicle performance results obtained in response to the ambient condition evaluation points. The energy consumption application of the binning and weighting process will be continued for illustrative purposes.

Recall the temperature bins and corresponding evaluation points introduced in the Single Variable Binning section. The weights associated with these bins, originally calculated using nationally aggregated ambient conditions, may also be evaluated using only local conditions. Consider the two ambient temperature distributions shown in Figure 13, describing conditions for both Anchorage and Miami. The distributions are intuitive: lower temperatures are more frequent in the Anchorage case relative to the national distribution (included in gray for comparison) while warmer temperatures are more frequent in Miami. The evaluation points, defined as the 50th percentile on the national temperature distribution, are also shown on each of the local distributions.

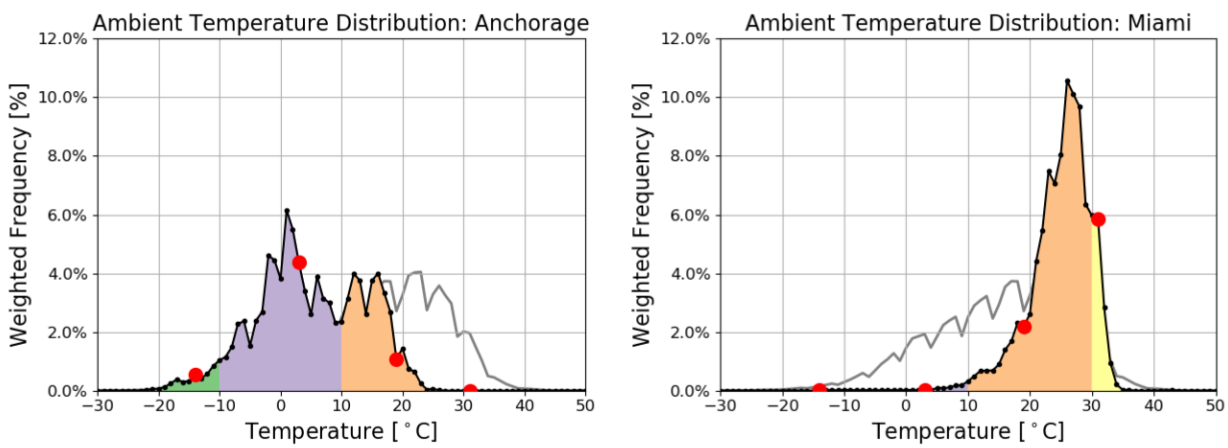

Figure 13. Discretization of the nationally aggregated temperature distribution using four equidistant bins. Red points are included describing the 50th percentile temperature value from the national-level analysis, or operating point, for each bin.

Table 1 quantifies the weights for the four temperature bins at the national level, Anchorage, and Miami (calculated using Equation 7). These local weights are applied the same operating points and thus enable calculating locally specific performance, such as energy consumption, without performing additional tests or simulations. This will localize representative energy consumption values to Anchorage, Miami, or any other area of interest for which ambient condition data are available without requiring any additional energy consumption evaluations.

Table 1. Example weight comparison at the national level, Anchorage, and Miami.

	Bin Limits	Bin Evaluation Point	National Weight	Anchorage Weight	Miami Weight
Extremely Cold	-30°C, -10°C	-14°C	1.4%	3.9%	0.1%

	Bin Limits	Bin Evaluation Point	National Weight	Anchorage Weight	Miami Weight
Moderately Cold	-10°C, 10°C	+3°C	26.5%	62.2%	0.9%
Moderately Warm	10°C, 30°C	+19°C	63.7%	33.9%	83.1%
Extremely Warm	30°C, 50°C	+31°C	8.4%	0.0%	15.9%

This process of calculating weights for a single variable, local ambient temperature, can also be extended for multiple variables simultaneously. Just as locally specific weights were established for Anchorage and Miami for only temperature, locally specific weights could be also calculated for the multivariate example involving ambient temperature and solar irradiance simultaneously (Figure 12). Localized vehicle performance calculations could also be performed with bins defined by all ambient conditions discussed in this paper (temperature, solar irradiance, humidity, and air density).

Evaluating VMT Coverage—Nationally and by Area

In both the single variable and multivariable examples, the amount of VMT covered in each bin (the bin weight) varies highly. In cases with a high quantity of bins and limited resources to perform tests or simulations, bins may be prioritized in accordance to their amount of VMT coverage. For instance, consider the bin weights produced in Figure 12, considering four temperature and four solar setpoints (16 bins). Excluding the five lowest bins reduces the number of required simulations to 11 while still covering 99.1% of VMT at the national level. Exclusion of the combined 0.9% of VMT coverage across bins 1a, 1b, 2a, and 4a would slightly reduce the accuracy of the aggregate result but also result in over a 30% reduction in the requisite simulations or tests (from 16 to 11).

Figure 14 explicitly visualizes the VMT coverage tradeoffs associated with bin selection. The 16 bins, defined by both ambient temperature and solar irradiance thresholds, are ranked on the x-axis in descending order of VMT coverage. Discrete and aggregate VMT coverage are plotted, illustrating the diminishing returns associated with evaluating vehicle performance for each bin. Use of the figure provides a means for prescribing a desired amount of VMT coverage nationally and identifying the fewest number of bins necessary to achieve requisite cumulative coverage.

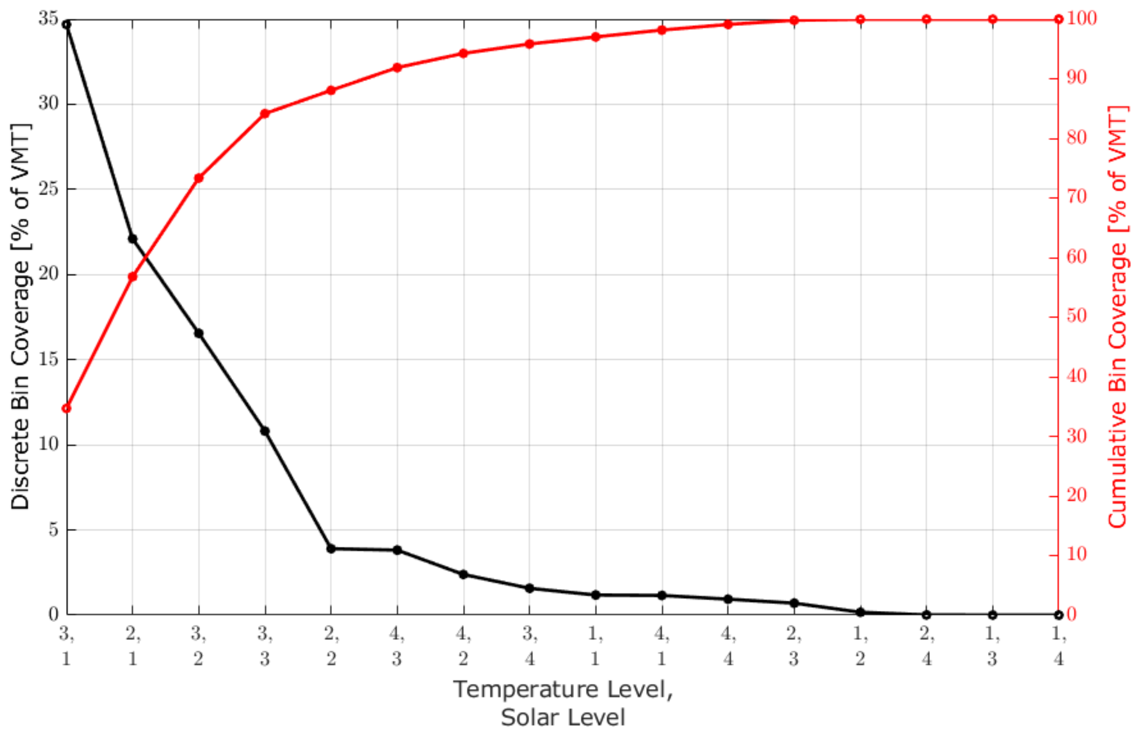

Figure 14. Discrete and cumulative VMT coverage for the 16 bins defined by temperature and solar irradiance.

It is important to note, however, that the importance of a given ambient condition combination toward producing a representative energy consumption value is not synonymous with the amount of VMT coverage the combination is associated with. While extreme conditions, such as very cold temperatures or very high solar irradiance may be infrequent, these conditions may disproportionately contribute to energy consumption penalties. Bins should only be excluded when engineering knowledge informs the lack of relevance between certain setpoints and technology. In cases where bins relevant to energy consumption are excluded, implications surrounding the reduced accuracy of the results should be understood. Removing ambient condition permutations from evaluation should be done after deliberation and with an understanding of the associating error introduced.

Finally, note that reductions in VMT coverage associated with the exclusion of ambient condition bins varies by location. For instance, consider the bin weights presented in Table 1 outlining the prevalence of extremely cold, moderately cold, moderately warm, and extremely warm conditions at the national level as well as for Anchorage and Miami. If resources are only available for 3 simulations, the "extremely cold" bin is a good candidate to exclude given the minor amount of VMT coverage associated with the condition nationally (1.4%). Exclusion of the "extremely cold" bin from the set of four temperature bins still produces 98.6% VMT coverage nationally versus 96.1% in Anchorage and 99.9% in Miami. The impact of VMT coverage reduction associated with the exclusion of the "extremely cold" bin is further illustrated in Figure 15, showing the percent coverage for all 504 areas considered in this analysis sorted in order of decreasing coverage. While aggregate energy consumption calculations excluding the "extremely cold" bin may be accurate nationally, localized to Miami, and perhaps localized to Anchorage, an aggregate value would be highly inaccurate for Fairbanks, Alaska, for which only 68% of local VMT is covered. Users of this methodology must be sensitive to these considerations when defining bins and producing aggregate vehicle performance conclusions.

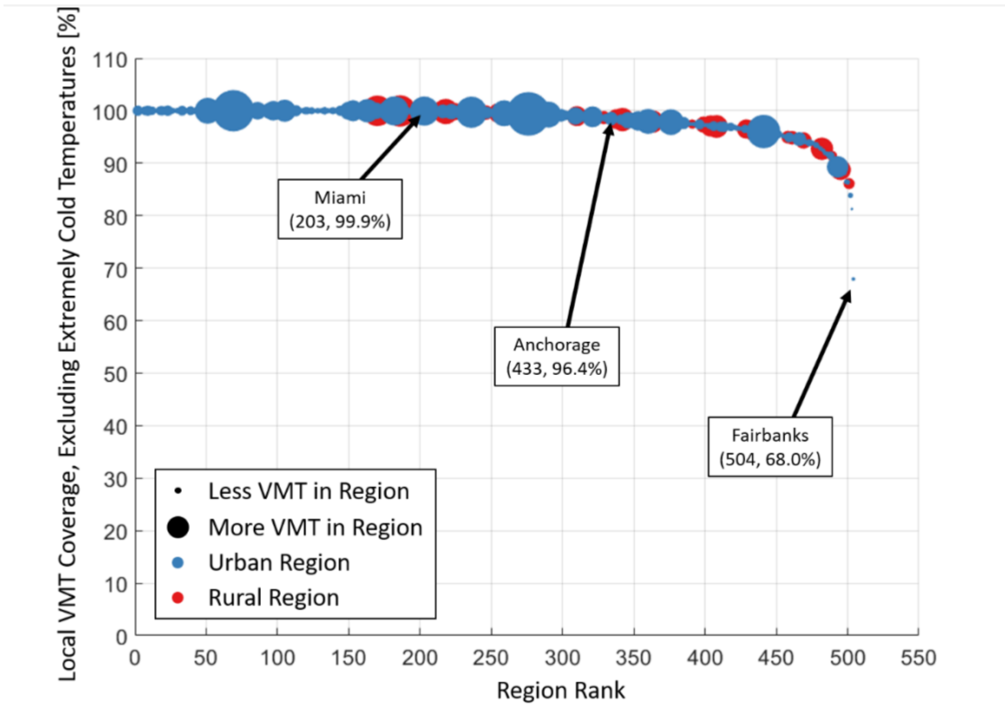

Figure 15. Area-level coverage impacts associated with excluding extremely cold temperatures (<-10°C) from consideration.

Future Work, Limitations

Limitations of this process arise from assumptions surrounding bin selection and the use of the 50th percentile as setpoints representing entire bins. The accuracy of energy consumption results calculated using the binning methodology is highly reliant upon both the number of bins specified in each ambient condition dimension as well as the number of bins maintained. In addition, assumptions surrounding the 50th percentile being the most representative value for use as a setpoint may vary depending on the vehicle or technology being evaluated and the size of the bin being represented. A small number of bins reduces accuracy in two ways: fewer bins result in a courser definition of experienced conditions (such as extremely cold and moderately cold being generalized into a bin for all cold conditions) and setpoints become less representative as bin thresholds widen. A comprehensive validation exercise is planned to understand the tradeoffs associated with bin selection, the accuracy of bin setpoints for representing all conditions within a bin, and finally the accuracy of localized energy consumption results. Although quantitative analyses describing tradeoffs associated with bin selection and bin setpoints are pending, the binning methodology does provide an effective means of identifying important ambient conditions nationally and specific to each area by incorporating when and where vehicles are driven.

An additional limitation is the exclusion of driving profiles during which vehicle performance may be evaluated. Explicitly capturing driving effects was considered to be outside the scope of this study; instead, the ambient condition characterization methodology introduced should accompany a more comprehensive set of representative drive cycles. An accompanying evaluation by the National Renewable Energy Laboratory exploring representative drive cycles is pending, featuring a similar emphasis on reducing the number of simulations or tests required while maintaining the ability to produce nationally representative results that may be localized to any area of interest.

Finally, the authors acknowledge that ambient conditions also affect vehicle performance while vehicles are not being driven, primarily by influencing the characteristics of initial thermal state. Vehicles in regions with ambient conditions consisting disproportionately of low temperatures and solar irradiance may experience a higher frequency of cold starts versus vehicles in warmer areas. A similar methodology is needed for characterizing ambient conditions between driving events to better represent the full impact ambient conditions pose toward vehicle energy consumption.

Conclusions

The analysis introduced within this paper provides a framework for quantifying ambient conditions experienced by LDVs. Open-source TMY3 meteorological data are adjusted to properly account for when vehicles are driven by comparing against survey data and where vehicles are driven by incorporating HPMS information as well as vehicle registrations. The properly weighted ambient condition data are then aggregated into four nationally representative distributions (temperature, solar irradiance, relative humidity, air density), which provide an understanding of how frequently LDVs experience certain conditions (such as extremely cold or extremely hot temperatures).

The ambient condition distributions are a foundation for the binning and weighting methodology, which enables calculation of representative vehicle performance across a wide range of ambient conditions. The methodology is highly flexible, informing conditions at regional levels and nationwide. In addition, the procedure is valuable in its use of exclusively open-source data sets available to all researchers and agnostic in its application to LDVs. While understanding conditions relating to LDV use was the emphasis for this paper, a similar temporal and geographic weighting could be applied to the underlying TMY3 weather data for separate applications, such as heavy-duty trucking or even other industries entirely.

Acknowledgment

This work was supported by the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308 with Alliance for Sustainable Energy, LLC, the operator of the National Renewable Energy Laboratory, and Hyundai America Technical Center, Inc. under Agreement No. TSA-18-01063. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Abbreviations

EPA	U.S. Environmental Protection Agency
GEO	geographic
GHG	greenhouse gas

HPMS	Highway Performance Monitoring System
LDV	light-duty vehicle
NHTS	National Household Transportation Survey
TMY3	Typical Meteorological Year 3
TOD	time of day
VMT	vehicle miles traveled

Nomenclature

VR_i	vehicle registrations in ZIP code i
VR_A	total registrations within area A
VMT_A	total VMT for area A
VMT_i	VMT in ZIP code i
VMT_{TMY_N}	VMT attributed to TMY3 station N
TOD_h	the time of day weight applied to ambient weather data occurring on hour h of the day
$VMT_{TOD\,h}$	the total VMT occurring during a given hour as calculated from the NHTS
GEO_{TMY_N}	GEO weight associated with all ambient conditions reported by TMY3 station N
$w_{j,N}$	the hour-level weight associated with an ambient condition at hour of day *j* and reported by TMY3 station *N*
TOD_j	VMT-TOD weighting, or the percentage of VMT occurring at hour of day j
w_i	weight associated with bin i
$f(T)$	weighted frequency distribution for ambient temperature
T_l, T_u	lower and upper temperature values defining bin i
η_r	representative vehicle performance across all temperatures
η_i	vehicle performance for bin i
$\lvert T_{bins} \rvert$	set of all temperature bins
$f(T,S)$	weighted frequency distribution considering ambient temperature and solar irradiance
S_l, S_u	lower and upper solar irradiance values defining bin i

References

[1] United States Environmental Protection Agency, "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2017," 2019.

[2] United States Environmental Protection Agency, "National Emissions Inventory Report," 2014.

[3] United States Environmental Protection Agency, "The 2018 EPA Automotive Trends Report," 2019.

[4] Y. K. Lee, J. I. Park and J. H. Lee, "Analysis of the Effect of Cold Start on Fuel Economy of Gasoline Automatic Transmission Vehicle," *International Journal of Automotive Technology,* vol. 15, pp. 709-714, 2014.

[5] F. Jehlik and E. Rask, "Development of Variable Temprature Brake Specific Fuel Consumption Engine Maps," *SAE International,* 2010.

[6] J. Jaguemont, L. Boulon and Y. Dube, "A Comprehensive Review of Lithium-Ion Batteries Used in Hybrid and Electric Vehicles at Cold Temperatures," *Applied Energy,* vol. 164, pp. 99-114, 2016.

[7] L. Horrein, A. Bouscayrol, W. Lhomme and C. Depature, "Impact of Heating System on the Range of an Electric Vehicle," *IEEE Transactions on Vehicular Technology,* vol. 66, no. 6, pp. 4668 - 4677, 2017.

[8] J. P. Rugh, C. Kreutzer, B. Kekelia, G. Titov and J. Lustbader, "U.S. Light-Duty Vehicle Air Conditioning Fuel Use and Impact of Solar/Thermal Control Technologies," *SAE International Journal of Passenger Cars - Mechanical Systems,* 2019.

[9] United States Environmental Protection Agency, "Fuel Economy Labeling of Motor Vehicles: Revisoins to Improve Calculation of Fuel Economy Estimates; Final Rule," 2006.

[10] X. Wang, G. Yunshan, L. Yu and X. Feng, "Effects of altitude on the thermal efficiency of a heavy-duty diesel engine," *Energy,* vol. 59, pp. 543-548, 2013.

[11] R. Farrington and J. Rugh, "Impact of Vehicle Air-Conditioning on Fuel Economy, Tailpipe Emissions, and Electric Vehicle Range," in *Earth Technologies Forum*, Washington, DC, 2000.

[12] J. Rugh and R. Farrington, "Vehicle Ancillary Load Reduction Project Close-Out Report," National Renewable Energy Laboratory, 2008.

[13] United States Environmental Protection Agency, National Highway Traffic Safety Administration, "Final Rulemaking for 2017-2025 Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards," 2012.

[14] T. Yuksel and J. J. Michalek, "Effects of Regional Temperature on Electric Vehicle Efficiency, Range, and Emissions in the United States," *Environmental Science & Technology,* vol. 49, pp. 3974-3980, 2015.

[15] National Renewable Energy Laboratory, "National Solar Radiation Data Base: 1991-2005 Update: Typical Meteorological Year 3," [Online]. Available: https://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/tmy3/. [Accessed 1 November 2018].

[16] J. McNerney, Z. A. Needell, M. T. Change, M. Miotti and J. E. Tancik, "TripEnergy: Estimating Personal Vehicle Energy Consumption Given Limited Travel Survey Data," *Transportation Research Record,* pp. 58-66, 2017.

[17] Z. A. Needell, J. McNerney, M. T. Chang and J. E. Trancik, "Potential for Widespread Electrification of Personal Vehicle Travel in the United States," *Nature Energy,* 2016.

[18] F. Jehlik, N. Chevers, M. Moniot, Y. Song, H. Hirabayashi, M. Nomura and E. Wood, "Determining Off-cycle Fuel Economy Benefits of 2-Layer HVAC Technology," *SAE International,* 2018.

[19] E. Wood, J. Gonder, S. Lopp and F. Jehlik, "Simulated Real-World Energy Impacts of a Thermally Sensitive Powertrain Considering Viscous Losses and Enrichment," *SAE Inernational,* 2015.

[20] B. Borlaug, J. Holden and J. Lustbader, "A Method for Selecting Representative Light-Duty Drive Cycles from Real-World GPS Data," *To be Submitted for Publication.*

[21] United States Department of Transportation, "2017 National Household Travel Survey," 2017. [Online]. Available: https://nhts.ornl.gov/downloads. [Accessed 1 November 2018].

[22] United States Department of Transportation, "Highway Statistics 2016: Miles and Daily Vehicle-Miles Traveled," [Online]. Available: https://www.fhwa.dot.gov/policyinformation/statistics/2016/hm71.cfm. [Accessed 1 November 2018].

[23] IHS Markit, "Vehicle Market Analysis: Registrations and Vehicles-in-Operation," 2017. [Online]. Available: https://www.ihs.com/products/automotive-market-dataanalysis.html. [Accessed 1 November 2018].