TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2024, Vol. 2678(1) 707–720

DOI: 10.1177/03611981231172502
journals.sagepub.com/home/trr
Sage

Comprehensive Energy Modeling Framework for Multi-Powertrain Bus Transit Systems

Mahsa Arabi[1] **and Jimi Oke**[1]

Abstract

With the increasing emphasis on sustainable transportation solutions, there is an urgent need for a system that can accurately assess vehicle-specific energy consumption in real-time. However, the current modeling frameworks have a significant gap in that they are not demand- and congestion-responsive, nor are they capable of modeling the system-wide energy consumption of bus fleets at the vehicle-level, encompassing all three types of powertrains—conventional, hybrid, and electric. To address this gap, we present a comprehensive modeling framework that takes into account passenger loading, considers the dynamics of the system including acceleration, and is capable of predicting the real-time energy usage of each bus in a transit network. Using the Pioneer Valley Transit Authority system as a case study, we calibrated existing powertrain-specific models for conventional diesel, hybrid, and battery electric buses. Our demand- and congestion-responsive energy consumption modeling framework fills this critical gap and offers significant potential for future sustainable transportation planning. Relying on detailed trajectories computed from one month of bus location data, along with passenger counts and road grade, our framework is a novel contribution in that it provides spatiotemporal energy predictions for an entire bus system. As it is readily calibrated with potential improvements in accuracy given more data, this framework can be applied by transit agencies for the purposes of real-time energy consumption monitoring, fleet electrification planning, equity analyses, and carbon emission reduction strategies.

Keywords

planning and analysis, decision-making, performance models, planning data analysis, public transportation

In 2021, the transportation sector was responsible for 28% of the total energy consumption in the U.S. (*1*). Although public transportation accounted only for 6% of all trips taken across the country, the energy intensity of buses (0.11 kWhpkm) is significantly lower than that of cars (0.22 kWhpkm) (*2*, *3*). This indicates that public transportation can significantly reduce transportation-related energy consumption and carbon emissions by providing an alternative to individual car use. Thus, public transportation has an important role in climate change mitigation. Given the increasing urgency of the climate situation, there is a need for readily deployable energy models to facilitate decision-making for sustainability in bus transit networks.

Hybrid electric bus (HEB) and battery electric bus (BEB) are low- and zero-emission vehicles. By adopting HEB and BEB into public transportation fleets, we have the potential to significantly reduce the transportation sector's overall emissions and move toward a more sustainable future (*4*). Furthermore, low- and zero-emissions vehicles are cheaper to run, as the annual energy costs of HEB and BEB are 34% and 73% less than that of conventional diesel bus (CDB) (*5*). Hybrid and electric buses that are actively operating at transit agencies across the U.S. account for 22% and 2% of the total buses, respectively (*6*).

Currently, transit agencies do not have any system to accurately track the per-vehicle energy consumption of their fleet on a daily or hourly basis. This highlights the need for energy models to analyze system-wide energy

[1]Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Watertown, MA

Corresponding Author:
Mahsa Arabi, marabi@umass.edu

consumption. Vehicle-specific models are essential for decision-making, especially under constraints on budget and time. For instance, an energy consumption model can aid in identifying the most effective investment strategies and prioritizing routes for electrification. These models can be applied to evaluate the potential impacts of electrification on energy usage and emissions. Furthermore, energy consumption models can help policymakers to predict future energy consumption patterns at the vehicle level, assess the economic viability of different electrification strategies, and estimate the potential environmental outcomes of these strategies. Given these factors, an energy consumption model can play an instrumental role in promoting sustainable transportation and is an initial step toward helping policymakers design effective strategies for reducing greenhouse gas emissions.

Related Work

The modeling of energy consumption in public transportation systems has been widely studied, and several methods have been explored to develop models to estimate energy consumption by taking into account a range of factors and parameters (*7–15*). The earliest framework to model fuel consumption and emissions in heavy-duty diesel vehicles was introduced in 2004. The model estimates fuel rates based on variables such as engine power, friction factor, speed, displacement, and engine efficiency (*16*). This model has been applied to a wide range of fleets, including light- and medium-duty vehicles. An enhancement was made to the previous model by removing an unrealistic condition under which buses have to either accelerate at full throttle or brake at full braking to minimize their fuel consumption levels (*17*). In another improvement to the physics-based energy consumption model, a second-order polynomial energy consumption function was introduced (*18*). This model takes into consideration several factors including vehicle resistance force, rotational mass, gear ratio, the total mass of bus, acceleration, and speed as the dependent variables. In another study, a model was developed to estimate energy consumption by taking into account the distance traveled, route taken, and time (*19*). This model captures auxiliary system variables including heating, ventilation, air conditioning, and other electrical components in the energy consumption model. A more recent study focused on developing a system-wide energy consumption model for Massachusetts Bay Transportation Authority trains (*15*). This machine-learning-based model is capable of predicting energy consumption models using train movement and operation data, ridership, and ambient temperature data. Despite all these efforts, no approach has been proposed to provide accurate system-wide energy estimates based on real-time vehicle-specific models.

Table 1. Glossary of Abbreviations and Units

Abbreviation	Description
BEB	battery electric bus
CDB	conventional diesel bus
gal	gallon(s)
GPS	global positioning system
HEB	hybrid electric bus
kWh	kilowatt hour (unit of consumed energy)
mi	mile (unit of traveled distance)
MPG	miles per gallon (unit of fuel economy)
mph	miles per hour (unit of speed)
mpkWh	miles per kilowatt-hour (unit of energy economy)
m/s^2	meters per second (unit of acceleration)
PVTA	Pioneer Valley Transit Authority
VRIE	vehicle refueling interval energy (gal)

This paper fills the gap in the literature by developing a modeling framework that is demand- and congestion-responsive. The framework is capable of modeling the system-wide energy consumption of a bus fleet at the vehicle-level, including all three types of conventional, hybrid, and electric powertrains. We calibrate vehicle-specific energy models for various powertrains to arrive at accurate system-level and vehicle-level predictions. We also develop a systematic approach to build an energy model, which is not only highly interpretable and high-performing, but is also simply built based on vehicle-specific parameters and trajectories.

The results of this energy modeling framework can be considered as an initial step to build a decision support tool to help transit agencies in planning the electrification of their fleets. This model can be used to plan future public transport systems, reduce energy consumption, facilitate the transition toward more sustainable fleets with lower emissions, and make better policy decisions about fleet replacement.

As an aid to the reader, we summarize the abbreviations used in this paper in Table 1. The rest of the paper is organized as follows. We describe the case study area in the next section. Following this, we describe the methods used (data and models). Then we present and discuss the results and close with a conclusion.

Case Study

Bus Network and Fleet

The Pioneer Valley Transit Authority (PVTA) is the second-largest transit agency in Massachusetts, U.S. It served over 6 million passenger trips in the Pioneer Valley region (Hampden and Hampshire counties) in 2022. Figure 1 shows the location of bus stops in PVTA network. PVTA operates 178 buses on 45 fixed routes (*20*). Out of these buses, 88% are CDB, 5% are HEB, and 7% are BEB. There are 10 different types of bus engines in the network:

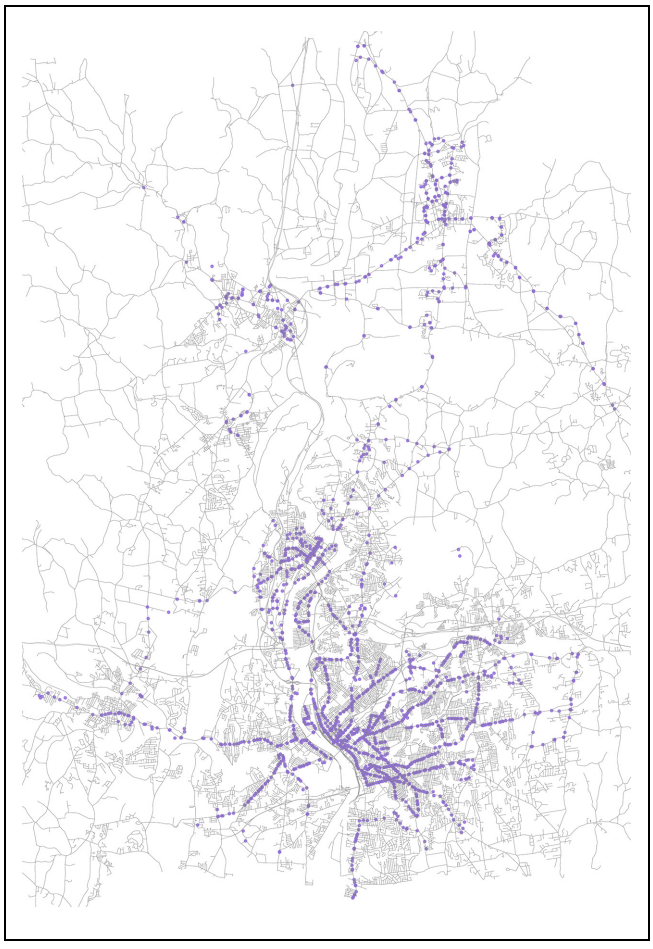

Figure 1. Pioneer Valley Transit Agency bus stop locations.

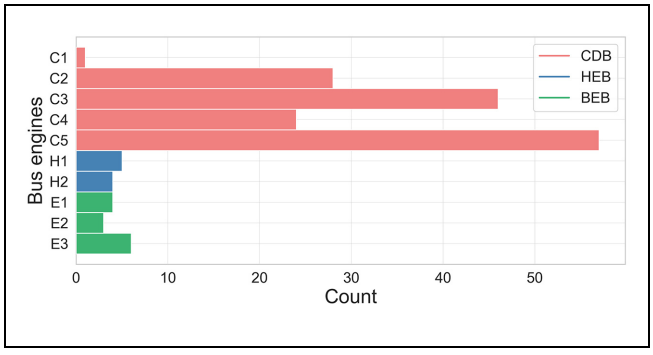

Figure 2. Histogram of vehicle engines in Pioneer Valley Transit Agency network.

five conventional ("C1"–"C5"), two hybrid ("H1" and "H2") and three electric ("E1"–"E3") powertrain engines. Figure 2 displays the histogram of PVTA bus engines.

Energy Data

To calibrate and validate the powertrain-specific energy models, we used timestamped refueling data indicating the amount of refueling tickets for each bus in April 2022. Refueling tickets are collected each time buses go for refueling to the garages. To better understand the frequency of refueling, we plotted a histogram of refueling intervals, as shown in Figure 3. This analysis revealed that refueling intervals are irregular, with an average interval of 34.8 h and a median interval of 24.8 h.

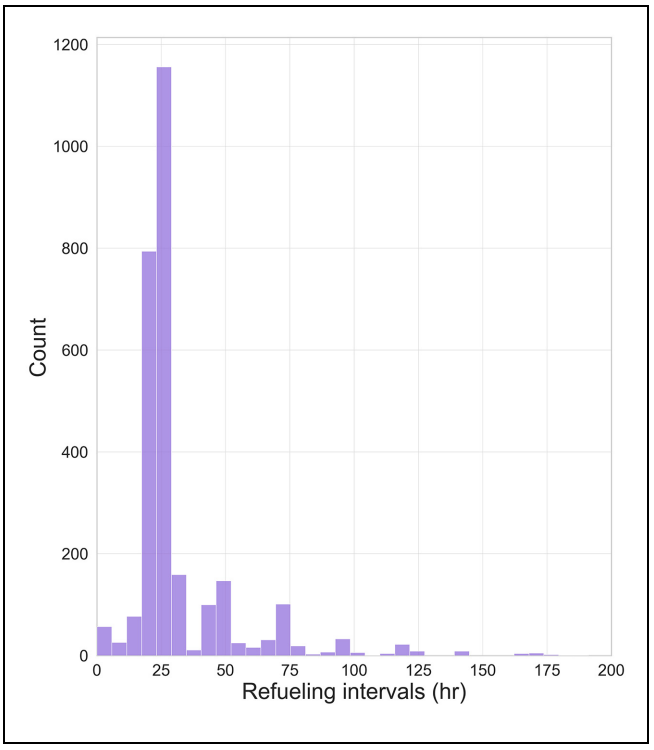

Figure 3. Histogram of refueling intervals.

For the purpose of validation, we define the vehicle refueling interval energy (VRIE) as $VRIE_{i,p,R_l}$ (gal) for vehicle of powertrain type p and refueling interval t_r. (See Equation 13 for more details on how it is computed.)

Figure 4 shows histograms of VRIE (gal) for CDB and HEB. The mean VRIE is about 30 gal for CDB and 26 gal for HEB.

Methods

Trajectory Computation

First, we cleaned the timestamped vehicle GPS location data by removing all data points with a time interval of less than 2 s. Each timestamp corresponds to the departure of a bus from a given stop. These spurious observations comprised about 3% of all the location timestamps in the dataset. Afterward, we computed trajectories (distance s_i, speed v_i, and acceleration a_i) of each bus i in the network at timestamp t_j, where j is the index of the timestamp. To calculate the distance, we used the Haversine distance function (*21*). The following equations were used:

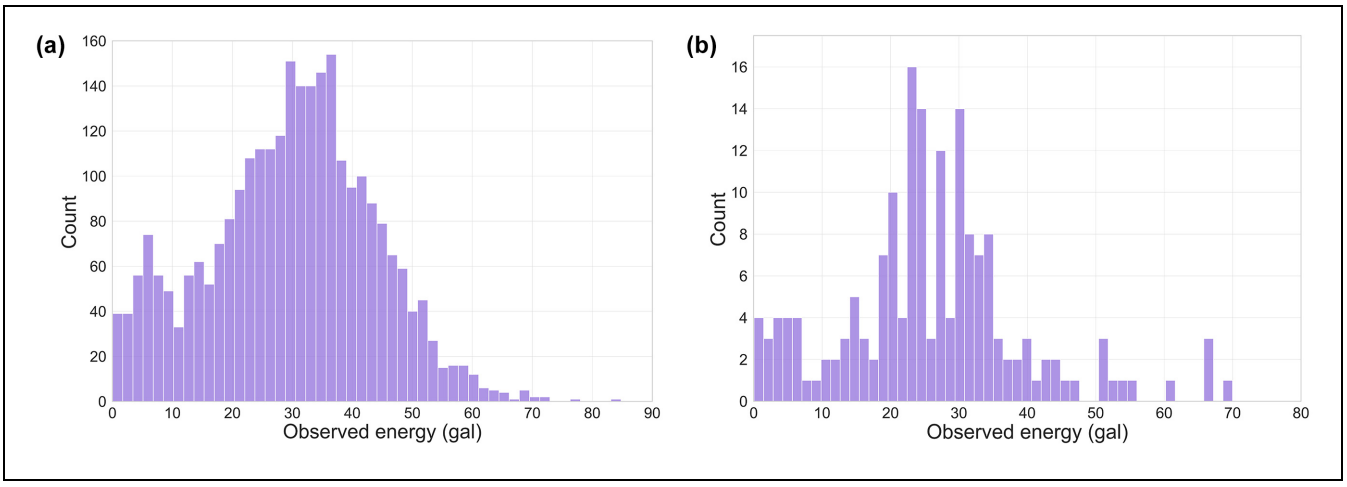

Figure 4. Histograms of vehicle refueling interval energy (gal) for: (*a*) conventional diesel bus and (*b*) hybrid electric bus.

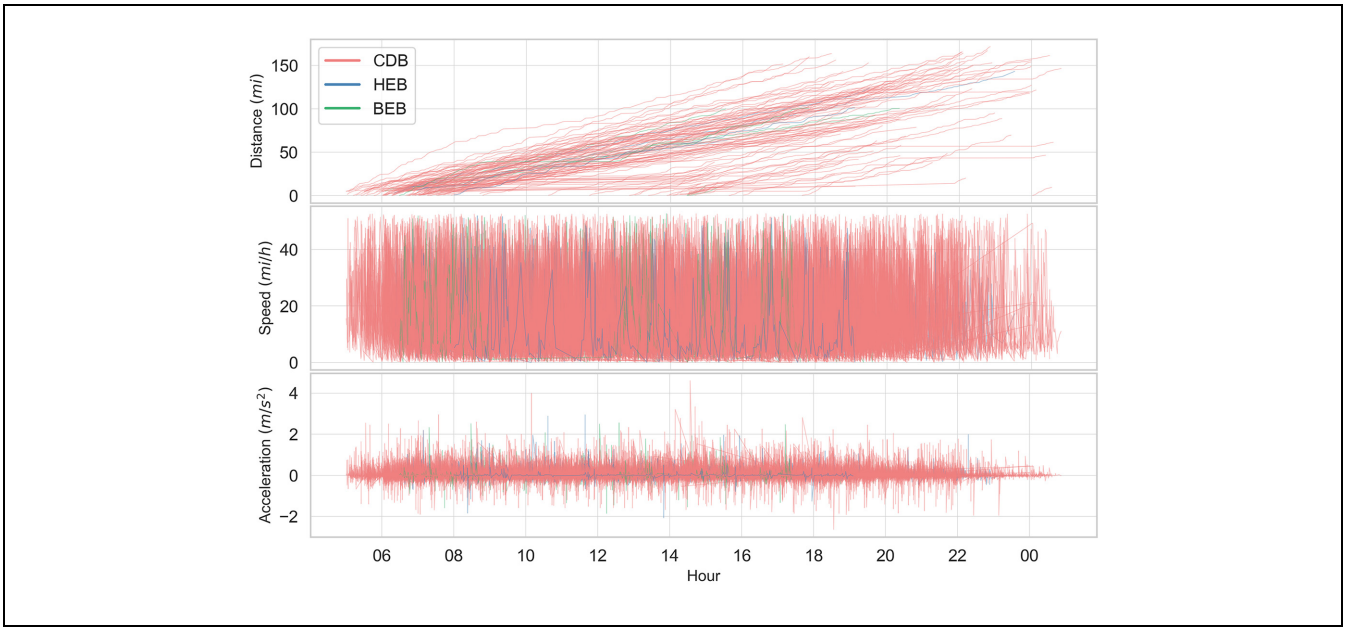

Figure 5. Time series of trajectories (cumulative distance [*top*], speed [*middle*], and acceleration [*bottom*]) for all 178 Pioneer Valley Transit Authority buses in the dataset over a 24 h period (5:00 a.m., April 4, 2022, to 5:00 a.m., April 5, 2022.
Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus.

$$s_i(t_j) = 2r \arcsin\left[\sin^2\left(\frac{\varphi_i(t_j)-\varphi_i(t_{j-1})}{2}\right) + \cos\left(\varphi_i(t_{j-1})\right)\cos\left(\varphi_i(t_j)\right)\sin^2\left(\frac{\lambda_i(t_j)-\lambda_i(t_{j-1})}{2}\right)\right]^{\frac{1}{2}} \quad (1)$$

$$v_i(t_j) = \frac{s_i(t_j) - s_i(t_{j-1})}{\Delta t_j} \quad (2)$$

$$a_i(t_j) = \frac{v_i(t_j) - v_i(t_{j-1})}{\Delta t_j} \quad (3)$$

$$\Delta t_j = t_j - t_{j-1} \quad (4)$$

where

r = the radius of the earth,

$\varphi_i(t_j)$ and $\varphi_i(t_{j-1})$ = the latitude of vehicle i at consecutive timestamps, and

$\lambda_i(t_j)$ and $\lambda_i(t_{j-1})$ = the longitude of vehicle i at consecutive timestamps.

Trajectories for a selected 24 h period are shown in Figure 5. Histograms of distances, speeds, acceleration, and time intervals are depicted in Figure 6. Summary statistics of trajectories for April, 2022, are shown in Table 2.

Energy Models

We adapted existing powertrain-specific physical energy models for CDB, HEB, and electric vehicles, which we

modified for BEB (*18*, *22*). Engine-specific parameters are shown in Table 3. All, except the age, are used in the powertrain models. The parameters common to all the energy models are summarized in Table 4. Model indices are provided in Table 5.

Table 2. Summary Statistics of Trajectories

Trajectory	Min	Median	Mean	Max
Cumulative distance (mi)	6.81	89.39	91.14	191.18
Speed (mph)	0		16.20	94.59
Acceleration (m/s^2)	−2.63	0	0.05	4.61

The instantaneous tractive power $P_{p,i}(t_j)$ output, in kW, of a vehicle i of powertrain type p at timestamp (t_j), is generally given by Wang and Rakha (*18*):

$$P_{p,i}(t_j) = \frac{1}{3600}\left[\frac{\rho}{25.92}C_D C_h A_{f,p} v_i^2(t_j) + \frac{m_i(t_j)gC_r}{1000}(c_1 v_i(t_j) + c_2) + 1.2m(t_j)a_i(t_j) + m_i(t_j)gG_i(t_j)\right]v_i(t_j) \quad (5)$$

where

- ρ = air density $\left(\mathrm{kg/m^3}\right)$,

Table 3. Engine Specifications

Engine	Mean age (years)	Length (ft)	Curb mass (lb)	Frontal area (m^2)	Powertrain
C1	17	30	25,700	6.99	CDB
C2	15	35	26,440	6.99	CDB
C3	15	40	27,180	6.99	CDB
C4	17	35	26,900	6.99	CDB
C5	17	40	28,250	6.99	CDB
H1	12	40	28,250	7.44	HEB
H2	10	60	41,500	7.44	HEB
E1	2	35	26,900	7.44	BEB
E2	2	40	32,770	7.44	BEB
E3	3	40	27,370	7.44	BEB

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus.

Table 4. Common Parameters for Powertrain Energy Models

Parameter	Description	CDB	HEB	BEB
α_2	quadratic fuel consumption coefficient		1.00×10^{-8}	na
η_{batt}	battery efficiency	na	na	0.95
$\eta_{d,p}$	driveline efficiency		0.94	0.92
η_m	motor efficiency	na	na	0.91
ρ	air density		1.23 kg/m ^{3}	
$A_{f,p}$	frontal area	6.99 m ^{2}	7.44 m ^{2}	7.44 m ^{2}
C_D	drag coefficient		0.78	
C_r	rolling coefficient		1.25	
C_1	rolling resistance coefficient 1		0.0328	
C_2	rolling resistance coefficient 2		4.575	
C_h	altitude correction factor		0.01 km	
d	engine size/displacement		8.71	
g	acceleration caused by gravity		9.81 m/s^2	

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus; na = not applicable.

Table 5. Model Indices and Mappings

Symbol	Description	Value
$i \in [1, N_p]$	index of bus	{ N_C: 156 (CDB), N_H: 9 (HEB), N_B: 13 (BEB)}
$p \in \{C, H, B\}$	powertrain type	{ C: CDB, H: HEB, B: BEB}
$j \in [1, T]$	index of timestamp in GPS location dataset	$T = 200,000$
$r \in [1, R]$	refueling ticket in observation dataset	$R = 3019$

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus.

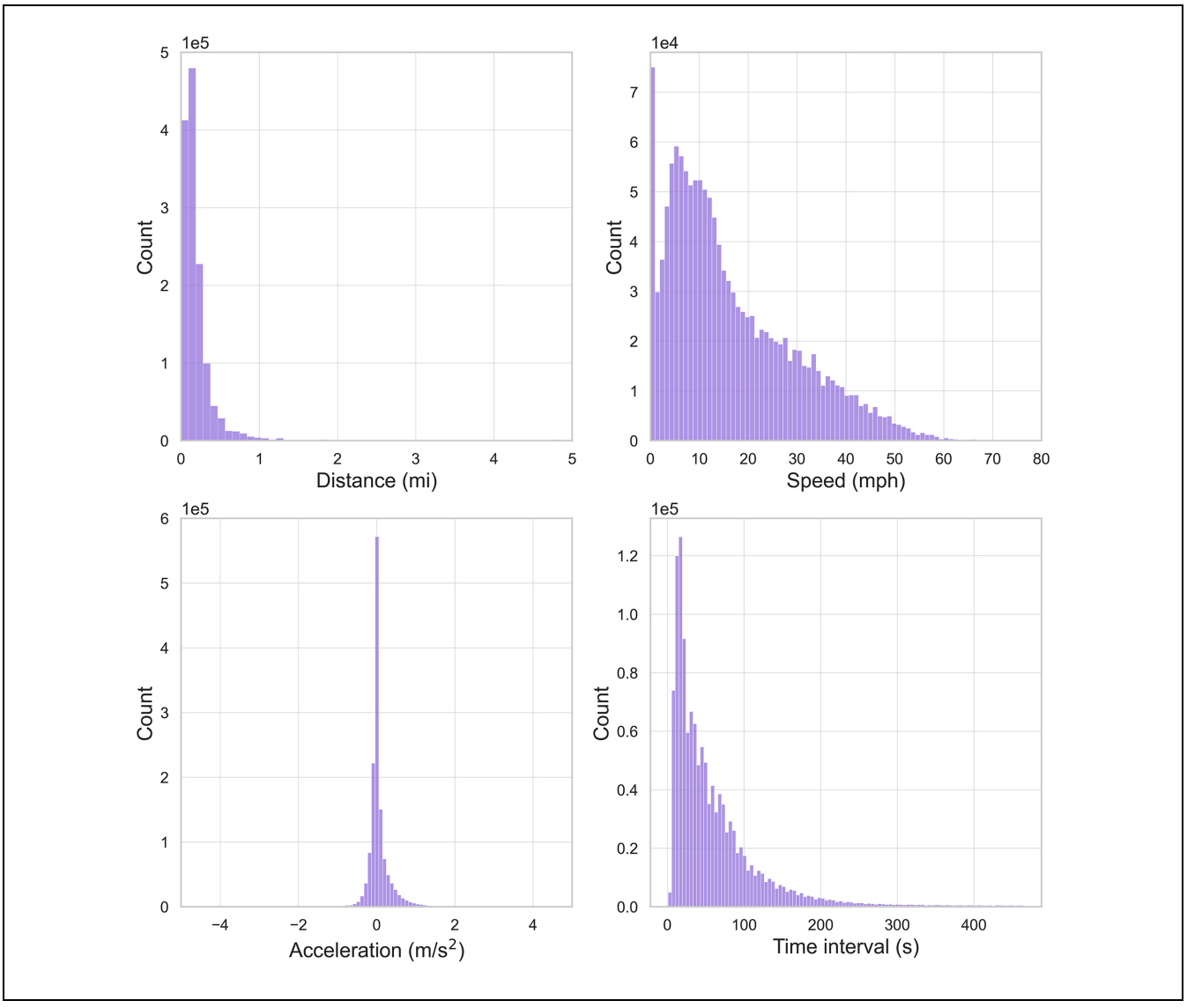

Figure 6. Histogram of trajectories: distance (*top left*), speed (*top right*), acceleration (*bottom left*), and time intervals (*bottom right*) in April, 2022. These histograms are trimmed for better clarity, with 3% of their data points removed.

- C_D = drag coefficient (unitless),
- C_h = altitude correction factor (unitless),
- $A_{f,p}$ = frontal area (m^2),
- C_r = rolling coefficient (unitless),
- c_1 = rolling resistance coefficient 1 (unitless),
- c_2 = rolling resistance coefficient 2 (unitless),
- g = gravitational acceleration (m/s^2),
- $v_i(t_j)$ = speed of bus i (km/h), at timestamp (t_j),
- $a_i(t_j)$ = acceleration of bus i (m/s^2), at timestamp (t_j),
- $G(t_j)$ = road grade (specified as the tangent of the angle) at the location of bus i at timestamp (t_j), and
- $m_i(t_j)$ = mass of bus i (kg), including passengers, at timestamp (t_j).

Conventional Diesel Bus (CDB) Model. The input power, $P_{C,i}(t_j)$ in kW, (or fuel power consumed) by vehicle i with a conventional diesel powertrain C, at timestamp t_j, is given by the quotient of the tractive power and the CDB driveline efficiency:

$$P_{C,i}(t_j) = \frac{1}{\eta_{d,C}} P_{C,i}(t_j) \tag{6}$$

where

$\eta_{d,C}$ = the driveline efficiency (0.94).

The fuel consumption rate $F_{C,i}(t_j)$ in liters per second (l/s) at timestamp t_j is then given by the piecewise function:

$$F_{C,i}(t_j) = \begin{cases} \alpha_{0,C} + \alpha_{1,C} P_{C,i}(t_j) \\ + \alpha_2 P_{C,i}^2(t_j) & \text{if} \quad P_{C,i}(t_j) \geq 0 \\ \alpha_{0,C} & \text{if} \quad P_{C,i}(t_j) < 0 \end{cases} \tag{7}$$

where
the constant $\alpha_{0,C}$ and $\alpha_{1,C}$ = powertrain-specific coefficients to be calibrated.
Thus, the diesel consumption in liters ($E_{C,i}(t_j)$) at timestamp (t_j) relative to the previous timestamp of a CDB is then computed as:

$$E_{C,i}(t_j) = \Delta_{t_j} \cdot F_{C,i}(t_j) \tag{8}$$

where
Δ_{t_j} = the duration of time interval between timestamp t_j and timestamp t_{j-1}.

Hybrid Electric Bus (HEB) Model. We assume the same energy model formulation for HEB as for CDB. Thus, instead of explicitly modeling the hybrid control unit mechanism, we account implicitly for the improved efficiency by calibrating $\alpha_{0,H}$ and $\alpha_{1,H}$ to fit the trajectories of the hybrid buses. Thus, the diesel consumption in liters ($E_{H,i}(t_j)$) of an HEB at timestamp t_j relative to previous timestamp t_{j-1} is then computed as:

$$E_{H,i}(t_j) = \Delta_{t_j} \cdot FH, i(t_j) \tag{9}$$

Battery Electric Bus (BEB) Model. To calculate the input power from the battery of a BEB, we account for regenerative braking, as well as the efficiencies of the electric motor and the battery (*22*). Thus, the power consumed, $P_{B,i}(t_j)$ (kW), is given by:

$$P_{B,i}(t_j) = \frac{\eta_{rb}\eta_{batt}}{\eta_m \eta_d} P_{B,i}(t_j) \tag{10}$$

where

- η_{batt} = battery efficiency (unitless),
- η_m = motor efficiency (unitless), and
- η_{rb} = regenerative braking efficiency (unitless).

The regenerative braking efficiency is specified by:

$$\eta_{rb} = \begin{cases} \exp\left[-\left(\frac{\gamma}{|a(t_j)|}\right)\right] & \text{if} \quad a(t_j) < 0 \\ 1 & \text{if} \quad a(t_j) \geq 0 \end{cases} \tag{11}$$

where
γ = the regenerative decay parameter and is calibrated to fit the dataset.
Then, energy (kWh) consumed at timestamp t_j relative to the previous timestamp t_{j-1} of a BEB is computed as:

Table 6. Calibrated Parameters

Parameter	CDB	HEB	BEB
α_0	✓	✓	
α_1	✓	✓	
γ			✓

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus.

$$E_{B,i}(t_j) = \frac{1}{3600} \Delta_{t_j} P_{B,i}(t_j) \tag{12}$$

Calibration Procedure

We selected five parameters for calibration. These are shown in Table 6:

Since the unit of observation for energy consumption was the duration of the refueling interval, we specify the energy consumed therein—vehicle refueling interval energy, $VRIE_{i,p,t_r}$—as:

$$VRIE_{i,p}(t_r) = \sum_{t_j \in R} E_{i,p}(t_j), \tag{13}$$

where
r = the index for the refueling interval ($t_r = \{0, 1, \ldots, R\}$).
$VRIE_{i,p}(t_r)$ is measured in gallons of diesel (gal) for CDB and HEB.

To calibrate the powertrain-specific parameters $1\theta_p$, for CDB and HEB, we minimize the root mean squared error (RMSE) of $VRIE_{i,p}(t_r)$. The error metric is given by:

$$RMSE_p = \sqrt{\frac{1}{N_p R} \sum_{i=1}^{N_p} \sum_{r=1}^{R} \left(VRIE_{i,p}(t_r) - \widehat{VRIE}_{i,p}(t_r) \right)^2}, \tag{14}$$

where
N_p = the number of vehicles of powertrain type $p \in \{C, H\}$,
R = the number of refueling intervals in the observed energy data, and $\widehat{VRIE}_{i,p}(t_r)$ = the predicted VRIE.
Thus, the minimization problem is as follows:

$$\boldsymbol{\theta}_p^* = \underset{\boldsymbol{\theta}_p}{\arg\min}\, RMSE_{p(\boldsymbol{\theta})} \tag{15}$$

where
$\boldsymbol{\theta}_{CDB} = (\alpha_{0,C}, \alpha_{1,C})$, and
$\boldsymbol{\theta}_{HEB} = (\alpha_{0,H}, \alpha_{1,H})$.

To minimize $RMSE_p$, we performed a grid search for the parameters using the first 77% of the 2,749 refueling ticket observations for training. In the grid searches, we evaluated the model at each point and selected the

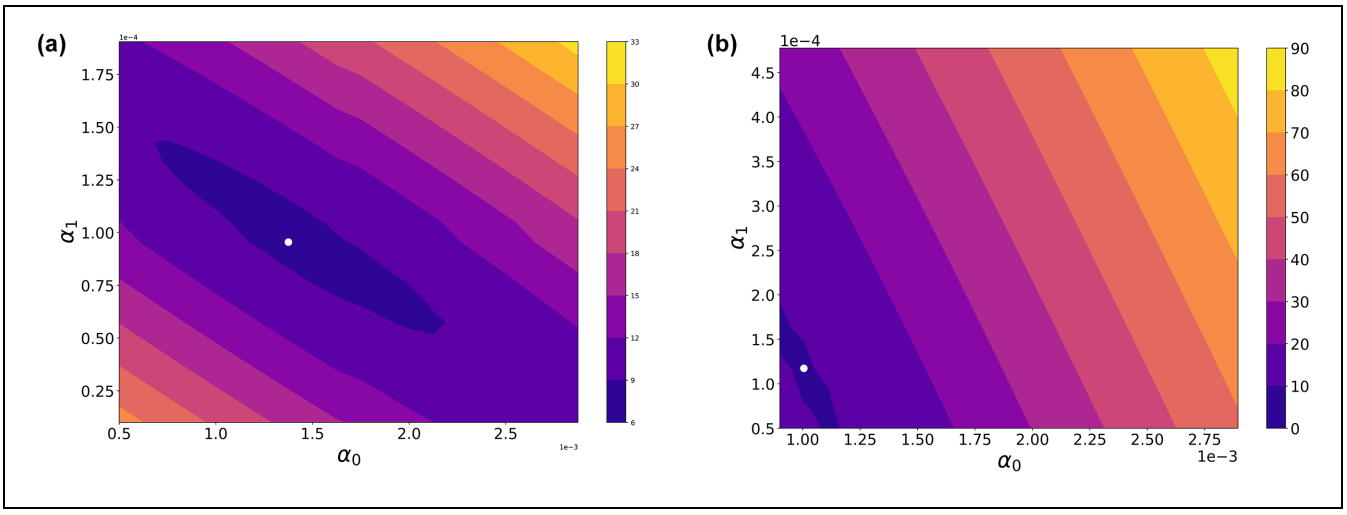

Figure 7. The conventional diesel bus (CDB) model energy root mean squared error (gal) resulting from calibration of α_0 and α_1 in the: (*a*) CDB model and (*b*) hybrid electric bus model.
Note: white markers = optimal points.

optimal parameters. This left the 691 data points corresponding to the last 7 days in April, 2022, for testing. For the CDB model, we considered a 100×100 grid of values for $\alpha_0 \in [5 \times 10^{-4}, 3 \times 10^{-3}]$ and $\alpha_1 \in [1 \times 10^{-5}, 2 \times 10^{-4}]$. In the HEB case, we initially conducted a coarse grid search for $\alpha_0 \in [8 \times 10^{-6}, 0.02]$ and for $\alpha_1 \in [1 \times 10^{-6}, 4 \times 10^{-3}]$. We then narrowed the search over a 100×100 grid of values for $\alpha_0 \in [9 \times 10^{-4}, 3 \times 10^{-3}]$ and $\alpha_1 \in [5 \times 10^{-5}, 5 \times 10^{-4}]$. Contour plots of the objective function for CDB and HEB are shown in Figure 7.

Energy consumption data for BEB were unavailable for April, 2022. Thus, we fitted our model a mean BEB energy economy ($\overline{eco}_B$) of 0.44 mpkWh (*23*). To calibrate the BEB model, first we obtain energy predictions at each timestamp, $\hat{E}_{i,B}(t_j)$. Then, we compute the predicted energy economy at that same timestamp, $\widehat{eco}_{i,B}(t_j)$, as:

$$\widehat{eco}_{i,B}(t_j) = \frac{s_i(t_j)}{\hat{E}_{i,B}(t_j)} \tag{16}$$

Finally, we obtain the optimal γ^* by minimizing the root mean squared deviation ($RMSD_B$) objective:

$$RMSD_B = \sqrt{\frac{1}{N_B \cdot T} \sum_{i=1}^{N_B} \sum_{j=1}^{T} \left(\widehat{eco}_{i,B}(t_j) - \overline{eco}_B\right)^2} \tag{17}$$

Thus, we conducted a search for $\gamma \in [0.02, 5]$ over 200 equally spaced points to solve the problem:

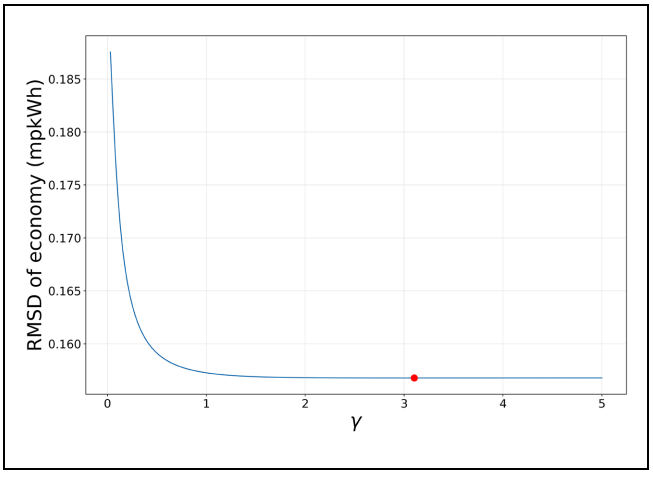

Figure 8. The battery electric bus model economy root mean squared deviation (RMSD) (mpkWh) resulting from calibration of the regenerative decay parameter (γ).
Note: red marker = the optimal point.

$$\boldsymbol{\theta}^*_{BEB} = \gamma^* = \underset{\gamma}{\arg\min}\, RMSD_{B(\gamma)} \tag{18}$$

The objective trajectory is shown in Figure 8.

Results and Discussion

Model Performance

The optimal parameter values resulting from the model calibration are shown in Table 7. We then computed the final performance of the models using 23% of the VRIE observations at vehicle level. The test RMSE is 8.19 gal for CDB, 8.01 gal for HEB, and 0.12 mpkWh

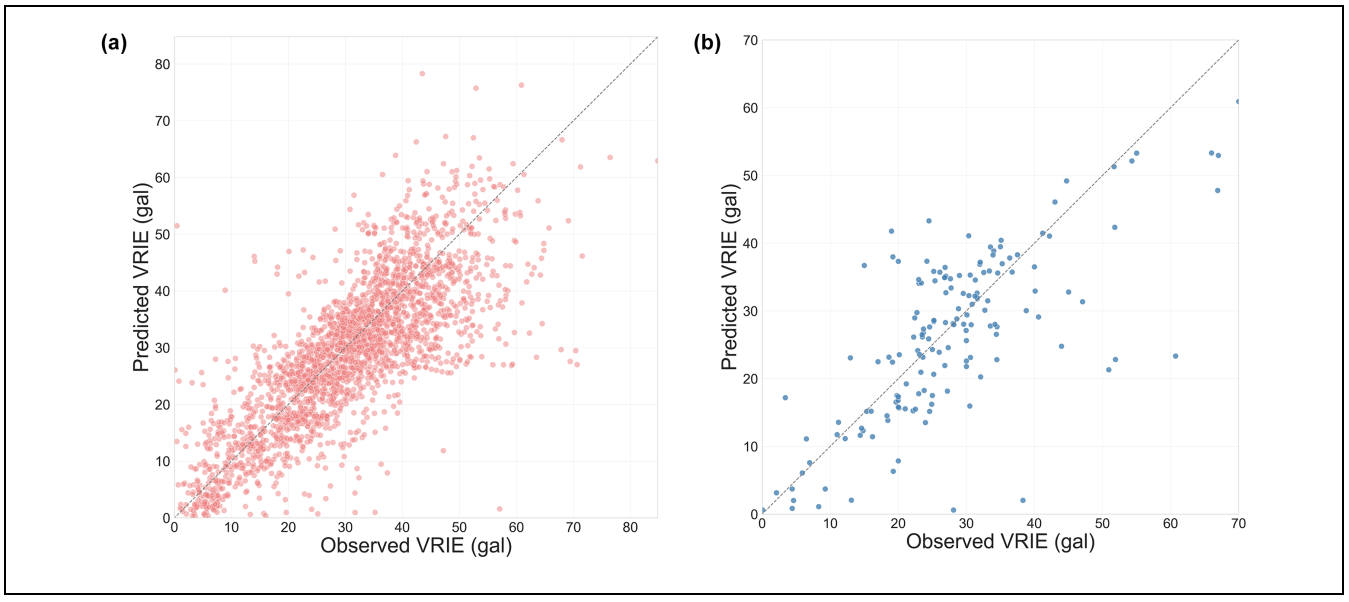

Figure 9. Predicted versus observed vehicle refueling interval energy (VRIE) (gal) for: (*a*) CDB and (*b*) HEB.

Table 7. Calibrated Parameters

Parameter	Description	CDB	HEB	BEB
$\hat{\alpha}_0$	intercept of fuel consumption rate function	1.68×10^{-3}	1.4×10^{-3}	na
$\hat{\alpha}_1$	linear term of fuel consumption rate function	8.82×10^{-5}	6.82×10^{-5}	na
$\hat{\gamma}$	regenerative braking efficiency decay parameter	na	na	3.10

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus; na = not applicable.

Table 8. Model Training and Test Performance Metric (RMSE) at Vehicle Level (Based on Vehicle Refueling Interval Energy)

Powertrain	RMSE	
	Train	Test
CDB (gal)	8.63	8.19
HEB (gal)	9.66	8.01
BEB (mpkWh)	0.15	0.12

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus; RMSE = root mean squared error.

Table 9. Model Training and Test Performance Metric (MAPE) at Vehicle Level (Based on Vehicle Refueling Interval Energy) and System Level (Weekly Aggregated)

	Vehicle-level MAPE (%)		System-level MAPE (%)	
Powertrain	Train	Test	Train	Test
CDB	47	28	4	5
HEB	35	26	9	2
BEB	26	18	na	na

Note: BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus; MAPE = mean absolute percentage error; na = not applicable.

for BEB (Table 8). To further assess performance, we calculated mean absolute percentage error (MAPE) for each powertrain-specific model (Table 9). At the vehicle level, the test MAPE is 28% for CDB, 26% for HEB, and 18% for BEB. Average fuel economies for CDB and HEB were computed as 5.35 MPG and 6.44 MPG, respectively (which are close to 4.82 MPG and 5.84 MPG from the literature). Despite the CDB and HEB MAPEs being relatively high when computed at the vehicle level, these models exhibited better performance when evaluated on an aggregated weekly system-level basis. In this case, the test MAPE is 5% for CDB, and 2% for HEB. We were unable to compute system-level MAPE values for BEB, as charging data were unavailable.

In Figure 9, the model predictions of VRIE (gal) are plotted against the observed values for CDB and HEB. It is evident that both the CDB and HEB models perform well in estimating energy. The HEB model is slightly inferior because of the relatively small size of the HEB fleet operating on the PVTA network. (HEB on the PVTA network comprise only 5% of the total fleet, while CDB comprise 88%). This limited size of the HEB fleet restricts the ability to calibrate the HEB model parameters for better performance.

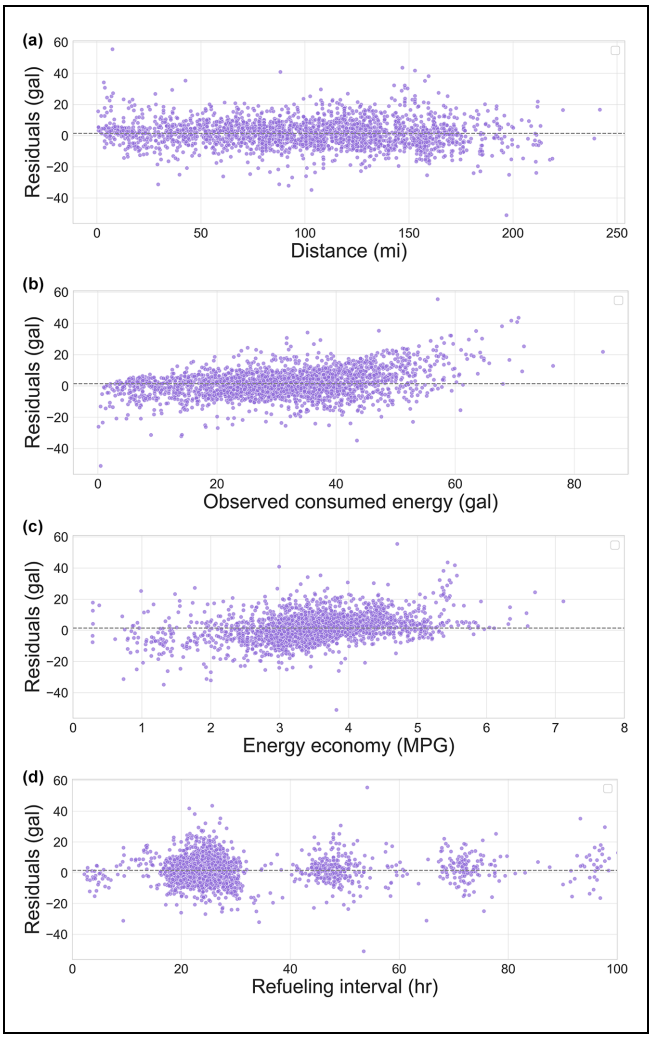

Figure 10. Vehicle refueling interval energy residuals (gal) versus: (*a*) distance (mi), (*b*) observed energy (gal), (*c*) fuel economy (MPG), and (*d*) refueling interval (h) for conventional diesel bus.
Note: dashed horizontal lines = the average residual.

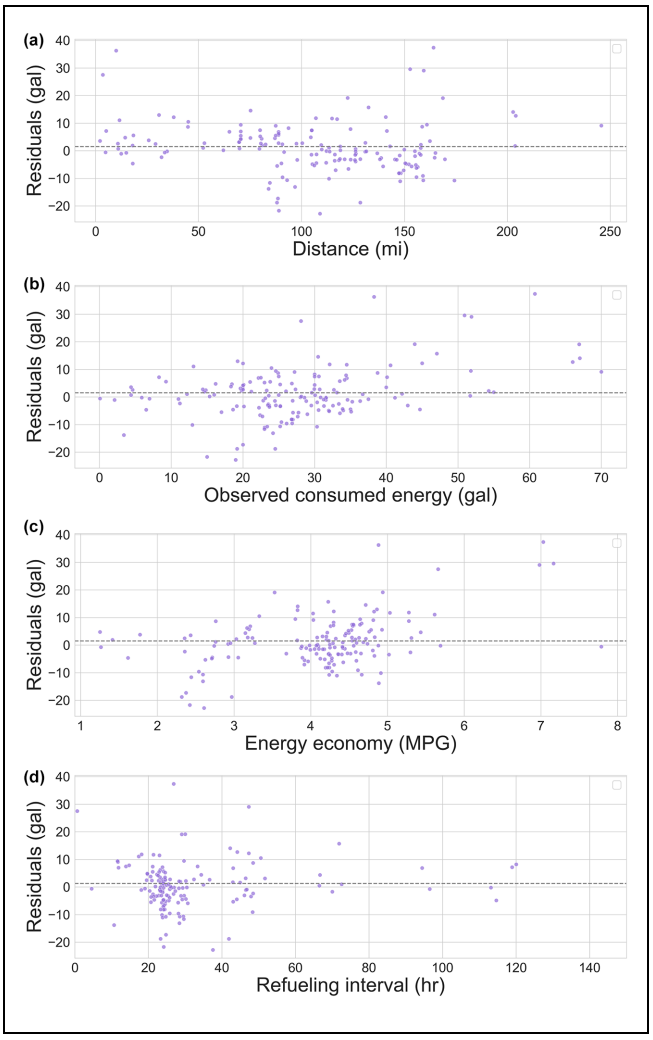

Figure 11. Vehicle refueling interval energy residuals (gal) versus: (*a*) distance (mi), (*b*) observed energy (gal), (*c*) fuel economy, and (*d*) refueling interval (h) for hybrid electric bus.
Note: dashed horizontal lines = the average residual.

We analyzed the residuals to evaluate the fitted models (Figures 10 and 11). When plotted against distance, the residuals display a largely homoskedastic pattern with constant variance. However, when the residuals are plotted against energy and fuel economy, there appears to be a slightly positive correlation. This behavior indicates a systematic error, as there is a linear dependency the current models fail to capture. This discrepancy is further highlighted when we consider the residuals versus refueling interval plots (Figures 10*d* and 11*d*). We observe that the absolute value of the residuals decreases with the refueling interval. This pattern may also point to the presence of a measurement error in the distances, as the timestamped locations do not include bus movements to and from depots. Thus, the impact of this error reduces as the refueling interval gets larger. We plan to further investigate and address these errors in future work.

Model predictions aggregated at the weekly level for CDB and HEB are compared with approximate weekly observations (Figure 12). The observed values are approximate given the irregularity of the refueling intervals and their variability by vehicle, which results in only a partial observation of the number of gallons consumed in a given day. We show the mean energy economy from the BEB model (0.36 mpKWh) compared with that obtained from the literature (Figure 13).

Hourly Predicted Energy Metrics

We used the calibrated models to predict vehicle energy consumption (gallons for CDB and HEB and kWh for

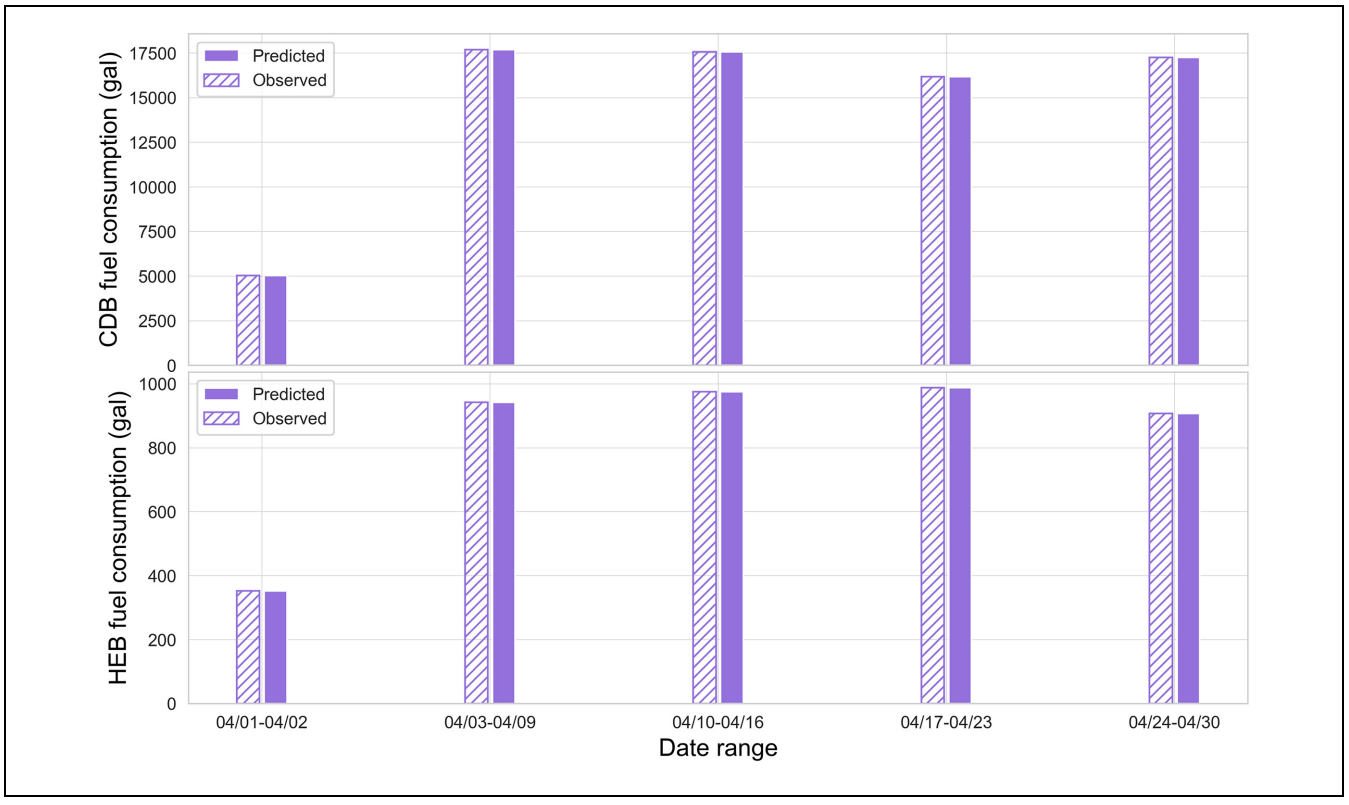

Figure 12. Weekly time series of system-wide energy consumption (gal) for conventional diesel bus (CDB) (*top*) and hybrid electric bus (HEB) (*bottom*) in 2022 (first 23 days: training; last 7 days: test).

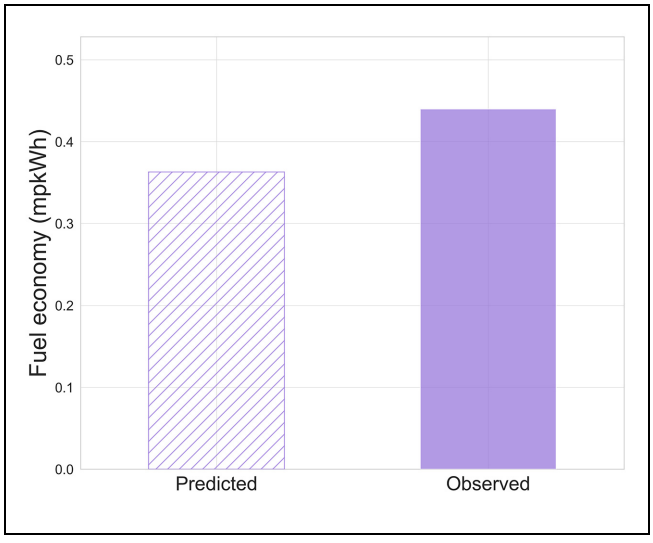

Figure 13. Battery electric bus mean predicted energy economy compared with observed value from the literature.

BEB) at each timestamp. We then computed other metrics to investigate system-wide performance for each powertrain type. These metrics are fuel efficiency (MPG for CDB and HEB, and mpkWh for BEB) and energy intensity (gallons per passenger for CDB and HEB, and kWh per passenger for BEB).

In the CDB cases, the period between 2:00 a.m. and 4:00 a.m. has the lowest average per-vehicle energy consumption with a median of 0.57 gal, since the ridership is minimum during this time. However, energy consumption peaks at 2:00 p.m. with a median of 242.24 gal, as ridership is high between 2:00 p.m. and 5:00 p.m. We observe the peak ridership at 4:00 p.m., with a median of 30,637 passengers. Additionally, at 6:00 p.m., vehicles travel a longer distance per gallon (6.02 MPG) possibly because of lower passenger load. Energy consumption is more intense per passenger between 2:00 a.m. and 5:00 a.m. with a median of 0.007 gal per passenger, as ridership is at its minimum during this time period (Figure 14*a*).

For HEB, energy consumption peaks at 7:00 p.m. with a median 9.15 gal, while we observe almost no energy consumed at 4:00 a.m. The highest fuel economy occurs at 4:00 a.m., with a median of 8.64 MPG, and the most intense energy consumption is at 1:00 a.m. with a median of 0.007 gal per passenger, when there are a few passengers onboard. Ridership is at its peak at 4:00 p.m., with a median of 1,812 passengers (Figure 14*b*).

Energy consumption peaks at 2:00 p.m. for BEB. We observe the highest energy economy (0.73 mpkWh) at 11:00 a.m. and energy intensity (3.41 kWh per passenger) at 11:00 p.m. Ridership peaks at 2:00 p.m. with a median

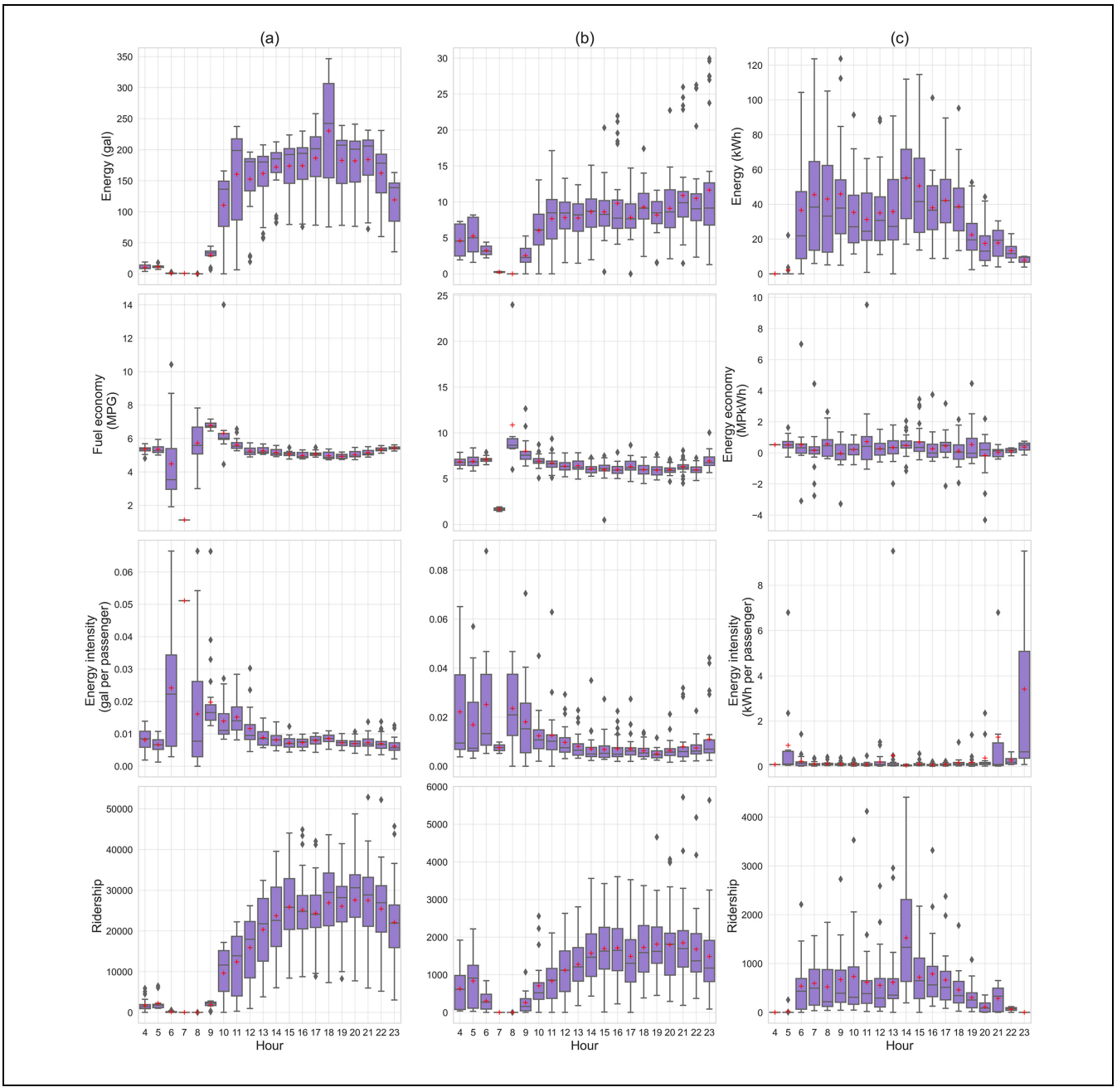

Figure 14. Hourly system-level energy consumption, energy economy, energy intensity, and ridership for: (*a*) CDB, (*b*) HEB, and (*c*) BEB.
Note: the horizontal line inside the purple box = the median value; the red marker inside the purple box = the mean value; BEB = battery electric bus; CDB = conventional diesel bus; HEB = hybrid electric bus.

of 1,333 passengers (Figure 14*c*). Summary statistics for all three powertrain-specific models are provided in Tables 10 and 11.

Potential Applications

The proposed systematic approach to model energy consumption has several potential applications, as it provides highly granular energy predictions at the vehicle level. Thus, it can be used to investigate the impacts of demand and congestion at different times of the day. The framework can also be extended to model and predict emissions. With this capability, both energy and emissions can be computed and monitored spatiotemporally. Planning agencies can use this information to understand the zones and routes that consume the most energy or that are exposed to the greatest emissions. Ultimately, using this framework, transit agencies can

Table 10. Summary Statistics of Hourly Energy Prediction Metrics for CDB and HEB

Metric	CDB		HEB	
	Mean	Median	Mean	Median
Energy (gal)	120.24	138.79	7.54	7.36
Fuel economy (MPG)	5.34	5.24	6.44	6.30
Energy intensity (gallons per passenger)	0.009	0.008	0.009	0.006
Ridership (passengers)	16, 254	16, 255	1, 266	1, 136

Note: CDB = conventional diesel bus; HEB = hybrid electric bus.

Table 11. Summary Statistics of Hourly Energy Prediction Metrics for BEB.

Metric	Mean	Median
Energy (kWh)	35.97	28.71
Energy economy (mpkWh)	0.36	0.28
Energy intensity (kWh per passenger)	0.23	0.08
Ridership (passengers)	602	355

Note: BEB = battery electric bus.

develop effective fleet electrification strategies. They can prioritize which routes or vehicles to replace based on the powertrain-specific model outputs. Analyses can also include the investigation of energy consumption and emissions exposure for various socioeconomic and demographic segments of the population, given the spatial resolution of the model. Thus, planners can incorporate equity considerations for sustainable decision-making, including electrification and disaster response, among others.

Conclusion

We have developed a framework for predicting the energy consumption of buses based on calibrated vehicle-specific engines, using real-time bus movement data. The framework accurately predicts energy consumption for conventional (CDB), hybrid (HEB), and electric buses (BEB). To demonstrate the effectiveness of our approach, we conducted a case study using the PVTA bus network.

Our integrated framework uses GPS location data to compute real-time vehicle energy consumption using calibrated powertrain-specific models. This novel framework is the first to demonstrate, to the best of our knowledge, a capability to predict highly spatiotemporally resolved vehicle-specific energy consumption for an entire bus network. Further avenues for research include investigating the effects of seasonal fluctuations on energy consumption and refining model calibration via a hierarchical Bayesian estimation method. The models can also be extended to predict vehicle-specific emissions, as well. Ultimately, the framework can be applied by transit agencies to various networks, serving as a valuable decision-support tool for equitable and effective fleet management and replacement, carbon mitigation, and cost reduction strategies.

Acknowledgments

The authors thank Alex Forrest, Tolu Oke, and Benjamin Amand (Pioneer Valley Transit Authority) for supporting this research and providing the data used in the analysis. Their valuable insights and feedback contributed to the quality of this study. We also thank Glenn Barrington, Thomas Vincent, and Matt Moretti (University of Massachusetts Transit Service) for providing further validation data for the study.

Author Contributions

The authors confirm their contribution to the paper as follows: study conception and design: J. Oke, M. Arabi; data collection: M. Arabi; analysis and interpretation of results: M. Arabi, J. Oke; draft manuscript preparation: M. Arabi, J. Oke. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research presented in this paper was funded by the Pioneer Valley Transit Authority via the Federal Transit Administration's Helping Obtain Prosperity for Everyone (HOPE) Program.

ORCID iDs

Mahsa Arabi https://orcid.org/0000-0003-1784-0906
Jimi Oke https://orcid.org/0000-0001-6610-445X

Data Accessibility Statement

All the code used in generating the results and figures in this paper are publicly available at https://github.com/narslab/bus-transit-energy.

References

1. Energy Information Association (EIA). Independent Statistics and Analysis. February 2023. https://www.eia.gov/index.php.
2. American Public Transportation Association (APTA). *Public Transportation Fact Book*. May 2021. https://www.apta.com/research-technical-resources/transitstatistics/public-transportation-fact-book.
3. International Energy Agency (IEA). Energy Intensity of Passenger Transport Modes, 2018 – Charts – Data & Statistics. 2018. https://www.iea.org/data-and-statistics/charts/energyintensity-of-passenger-transport-modes-2018.
4. Holland, S. P., E. T. Mansur, N. Z. Muller, and A. J. Yates. The Environmental Benefits of Transportation Electrification: Urban Buses. *Energy Policy*, Vol. 148, 2021, p. 111921. https://doi.org/10.1016/j.enpol.2020.111921.
5. Immonen, P., P. Lindh, and J. Pyrhönen. Analyses of Energy Saving in a Hybrid and Electric Bus. *Proc., IEEE International Electric Machines and Drives Conference (IEMDC)*, Miami, FL, May 21–24, 2017, IEEE, New York, pp. 1–6. https://doi.org/10.1109/IEMDC.2017.8002226.
6. Federal Transit Administration (FTA). The National Transit Database (NTD). 2022. https://www.transit.dot.gov/ntd/ntd-data.
7. Wang, J., and H. A. Rakha. Modeling Fuel Consumption of Hybrid Electric Buses: Model Development and Comparison with Conventional Buses. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2539: 94–102.
8. Wang, J., and H. A. Rakha. Fuel Consumption Model for Heavy Duty Diesel Trucks: Model Development and Testing. *Transportation Research Part D: Transport and Environment*, Vol. 55, 2017, pp. 127–141. https://doi.org/10.1016/j.trd.2017.06.011.
9. Wang, J. *Multi-Modal Energy Consumption Modeling and Eco-routing System Development*. PhD thesis, Virginia Tech, July 2017.
10. Guensler, R., H. Liu, Y. (Ann) Xu, A. Akanser, D. Kim, M. P. Hunter, and M. O. Rodgers. Energy Consumption and Emissions Modeling of Individual Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2627: 93–102.
11. Elmi, S., and K.-L. Tan. DeepFEC: Energy Consumption Prediction Under Real-World Driving Conditions for Smart Cities. *Proc., WWW '21: The Web Conference 2021*, Ljubljana, Slovenia, April 19–23, 2021, Association for Computing Machinery, New York, NY, pp. 1880–1890. https://doi.org/10.1145/3442381.3449983.
12. Śmieszek, M., and V. Mateichyk. Determining the Fuel Consumption of a Public City Bus in Urban Traffic. *IOP Conference Series: Materials Science and Engineering*, Vol. 1199, No. 1, 2021, p. 012080. https://doi.org/10.1088/1757-899X/1199/1/012080.
13. Sun, R., Y. Chen, A. Dubey, and P. Pugliese. Hybrid Electric Buses Fuel Consumption Prediction Based on Real-World Driving Data. *Transportation Research Part D: Transport and Environment*, Vol. 91, 2021, p. 102637. https://doi.org/10.1016/j.trd.2020.102637.
14. Yang, Y., N. Gong, K. Xie, and Q. Liu. Predicting Gasoline Vehicle Fuel Consumption in Energy and Environmental Impact Based on Machine Learning and Multidimensional Big Data. *Energies*, Vol. 15, No. 5, 2022, p. 1602. https://doi.org/10.3390/en15051602.
15. Han, Z., E. Gonzales, E. Christofa, and J. Oke. Modeling System-Wide Urban Rail Transit Energy Consumption: A Case Study of Boston. *Transportation Research Record: Journal of the Transportation Research Board*, 2022. 2676: 627–640.
16. Barth, M., G. Scora, and T. Younglove. Modal Emissions Model for Heavy-Duty Diesel Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, 2004. 1880: 10–20.
17. Wang, J., and H. A. Rakha. Fuel Consumption Model for Conventional Diesel Buses. *Applied Energy*, Vol. 170, 2016, pp. 394–402. https://doi.org/10.1016/j.apenergy.2016.02.124.
18. Wang, J., and H. A. Rakha. Convex Fuel Consumption Model for Diesel and Hybrid Buses. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2647: 50–60.
19. Hjelkrem, O. A., K. Y. Lervåg, S. Babri, C. Lu, and C.-J. Södersten. A Battery Electric Bus Energy Consumption Model for Strategic Purposes: Validation of a Proposed Model Structure with Data from Bus Fleets in China and Norway. *Transportation Research Part D: Transport and Environment*, Vol. 94, 2021, p. 102804. https://doi.org/10.1016/j.trd.2021.102804.
20. Pioneer Valley Transit Authority (PVTA). FY21 Annual Report. 2021. https://pvta.com/documents/annualReports.
21. Azdy, R. A., and F. Darnis. Use of Haversine Formula in Finding Distance Between Temporary Shelter and Waste End Processing Sites. *Journal of Physics: Conference Series*, Vol. 1500, No. 1, 2020, p. 012104. https://doi.org/10.1088/1742-6596/1500/1/012104.
22. Fiori, C., K. Ahn, and H. A. Rakha. Power-Based Electric Vehicle Energy Consumption Model: Model Development and Validation. *Applied Energy*, Vol. 168, 2016, pp. 257–268. https://doi.org/10.1016/j.apenergy.2016.01.097.
23. Ma, X., R. Miao, X. Wu, and X. Liu. Examining Influential Factors on the Energy Consumption of Electric and Diesel Buses: A Data-Driven Analysis of Large-Scale Public Transit Network in Beijing. *Energy*, Vol. 216, 2021, p. 119196. https://doi.org/10.1016/j.energy.2020.119196.

The authors are solely responsible for the facts, the accuracy of the data and analysis, and the views presented here.