AIMS Energy

AIMS Energy, 14(3): 732–759.
DOI: 10.3934/energy.2026030
Received: 01 March 2026
Revised: 30 May 2026
Accepted: 15 June 2026
Published: 30 June 2026

https://www.aimspress.com/journal/energy

Research article

TALe-Bus: An interactive dashboard for electric bus fleet electrification planning

Kareem Othman[1,2,*], Diego Da Silva[1], Amer Shalaby[1] and Baher Abdulhai[1]

[1] Department of Civil and Mineral Engineering, University of Toronto, Toronto, Ontario, Canada
[2] Public Works Department, Faculty of Engineering, Cairo University, Giza, Egypt

* **Correspondence:** Email: kareem.othman@utoronto.ca; karemmohamed1993@cu.edu.eg. Tel: +1(416)978-3099.

Abstract: This study presents the development of the Transit Analytics Lab Electric Bus (TALe-Bus) Dashboard, an integrated decision-support tool designed to support transit agencies in planning the transition from conventional diesel buses to battery-electric buses. The dashboard combines predictive modeling, data analytics, and interactive visualization to estimate electric bus energy consumption rates, fleet size requirements, replacement factors, and maximum operational range for the case of overnight depot charging. The dashboard was built based on real-world data collected from the operations of 60 battery-electric buses on 48 routes in Toronto. The modeling workflow includes data preprocessing, feature selection, and the development and comparison of multiple statistical and machine learning energy prediction models. Tree-based modeling techniques outperformed the other techniques, demonstrating strong capability in capturing the relationships between operational, environmental, vehicle, and route characteristics and the energy consumption rate, achieving a root mean square error (RMSE) of 0.13 kWh/km. These techniques were therefore adopted as the core predictive engine of the system. The fleet-sizing module integrates traditional transit planning formulations (for diesel fleets) with electric-bus energy and range constraints to estimate both electric and diesel fleet requirements and compute the replacement factor, a key indicator reflecting the relative fleet needs for electrification. The dashboard also estimates the maximum operational bus range based on the battery capacity and real-world operating conditions, supporting reliable service planning. The system is implemented as an interactive web-based platform that provides geospatial visualization of route electrification feasibility and a scenario-based interface for customized operational analysis. By translating complex predictive analytics into a practical planning tool, the TALe-Bus Dashboard supports informed, data-driven decision-making for fleet electrification and infrastructure planning.

Keywords: electrification; decision support tool; interactive dashboard; transit planning; transition planning; sustainable transportation

1. Introduction

The global move toward sustainable and low-carbon urban transportation systems has brought the issue of the electrification of public transit systems to the forefront as a primary concern for policymakers and industry leaders [1,2]. Electric buses have emerged as a key solution for urban cities around the world to tackle the issue of greenhouse gas emissions, reduce fossil fuel energy consumption, and mitigate the negative environmental impacts of vehicle pollution [3–12]. However, the transition from conventional diesel-powered buses to fully electric-powered buses faces numerous complexities and intricacies that must be carefully planned and analyzed through the use of sophisticated predictive tools and methodologies [13]. The transition from diesel-powered to electric-powered buses cannot simply be seen as the transition from one type of vehicle to another; rather, it represents a substantial adjustment of the entire transit system, fleet operations, and service delivery strategies. Issues such as range limitations, battery degradation, differences in energy consumption patterns, climatic conditions, and other operational characteristics make the transition process extremely complicated [14–19]. Moreover, the need to ensure the transition does not compromise cost-effectiveness and service reliability adds another dimension to the overall complexity of the transition process.

The literature on electric bus fleet planning has expanded rapidly over the last few years, offering various models and frameworks to support the electrification of bus fleets [20–23]. Early models primarily focused on estimating energy consumption using simplified linear models [24]. Subsequent research studies introduced more sophisticated approaches that incorporate vehicle dynamics, traffic conditions, terrain features, and external environmental variables, such as temperature and wind speed, for energy consumption prediction [25–30]. In addition, other studies have mainly focused on testing and evaluating different charging strategies (charging infrastructure and distinguishing between depot charging, opportunity charging, and on-route fast charging, as these choices directly affect operational feasibility and infrastructure requirements) in the transition process [31–35]. Moreover, other studies focused on the development of optimization models for fleet scheduling [36–39]. Mixed-integer linear programming (MILP), heuristic algorithms, and simulation-based approaches have been utilized to optimize electric bus deployment, charging schedules, and fleet assignments [40–44]. These models often aim to minimize total operational costs, reduce peak electricity demand, or maximize service reliability. For energy consumption prediction, multiple studies have focused on the development of energy prediction models using simulations, statistical regressions, and, more recently, machine learning techniques, which demonstrated improved accuracy and ability to capture complex, nonlinear relationships influencing energy consumption and fleet needs [45–58].

While the existing models provide valuable contributions, several research gaps remain. First, many existing models require access to detailed operational data, such as state-of-charge profiles or real-time energy consumption information, which may not be readily available at the early stages of the planning process. Second, only a limited number of existing tools offer the capability for both energy consumption prediction and fleet size estimation within the same platform, which can be easily used by non-specialist practitioners. Third, the majority of existing models do not offer the capability to dynamically take into account the impact of seasons and variations in bus characteristics, which can play important roles in cold-climate contexts. Finally, the availability of visualization tools for planners to explore scenarios at the city or route level is limited. Thus, this study introduces the TALe-Bus Fleet Planning Dashboard, a comprehensive and accessible decision-support tool designed to assist transit agencies in planning their electrification strategies. The dashboard uniquely combines machine learning–based energy consumption prediction with fleet size estimation, incorporating operational, route, vehicular, and environmental variables available to transit agencies during the planning phase. The predictive models were developed using real-world year-round operational data from Toronto's electric bus fleet, ensuring that the dashboard captures the complexities of cold-weather operations, varied route geometries, and diverse bus configurations. Machine learning methods such as XGBoost and random forest were employed to model energy consumption and fleet requirements, achieving superior predictive performance compared to traditional regression techniques. The dashboard provides an intuitive, map-based interface that enables users to visualize replacement factors (ratio of the required electric bus fleet size to the diesel engine fleet size to offer the exact same transit service) across the bus network of Toronto and to conduct customized predictive analyses by entering specific route, operational, vehicle, and external information. The TALe-Bus dashboard offers several key contributions and advantages over existing approaches in previous studies. The primary contribution of this study is the development of TALe-Bus, an interactive decision-support dashboard that integrates energy consumption prediction, fleet sizing estimation, replacement factor analysis, and geospatial visualization into a unified planning framework for transit electrification. By combining these analytical components within a user-friendly platform, the dashboard enables route-level assessment of electrification feasibility while accounting for seasonal and operational variability. Furthermore, the tool allows transit agencies to explore different deployment scenarios without requiring highly granular operational datasets, thereby facilitating evidence-based planning and evaluation of electric bus implementation strategies.

The remainder of this paper is organized as follows: Section 2 provides a general overview of the workflow and introduces the TALe-Bus dashboard framework. Section 3 details the data sources utilized to build the dashboard and highlights the variables used in this study. Section 4 focuses on comparing the energy behavior and performance of the different bus types used in this study across the different seasons. Section 5 describes the feature selection process and the use of Ward's hierarchical method for feature comparisons and removal. Section 6 focuses on the development of the energy prediction model and compares the performance of different modeling techniques. Section 7 provides details on the methodology used for estimating the fleet sizing and replacement factor. Section 8 discusses the dashboard architecture and its system integration. Section 9 focuses on discussing the outputs of TALe-Bus and their practical benefits for transit agencies. Finally, section 10 provides a brief summary of the key findings and benefits of the TALe-Bus dashboard.

2. General overview of the workflow and the dashboard framework

The TALe-Bus was developed as a decision support tool to help transit agencies manage the transition from conventional diesel-engine buses to battery-electric buses. TALe-Bus is an integrated, user-friendly, online dashboard (tool) that incorporates predictive modeling, interactive visualization, and data analytics. The dashboard has two main modules: one for the prediction of the energy consumption rate and another for fleet sizing and replacement factor (representing the ratio between the required electric bus fleet and the required diesel-engine fleet to offer the exact same transit service) estimation. The methodological framework of TALe-Bus is built based on statistical modeling, machine learning algorithms, and real data from the operations of electric buses to develop predictive models for energy consumption forecasting, fleet sizing, and replacement factor estimation. The dataset used in this study includes actual data from the operations of 60 battery-electric buses on 48 different bus routes in the city of Toronto between 2019 and 2021 (across different seasons). The dataset contains operational data of battery-electric buses from three different manufacturers: New Flyer, BYD, and Proterra. During the study period, the Toronto Transit Commission (TTC), which manages and operates the bus fleet in the City of Toronto, adopted an overnight depot charging strategy. This charging strategy is the most commonly adopted charging strategy in the early stages of transitioning to electric buses, given its lower costs, particularly when compared to on-route (opportunity) charging strategies.

The modeling pipeline of the dashboard began with data collection and preprocessing. The dataset used in this study includes detailed information about the operations of the buses and operational characteristics, route characteristics, vehicle specifications, environmental conditions, and energy consumption rates. Data preprocessing involved cleaning of missing data and ensuring data consistency across various bus models and seasonal conditions. After preprocessing the dataset, feature selection techniques are used to select the most important features to be used in the model. The feature selection techniques eliminate redundant information in the data set. For the purpose of unbiased validation of the developed models, it is essential to split the dataset into training and testing datasets. Thus, the data was divided into two datasets: one for training with 80% of the data and the other for testing with 20% of the data. A variety of predictive modeling techniques with different levels of complexity were developed, implemented, and evaluated in this study, including multiple linear regression, decision tree–based models, support vector regression, and artificial neural network–based models. For hyperparameter optimization, a comprehensive random search was used to identify the most effective models. Finally, the developed models were tested on the testing dataset, and their performance was evaluated using standard statistical measures, including RMSE, mean absolute error, and the coefficient of determination (R^2).

Following the development and validation of the predictive models, these models were integrated into the TALe-Bus interactive online dashboard to transform the analytical outputs into a practical planning and decision-support tool. The dashboard consists of two complementary modules. The first module is a geospatial visualization interface that displays route-level replacement factors on an interactive map and allows users to filter results by bus technology and season. The second module is a scenario-based prediction interface that enables users to specify operational, vehicle, and environmental conditions and obtain estimates of energy consumption, fleet requirements, and replacement factors. Figure 1 illustrates the operational architecture and information flow within the TALe-Bus dashboard rather than the detailed machine-learning model development process described earlier in this section.

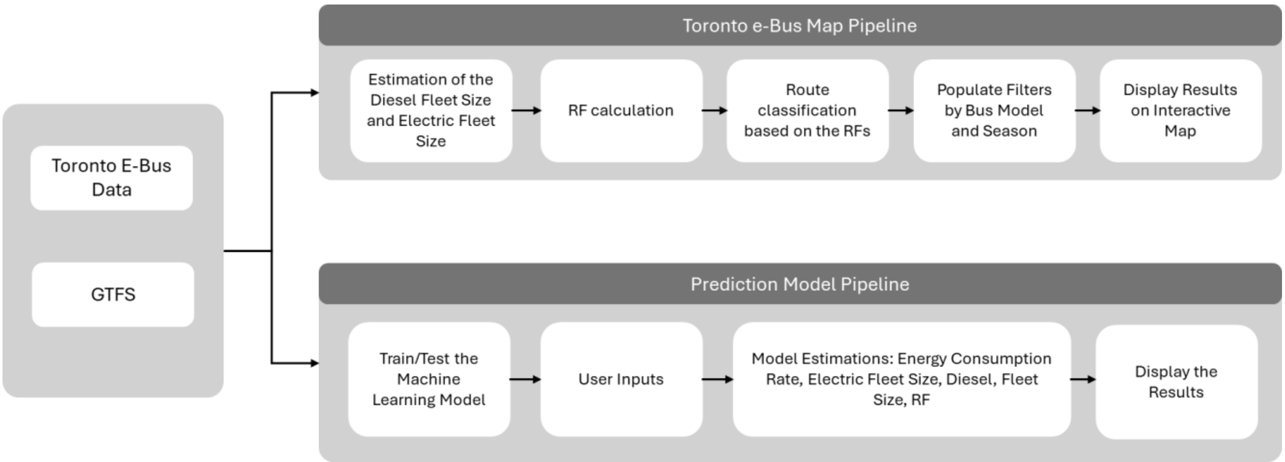

Figure 1. Overview of the workflow of the TALe-Bus dashboard.

3. Data source

In November 2017, the Board of the TTC approved the Green Bus Technology Plan. The plan developed a long-term procurement roadmap to transition the entire municipal bus fleet to zero-emission vehicles by 2040. The plan is consistent with the City of Toronto's TransformTO Action Plan, which represents the City of Toronto's overarching framework to reduce greenhouse gas emissions and counter climate change. Initially, the Green Bus Program focused on procuring electric buses from three manufacturers: BYD Canada, New Flyer Industries, and Proterra. In June 2019, 60 electric buses (25 New Flyer, 25 Proterra, and 10 BYD) were deployed and started offering transit service in the City of Toronto, positioning the TTC as the largest electric-bus operator in North America at that time. The technical specifications of these three bus technologies are presented in Table 1. A major strength of this deployment strategy lies in operating multiple bus technologies along identical routes and across different seasons, enabling direct and unbiased comparisons under comparable operational and environmental conditions, including variations in weather and temperature. These buses were deployed on 48 routes throughout the City of Toronto, and their energy consumption was systematically monitored to evaluate operational efficiency. This controlled evaluation framework ensured objective benchmarking and comparison of the different bus technologies and enabled the construction of a consistent empirical dataset, which was used to compare technology performance and to develop energy consumption prediction models. In addition to the energy consumption rates, the dataset also included operational variables (such as average route speed, stop density, and operational hours), vehicle specifications (such as battery capacity and curb weight), route specifications (such as route length and intersection density), and environmental variables (such as ambient temperature and average wind speed). The definitions of all variables used in this study to build the energy prediction model are summarized in Table 2. It should be noted that the energy consumption rates used in this study represent the average energy consumption (in kWh/km) per bus type and season for each bus route, enabling the models developed to be used in the planning stage. In addition, it should be noted that given the similarity in the energy consumption behavior of the different bus types as well as the environmental and operating conditions during the fall and spring seasons, data for the two seasons were merged together to minimize potential bias in the data and the developed prediction model. Including these data separately could artificially inflate the energy consumption observations under similar conditions and bias the model toward these scenarios. Thus, a total of 480 data points (collected for the different bus types across the different seasons while operating at different bus routes) were used for the development of the energy prediction model.

Table 1. Key characteristics of the three electric bus technologies deployed by the TTC in the City of Toronto.

Attribute	BYD K9M	New Flyer XE40	Proterra E2
Length	40 ft	40 ft	42.5 ft
Battery	360 kWh	400 kWh	440 kWh
Curb weight	31,750 lb	32,770 lb	29,849 lb
Frontal area	13,614 in^2	13,260 in^2	13,056 in^2
Engine power	200 HP	280 HP	338 HP

Table 2. Description of the variables used in this study.

Variable	Description
Route name	Identifier of the bus route as assigned by the TTC (excluded as an input for the energy consumption model).
Average speed	Average bus speed ($km \cdot h^{-1}$).
Number of Intersections	Total number of intersections along the full route, both directions (count).
Number of bus stops	Total number of designated bus stops along the route in both directions (count).
Cycle length	Total cycle length of the bus route across both travel directions (km).
Total distance travelled	Total distance covered daily by all buses operating on the route (km).
Number of cycles	Average number of round trips completed daily to meet service demand ($cycles \cdot day^{-1}$).
Stop spacing	Average spacing between two consecutive bus stops (m).
Average daily ridership	Average daily passenger boardings on the route ($passengers \cdot day^{-1}$).
Daily operational hours	Total daily operational hours accumulated by all buses on the route ($h \cdot day^{-1}$).
Bus model	Bus type, one-hot encoded into three categorical variables: *bus_model_NewFlyer*, *bus_model_BYD*, and *bus_model_Proterra*.
Bus age	Age of the buses in service at the start of the observed season (month).
Curb weight	Weight of the bus excluding passengers and cargo (lb).
Engine power	Bus engine power (HP).
Frontal area	Frontal cross-sectional area of the bus (in^2).
Battery capacity	Energy storage capacity of the bus battery (kWh).
Season	Seasonal factor, one-hot encoded into: *season_1* (fall/spring), *season_2* (winter), *season_3* (summer).
Average temperature	Average ambient temperature during the observation season (°C).
Average wind speed	Average wind speed during the observation season ($km \cdot h^{-1}$).
Net energy consumption	Average energy consumption rate per distance traveled per season per route ($kWh \cdot km^{-1}$).
Route location	The location of the bus route: 1 = route passes through downtown; 0 = route out of the downtown

4. Comparative energy performance across bus types and seasons

The three battery electric buses (BEB) evaluated in this study are different in terms of vehicle parameters as well as heating system configurations; thus, it is expected that there will be some differences in their energy consumption patterns. To analyze these differences, violin plots were generated (Figure 2a) to show the energy consumption distribution of each bus technology. In addition, given that ambient temperature has been widely recognized as a key determinant of electric bus performance, seasonal boxplots were also developed to assess how energy consumption varies across bus technologies under different climatic conditions, as presented in Figure 2a. Moreover, one of the most important features of BEB is the maximum bus range, which has different impacts on the required fleet size. The maximum bus range is mainly affected by the energy consumption rate and the battery capacity. Thus, the maximum bus range was estimated by dividing the bus battery capacity by the energy consumption rate, as shown in Eq (1). The resulting bus range distributions are summarized in the violin plots shown in Figure 2b.

$$Maximum\ Range\ (km) = \frac{Battery\ Capacity\ (kWh)}{Energy\ Consumption\ Rate\ (kWh/km)} \tag{1}$$

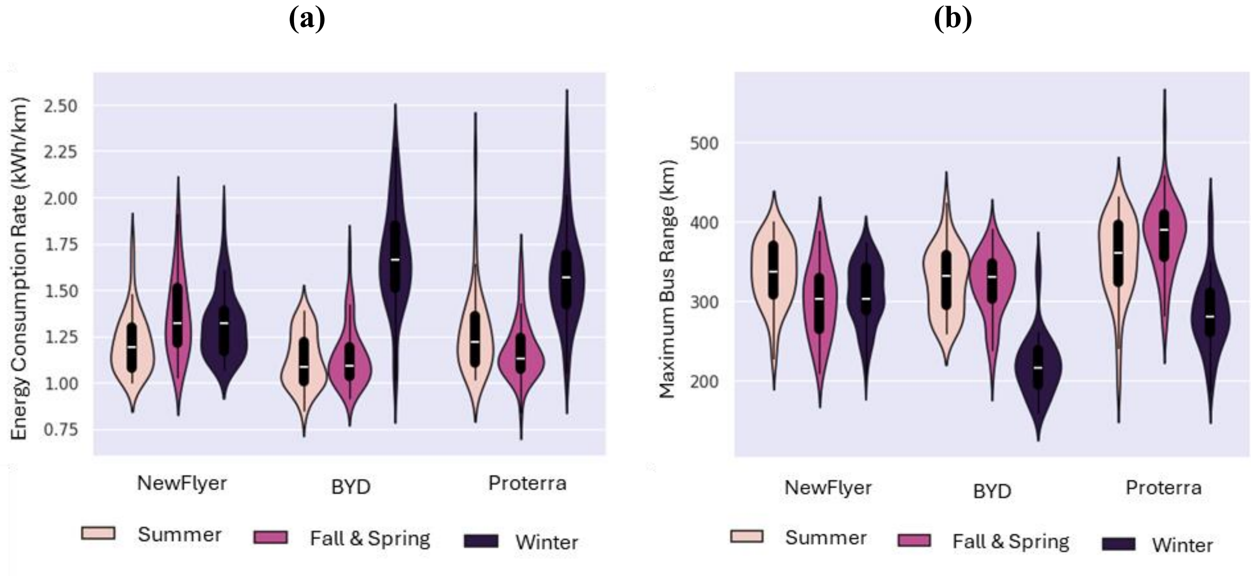

Figure 2. Performance of the different bus models across different seasons in terms of (a) energy consumption rate and (b) maximum bus range.

Comparing the energy performance of the three bus types reveals that there are considerable differences in the performance level and sensitivity to temperatures. In the summer, BYD has the lowest energy consumption, followed by Proterra and New Flyer, with similar performance and variability. In the fall and spring, BYD has the lowest energy consumption with the least variability, followed by Proterra, and then New Flyer with higher consumption and variability. In the winter, the lowest energy consumption with the least variability is observed in the case of New Flyer, followed by Proterra and then BYD. These results show that each bus technology behaves differently to temperature fluctuations. Notably, New Flyer buses display relatively stable energy consumption and traveling range across seasons, suggesting limited sensitivity to ambient temperature changes. Conversely, BYD buses show the largest energy variation and traveling range across the different seasons, indicating greater temperature sensitivity. The results clearly show that the energy consumption rate of BYD buses faces a large increase in the winter compared to the other seasons, which, in turn, reduces their traveling range. Similarly, Proterra buses show similar energy consumption patterns as BYD buses, with slightly higher bus ranges, given their larger battery capacities (BYD bus battery capacity = 360 kWh; Proterra bus battery capacity = 440 kWh). One key factor that influences the energy consumption pattern (variation) among the different types of buses is the heating system configuration. For instance, New Flyer buses rely on diesel auxiliary heaters, which reduces the energy required from the batteries in cold-weather conditions, while Proterra and BYD buses use their batteries for heating, which increases energy consumption in winter conditions. The observed increase in winter energy consumption for BYD and Proterra buses is consistent with findings reported in previous studies on battery electric buses operating in cold climates. Several studies have shown that winter energy consumption can increase substantially due to additional energy requirements for cabin heating, battery thermal management, and reduced battery efficiency at low temperatures. For example, Vehviläinen et al. [59] reported that the energy consumption of battery electric buses during winter can be 40%–45% higher than during summer, with cabin heating representing the primary contributor to the increase. Similarly, studies have identified auxiliary heating demand as one of the most influential factors affecting electric bus energy consumption in cold-weather operations. The results of the present study are consistent with these findings [60]. The relatively stable winter performance of the New Flyer buses can be attributed to their diesel auxiliary heating system, which reduces the heating load on the traction battery. In contrast, the BYD and Proterra buses rely on battery-powered heating systems, resulting in higher winter energy consumption and reduced operating range. The use of auxiliary diesel heaters to preserve battery range in cold climates has also been documented in the literature [61].

It should be noted that the heating-system configuration was not explicitly included as an input variable in the energy prediction model. Instead, the model incorporates both bus type and ambient temperature as predictor variables. Since the dataset contains observations for each bus technology under different seasonal and temperature conditions, the machine-learning model is able to implicitly capture technology-specific responses to temperature variations. Consequently, the influence of auxiliary heating systems is reflected indirectly through the historical energy-consumption patterns associated with each bus model rather than through an explicit representation of heating-system characteristics.

5. Feature selection

To identify the most relevant inputs for the predictive model, as well as eliminate redundant information, the Ward's hierarchical cluster method was used. This method is a variance-minimizing algorithm for agglomerative clustering, where each variable starts in its own cluster, and clusters are merged to minimize within-cluster variance. For each potential merge, the algorithm calculates the increase in the total sum of squared distances within the clusters and selects the merge that results in the smallest variance increase. This method groups variables based on similar statistical and functional characteristics, making it possible to reduce the dimensionality of the data and remove redundant variables by retaining only one representative variable from each cluster while retaining the majority of the information needed for the modeling process. The results of the clustering process are usually represented graphically through a dendrogram, which illustrates the cluster formation process. The horizontal axis of the dendrogram corresponds to the variables or clusters, while the vertical axis corresponds to the dissimilarity levels at which the clusters are merged. Then, a threshold value is specified to define the criteria for clustering or grouping the input variables. This threshold value is known as Ward's distance, and a value of 0.15 was set for this study based on literature recommendations [8]. From each cluster identified below this threshold value, only one variable can be selected as an input. The dendrogram developed using the input variables used in this study is shown in Figure 3, showing the different clusters and the cutoff threshold (Ward's distance threshold). For example, the first cluster consists of the average daily ridership, total distance traveled, and operational hours. From this cluster, only the average daily ridership was retained and used to develop the prediction model. In addition, another three clusters were identified and reduced to single variables. Cluster two consists of the number of intersections and stop spacing, from which the number of intersections was used. Cluster three consists of the average temperature and wind speeds, from which the average temperature was retained as the single representative variable. Finally, cluster four consists of the battery capacity, engine power, and frontal area, from which battery capacity was used.

It should be noted that the net energy consumption was included in the dendrogram solely for exploratory analysis purposes to evaluate its relationship with the explanatory variables. It was treated as the target variable throughout the modeling process and was not used as an input feature during feature selection or model training.

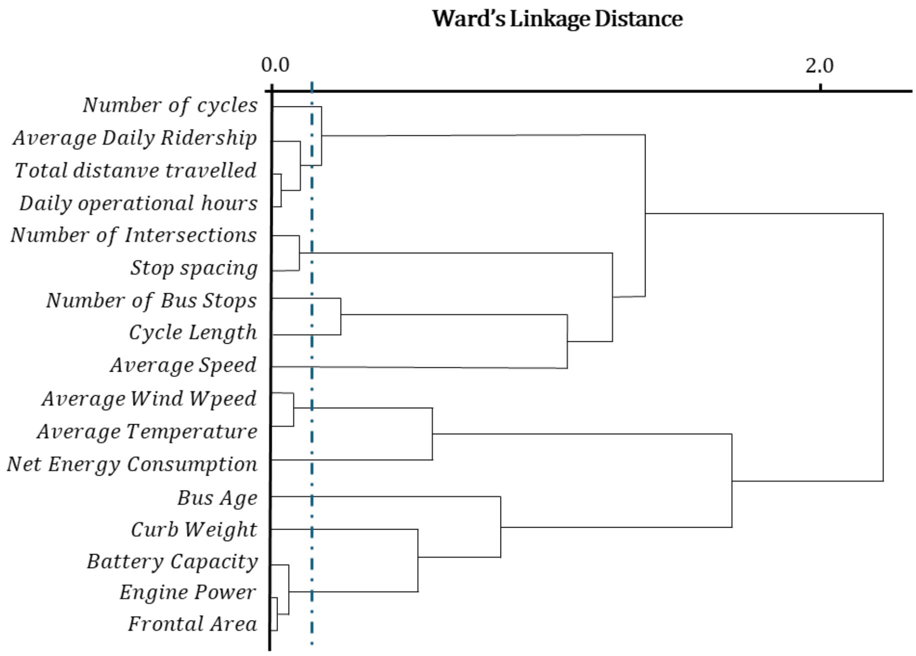

Figure 3. Dendrogram from Ward's hierarchical clustering illustrating the identified clusters.

6. Development of the energy prediction model

6.1. Comparing different modeling techniques

The existing literature reveals the absence of a comprehensive framework for the evaluation of data-driven approaches for estimating the energy consumption of electric buses and associated fleet size requirements. Although a large number of predictive models have been presented in the existing literature, the majority of the studies have been conducted using a single modeling technique without any specific justification for the choice of the technique, thereby raising a degree of ambiguity in the choice of the most appropriate technique. The present study aims to fill this gap by presenting a comprehensive benchmarking of the different modeling families for the prediction of energy consumption rate. A total of 11 models from the four most commonly used families were developed and tested: linear regression, tree-based, support vector regression, and artificial neural networks. The linear modeling family included ordinary least squares (OLS), ridge regression, Bayesian ridge regression, stochastic gradient descent regression, and automatic relevance determination regression. Tree-based methods included decision trees, random forests, gradient boosting, XGBoost, and LightGBM. Artificial neural networks were used to capture complex nonlinear relationships between the explanatory variables and energy consumption, while support vector regression was evaluated as a kernel-based nonlinear regression technique. These model families were selected because they represent a range of modeling complexities and have demonstrated strong performance in transportation and energy prediction applications.

The dataset was divided into two subsets: a training dataset containing 80% of the observations and a testing dataset containing the remaining 20%. To ensure a representative distribution of the different bus technologies, routes, and seasonal conditions while minimizing potential data leakage, a grouped sampling approach was adopted when constructing the training and testing datasets. This procedure ensured that the diversity of operating conditions represented in the full dataset was maintained across both subsets and provided a more reliable evaluation of model generalization performance. Then, the training dataset was used to develop the models, while the testing dataset was used to evaluate and compare the performance of the models on new (or unseen) data. For model evaluation, four statistical indicators were estimated for each model. The first measure was the RMSE, which represents the square root of the average of the squared differences between the predicted and observed values. The second measure was the coefficient of determination (R^2), which assesses the explanatory power of the target variable. The third measure was the mean absolute error (MAE), which represents the average of the absolute errors between the predicted and observed values. The fourth measure was the maximum absolute error, which represents the maximum difference between the predicted and observed values. In general, lower values of RMSE, MAE, and maximum error indicate higher predictive accuracy, while higher values of R^2 indicate higher explanatory capability. Figure 4 shows the performance of the different models on the testing dataset. The results clearly show that models of the tree-based modeling family outperform other approaches, showing better performance in terms of accuracy and capacity to recognize the complex patterns in the dataset. The results clearly show that tree-based models represent the most appropriate approach to model energy consumption in electric buses while maintaining balanced generalization capacity and preventing overfitting and underfitting phenomena. On the contrary, the performance of linear models is relatively low. This clearly shows that the energy consumption of electric buses in relation to input variables, such as vehicles, operations, routes, and environmental conditions, is nonlinear in nature. In addition, the MLP model shows the worst performance among all models, indicating overfitting, given that the dataset used in this study is relatively small for such a complex modeling approach, which requires large data for training. Finally, the results show that the random forest model outperforms all other models; thus, this technique was selected to build the final predictive model that will be used in the TALe-Bus dashboard after hyperparameter tuning through a comprehensive search process.

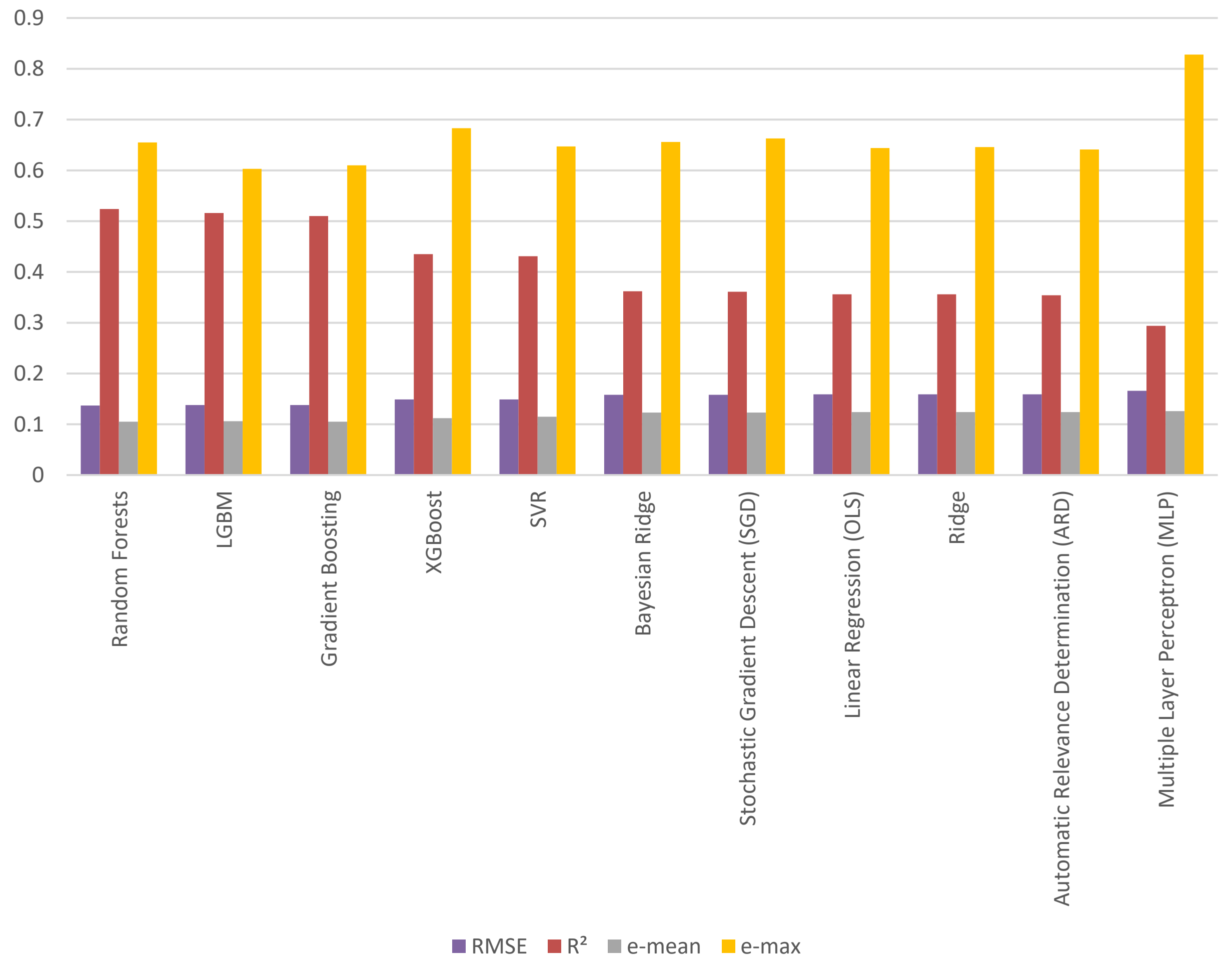

Figure 4. Performance comparison of different modeling approaches on the testing dataset.

6.2. Final energy prediction model

The previous section identified the random forest (RF) model as the most efficient technique in forecasting the energy consumption of electric buses. Based on this, the aim of this section is to build the final RF model using the optimal set of hyperparameters obtained from an extensive search process. Six key hyperparameters were systematically evaluated: the number of estimators (n_estimators), which defines the number of trees in the ensemble; the minimum number of samples needed to split an internal node (min_samples_split), which controls the tradeoff between over- and underfitting; the minimum number of samples in each leaf node (min_samples_leaf), which determines the leaf node size and controls the level of complexity; the maximum number of features to consider at each node (max_features), which incorporates randomness to the model; the maximum depth of the tree (max_depth), which determines the level of complexity; and the bootstrap option, which defines the use of bootstrap sampling to increase the robustness of the model. A comprehensive hyperparameter search process was conducted to evaluate model performance under different parameter combinations. The optimal parameters used in the final model and the tested values are as follows:

- n_estimators: 100 (search: 5, 20, 50, 100, 500)
- min_samples_split: 6 (search: 2, 6, 8, 10)

- min_samples_leaf: 1 (search: 1, 2, 3, 4, 5)
- max_variables: sqrt (search: auto, sqrt)
- max_depth: 60 (search: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120)
- bootstrap: True (search: True, False)

Using these hyperparameters, the final RF model was developed. The model was able to predict energy consumption with a RMSE of 0.13 kWh/km, a mean absolute error of 0.11 kWh/km, a maximum absolute error of 0.32 kWh/km, and an R^2 value of 0.7. Figure 5 shows a comparison between observed and predicted energy consumption values for the testing dataset by using a 45° line. In addition, to provide context for these error metrics, the average energy consumption rate in the dataset was 1.35 kWh/km. Accordingly, the RMSE and MAE correspond to approximately 9.6% and 8.1% of the average energy consumption rate, respectively. These results indicate that the model is capable of capturing the primary relationships between operational, vehicle, route, and environmental characteristics while maintaining a level of prediction accuracy suitable for fleet planning and scenario analysis applications.

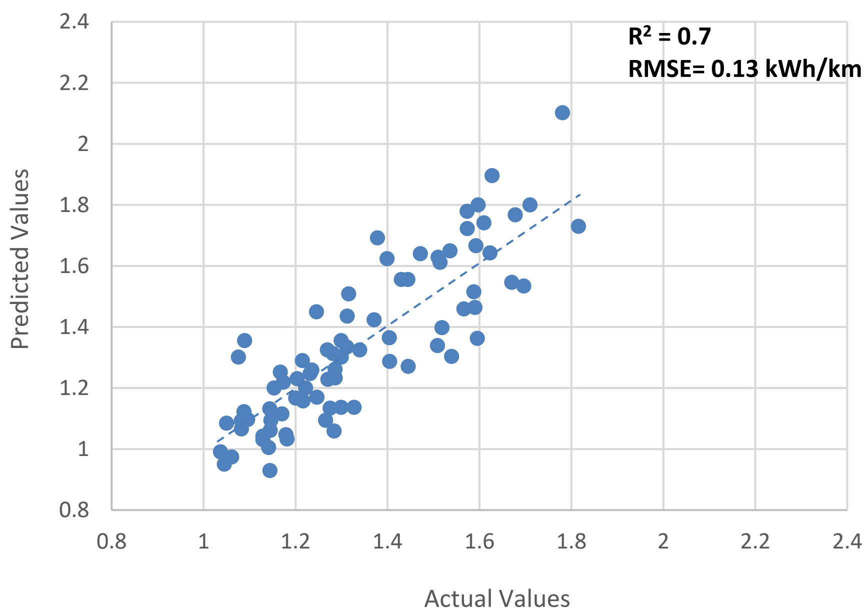

Figure 5. Observed vs. predicted values for the final random forest model on the testing dataset.

7. Fleet sizing and replacement factor estimation

The process of estimating the required electric bus fleet size to offer transit service is a multidimensional, complex process that goes beyond the scope of traditional fleet replacement. Unlike the case of diesel-engine buses, electric buses have limitations in terms of battery capacity, charging requirements, infrastructure, and variability of energy consumption. These limitations have a direct impact on their operational capability and scheduling flexibility. In order to address the issue of estimating the electric bus fleet size, a fleet-sizing estimation module was designed to assess the operational feasibility of the substitution of diesel-powered buses with electric-powered buses without compromising the level of service (offering the same transit service). This module integrates energy modeling, operational planning, and vehicle performance analysis to generate several key planning outputs, including the required electric fleet size, diesel fleet equivalency, replacement factor (representing the ratio between the required electric bus fleet and the required diesel fleet), maximum bus range, and fleet expansion requirements.

7.1. Diesel fleet size

The fleet size of buses that is needed to provide transit service along any particular route is primarily a function of the service headway along that particular route, as well as the operational characteristics of the buses that operate along that route. The service headway is usually determined in response to the passenger demand at the point of maximum load, in line with the formulation provided by Vuchic [62], as described in Eqs (2) and (3). The transit agencies then announce these headways in their service schedules. For Toronto, the Toronto Transit Commission publishes its schedules for each sign-up period, with the service frequency depending on the seasons. Once the headways have been estimated for any particular route, they are used to calculate the fleet size of buses that is needed to offer the service on that route by dividing the total cycle time by the headway, as expressed in Eq (4). The cycle time is defined as the sum of the running time and the terminal layover time, as per Eq (5). The running time can be determined from past operational historical data. However, it can also be estimated by making use of the average speed of the buses that operate along that route, as presented in Eq (6). These equations are widely adopted by transit agencies to determine the number of diesel-engine buses required to offer transit service on a certain route during a specified time interval. Using the published TTC headways and observed operating speeds, the fleet size for each route was computed, and the resulting values were then validated against the actual fleet deployment reported by the TTC, showing that both values match for all routes.

$$f = \frac{PHF * PD_{CLS}}{N * CC_v * \alpha} \tag{2}$$

$$h = \left[\frac{60}{f}\right]^{-} \tag{3}$$

$$F_{diesel} = \left[\frac{T_{cycle}}{h}\right]^{+} \tag{4}$$

$$T_{cycle} = T_{operational} + T_{terminal} \tag{5}$$

$$T_{operational} = \frac{L}{S} \tag{6}$$

where f denotes the service frequency; PD_{CLS} represents the passenger demand observed at the critical maximum-load segment of the route; PHF is the peak-hour factor reflecting temporal fluctuations in demand within the peak hour; α refers to the design load factor specified in service design standards, typically set below 1; N represents the number of vehicles per transit unit (equal to one for conventional bus operations); CC_v corresponds to the crush capacity of an individual vehicle; h represents the service headway; F_{diesel} is the required diesel-engine fleet size; T_{cycle} is the total cycle time; $T_{operational}$ is the operating time; $T_{terminal}$ is the terminal time; L represents the cycle length; S is the average operating speed; $[\ ]^{-}$ represents the floor operator (round down to the nearest integer), and [ ]$^+$ represents the ceiling operator (round up to the nearest integer).

7.2. Electric fleet size

For electric bus operations, the fleet size required on a given route during a specified time interval must satisfy passenger demand (following Eqs (2)–(5)) while also accounting for vehicle range limitations. In this context, both the energy consumption rate and battery capacity are critical in determining the number of vehicles needed to maintain the same level of service. The average energy consumption rates estimated through the prediction model were used to evaluate the fleet requirement from an energy perspective. The first step in this process focuses on estimating the energy consumed per operating cycle on the bus route by multiplying the route cycle length by the average energy consumption rate, as shown in Eq (7). Based on this value, the number of cycles that a bus can complete on a single charge was then derived using the battery capacity, as shown in Eq (8). In practice, transit agencies and manufacturers advise against fully depleting the battery in order to maintain a safety buffer. A minimum state of charge of 15% is commonly recommended; therefore, this study assumes that only 85% = α of the nominal battery capacity is utilized during operations. It should be noted that the calculations and approach followed in this study focus on overnight depot charging, which is the charging strategy applied in the City of Toronto (given that the dashboard is built based on this case). Under this approach, a bus remains in service until completing the number of cycles that reduce the battery to the minimum allowable state of charge (that is, larger than or equal to 15%), after which it is withdrawn from service and returned to the depot for overnight recharging before resuming operation the following day. Accordingly, the fleet size required from an energy perspective was determined by dividing the total number of daily cycles completed by the entire fleet serving a bus route by the maximum number of cycles that can be completed by a bus in one charge, as highlighted in Eq (9). The final electric fleet size was then defined as the larger of two values [as shown in Eq (10)]: (1) the fleet required to meet passenger demand as estimated using conventional diesel-based planning formulations (from Eqs (2)–(6)) and (2) the fleet necessary to satisfy electric-bus range constraints.

$$EC_{cycle} = L * EC_{avg} \tag{7}$$

$$NC_{bus} = \left[\frac{0.85*BC}{EC_{cycle}}\right]^{-} \tag{8}$$

$$F_{energy} = \left[\frac{NC_{route}}{NC_{bus}}\right]^{+} \tag{9}$$

$$F_{electric} = \max\begin{cases} F_{diesel} \\ F_{energy} \end{cases} \tag{10}$$

where EC_{cycle} represents the energy consumption per cycle in kWh; EC_{avg} is the average energy consumption rate in kWh/km; L represents the cycle length in km; NC_{bus} stands for the number of cycles that a bus can complete utilizing its capacity in a certain bus route; BC represents the bus battery capacity (kWh); NC_{route} is the number of daily cycles the bus fleet needs to complete to offer transit service at a certain bus route (cycle); F_{energy} represents the required fleet size needed to offer transit service given the limitation on the bus range (bus); $F_{electric}$ is the number of electric buses needed to offer transit service using electric buses (bus); [ ]$^-$ represents the floor operator (round down to the nearest integer); and [ ]$^+$ represents the ceiling operator (round up to the nearest integer).

The fleet-sizing methodology presented in this study assumes an overnight depot charging strategy, whereby buses are charged only at the depot between service days. However, the framework can be extended to accommodate opportunity charging or mixed charging strategies. In such cases, the effective energy available to a bus would consist of both the initial usable battery capacity and any energy replenished during intermediate charging events throughout the day. Consequently, the maximum number of route cycles that can be completed by a vehicle before requiring replacement would increase, potentially reducing the energy-constrained fleet requirement. This extension could be incorporated by modifying Eqs (8) and (9) to account for the cumulative energy supplied through opportunity charging events. Future versions of the dashboard could incorporate charging frequency, charging duration, charger power, and charger location as additional inputs to support the evaluation of alternative charging strategies. For example, under opportunity charging, the effective available energy could be represented as the sum of the usable battery capacity and the energy received from intermediate charging events during the service day. The resulting increase in available energy would increase the number of route cycles that a bus can complete before reaching its minimum state of charge and consequently reduce the fleet requirement associated with range limitations.

7.3. Replacement factor estimation and fleet size comparison

In addition to estimating the required electric and diesel fleets needed to offer transit service, the dashboard also calculates the replacement factor, defined as the ratio between the number of electric buses required and the number of diesel buses needed to offer the exact same transit service [as shown in Eq (11)]. This metric provides a concise and interpretable indicator of the relative fleet requirements when transitioning from conventional diesel technology to battery-electric buses under comparable service conditions. The replacement factor is particularly important because it captures the combined influence of operational constraints, vehicle energy performance, and battery limitations in a single value. Unlike simple fleet counts, this indicator reflects how differences in vehicle range, charging requirements, and energy consumption affect the number of buses needed to maintain the same level of service compared to diesel-engine buses (that have no range limitation). As such, it provides a realistic representation of electrification impacts on fleet planning. Moreover, the replacement factor offers a direct and practical basis for comparing electric and diesel fleet needs. Transit agencies can use this value to quickly assess whether electrification will require fleet expansion or not under specific operational conditions. This facilitates strategic decision-making related to capital investment, depot and charging infrastructure planning, operational scheduling, and long-term electrification planning. In addition, the replacement factor supports scenario-based analysis by allowing planners to evaluate how improvements in battery capacity, charging strategy, or operational efficiency may reduce the number of required electric buses over time. Consequently, it serves not only as a comparative indicator but also as a planning and policy-support tool that helps transit agencies quantify the operational implications of fleet electrification.

$$RF = \frac{F_{electric}}{F_{diesel}} \tag{11}$$

where RF represents the replacement factor.

It should be noted that the replacement factor cannot be less than one. This follows directly from Equation (10), where the required electric fleet size is defined as the maximum of the fleet required to satisfy passenger demand and the fleet required to satisfy electric-bus range constraints. Since the passenger-demand fleet requirement is equivalent to the diesel fleet size needed to provide the same service frequency and passenger capacity, the electric fleet size is always greater than or equal to the diesel fleet size. Consequently, replacement factor (RF) = 1 indicates that electric buses can directly replace diesel buses on a one-to-one basis without fleet expansion, while RF > 1 indicates that additional electric buses are required to maintain the same level of service.

Ultimately, a key planning question for transit agencies is whether electrification will require fleet expansion. Accordingly, the dashboard is designed to address this issue directly by comparing both electric and diesel fleet sizes and providing some information on whether electrification of a certain route under certain input conditions would require fleet expansion or not.

7.4. Maximum bus range

In overnight depot charging systems, electric buses depend on a single charging session during off-service hours to supply sufficient energy for the entire next day of operation. Under this operating strategy, accurately estimating the maximum bus range is essential to ensure service reliability, proper fleet allocation, and efficient scheduling. The proposed module computes the maximum range by dividing the portion of the bus battery capacity that can be utilized (85%) by the energy consumption rate, as shown in Eq (12). This formulation reflects operational rather than theoretical performance (reported by bus manufacturers), as the usable battery capacity and energy consumption rate are derived from real-world conditions. Thus, the estimation explicitly incorporates the effects of passenger load, ambient temperature and weather, Heating, ventilation, and air conditioning (HVAC) and auxiliary energy demands, stop frequency, etc. These factors significantly influence daily energy consumption and therefore determine the realistic distance a bus can travel before reaching its minimum allowable state of charge (15%). For overnight depot charging, the estimated maximum range is a critical determinant of service feasibility. Since buses are not recharged during the day, the calculated range must be sufficient to allow the bus to complete its assigned daily block. This includes the ability to perform multiple route cycles and maintain continuous operation during peak service periods. Overestimating the range may lead to in-service energy depletion and operational disruptions, while underestimating it may result in conservative planning, unnecessary fleet expansion, and reduced vehicle utilization. Moreover, maximum range estimation directly supports fleet planning and operational optimization. It enables transit agencies to evaluate whether specific routes are compatible with depot-only charging, assign appropriate service blocks to electric buses, and verify that battery capacity and consumption characteristics can sustain reliable daily operations across different seasonal and operating conditions. Thus, to enhance decision-making, the TALe-Bus dashboard presents the maximum operational range as a primary output indicator. This allows planners and operators to quickly assess whether scheduled service blocks fall within safe operating limits and to understand how variations in energy consumption or usable battery capacity influence daily service capability. By translating complex energy performance calculations into a clear and interpretable metric, the dashboard facilitates informed planning and reliable deployment of electric buses.

$$MBR = \frac{0.85 * BC}{EC_{avg}} \tag{12}$$

where MBR represents the maximum bus range (km); BC stands for the bus battery capacity (kWh);

and EC_{avg} is the average energy consumption rate (kWh/km)

8. Dashboard architecture and system integration

The TALe-Bus Dashboard was designed as a cloud modular, scalable, web-based application integrating predictive modeling, dynamic visualization, and user input processing. The architecture follows a model-view-controller architectural pattern that comprises three core layers:

- Data and model layer (model): Python-based backend implementing the trained random forest model to predict the energy consumption rate and estimate the required electric fleet, diesel fleet, and replacement factor.
- Application logic layer (controller): User queries, model inference, and database access orchestration.
- Visualization and user interface layer (view): Web technologies (HTML5, CSS, Leaflet.js) for interactive maps and form-based user experiences on the frontend.

The user inputs (route characteristics, bus model, seasonal selection) are processed in real time, and the parameters are passed to the backend models to predict energy consumption rates and estimate required fleet sizes and replacement factors. The visualization layer then dynamically updates the citywide map or displays the custom predictions based on the user's scenario. The dashboard was deployed on a cloud-based server with SSL encryption.

The final user interface was designed with a focus on accessibility, simplicity, and responsiveness. The goal was to allow transit professionals with varying levels of technical expertise to use the dashboard efficiently without requiring specialized training (https://talebusv2.ue.r.appspot.com/). The final dashboard is structured into two main interactive modules (Toronto eBuses and Prediction), and the user is capable of selecting the preferred module from the landing page, as shown in Figure 6, which shows the dashboard landing page. The first module (Toronto eBuses) focuses on visualizing electric bus data from the City of Toronto, as shown in Figure 7, and it consists of the following tools:

- Dynamic map visualization: An interactive GIS map displays Toronto bus routes color-coded based on the computed replacement factor. Green routes (RF = 1) represent routes suitable for direct one-to-one diesel-to-electric replacement, while red routes (RF > 1) represent routes requiring expanded electric fleets.
- Route filtering: Users can filter routes by selecting specific bus models (BYD, New Flyer, Proterra) and seasons (winter, summer, spring/fall). This visualization enables transit agencies to rapidly identify routes that are most suitable for electrification and assess the spatial distribution of electrification challenges across the transit network. By providing an intuitive representation of fleet replacement requirements, the dashboard supports strategic planning, route prioritization, and investment decision-making. The replacement factor visualization provides a practical decision-support tool for transit agencies planning fleet electrification programs. Routes displayed in green (RF = 1) can be considered strong candidates for early electrification because they can be converted without requiring additional vehicles, thereby minimizing capital investment and operational disruption. Conversely, routes displayed in red (RF > 1) highlight corridors where range limitations may necessitate fleet expansion, charging infrastructure enhancements, operational schedule adjustments, or future battery technology improvements. By enabling planners to quickly distinguish between these route categories across an entire transit network, the dashboard supports phased electrification strategies, prioritization of investment decisions, and identification of routes requiring detailed operational

assessment before deployment.

- Route details popup: Clicking on any route provides a detailed view of its predicted energy consumption rate, required electric fleet size, and replacement factor.

The second module allows users to define a range of parameters, including route characteristics, passenger demand profiles, service frequency, and environmental conditions such as temperature and weather. By inputting these scenario-specific variables, transit agencies can evaluate the impact of different operational strategies on key performance metrics. Specifically, the module calculates net energy consumption for the selected scenarios using the developed energy prediction model. Additionally, the module estimates the required fleet sizes for electric buses (as well as for diesel-engine buses) to maintain service levels under the scenario conditions, taking into account constraints such as battery capacity and maximum bus range. A further output of this module is the replacement factor, which quantifies the number of electric buses needed to replace a given diesel fleet while maintaining equivalent service levels. Moreover, the maximum bus range is also provided as one of the outputs of the dashboard. By providing a scenario-based analysis, this module not only supports short-term operational planning but also facilitates strategic decision-making for fleet electrification. Agencies can test “what-if” scenarios, such as introducing new routes, extending service hours, or deploying buses in extreme weather conditions, and immediately observe the resulting implications for energy consumption and fleet requirements. The outputs are designed to be interpretable and insightful, offering decision-makers insights into whether a particular operational change or electrification plan would necessitate fleet expansion, additional charging infrastructure, or adjustments to service frequency. Overall, this scenario-based interface empowers transit agencies to move beyond generic estimates, enabling data-driven planning that aligns with operational realities and supports efficient electrification of transit services. This module focuses on allowing the user to input the route, vehicle, and operational characteristics and external conditions, as shown in Figure 8. This module consists of the following tools:

- Input forms: Users input operational characteristics, vehicle specifications, operational characteristics, and external conditions.
- Instant prediction display: Upon submission, the dashboard displays predicted energy consumption, required fleet sizes, maximum eBus range, replacement factor, and whether the electrification of this route would result in fleet expansion or not.

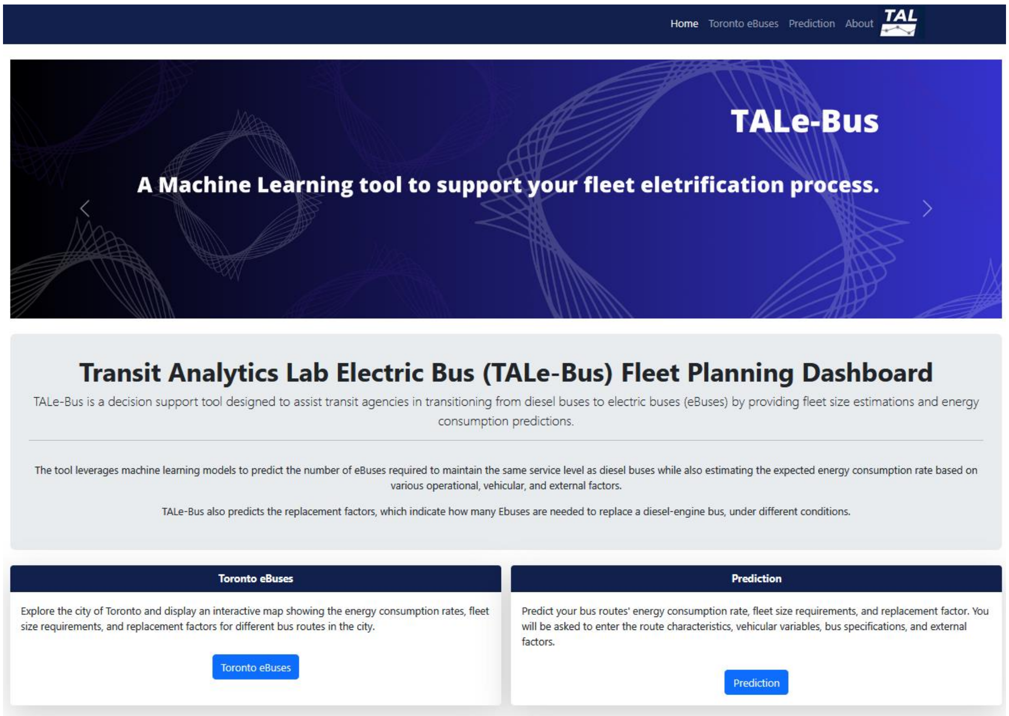

Figure 6. Landing page of the TALe-Bus dashboard.

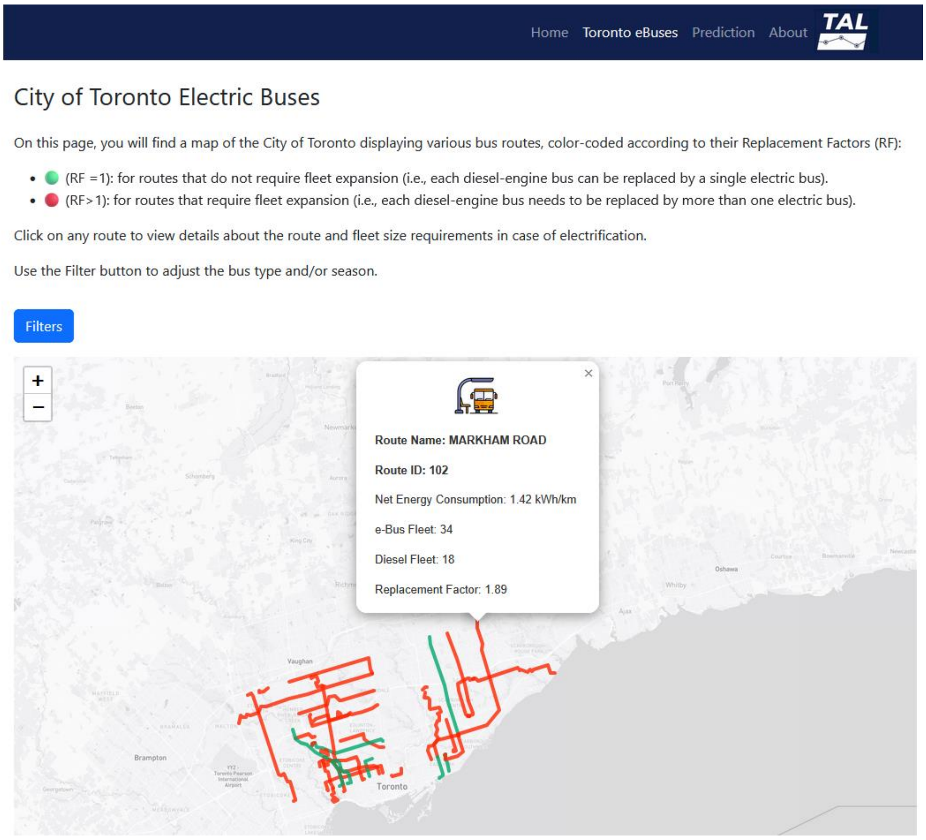

Figure 7. Screenshot of the Toronto eBuses module from the TALe-Bus dashboard.

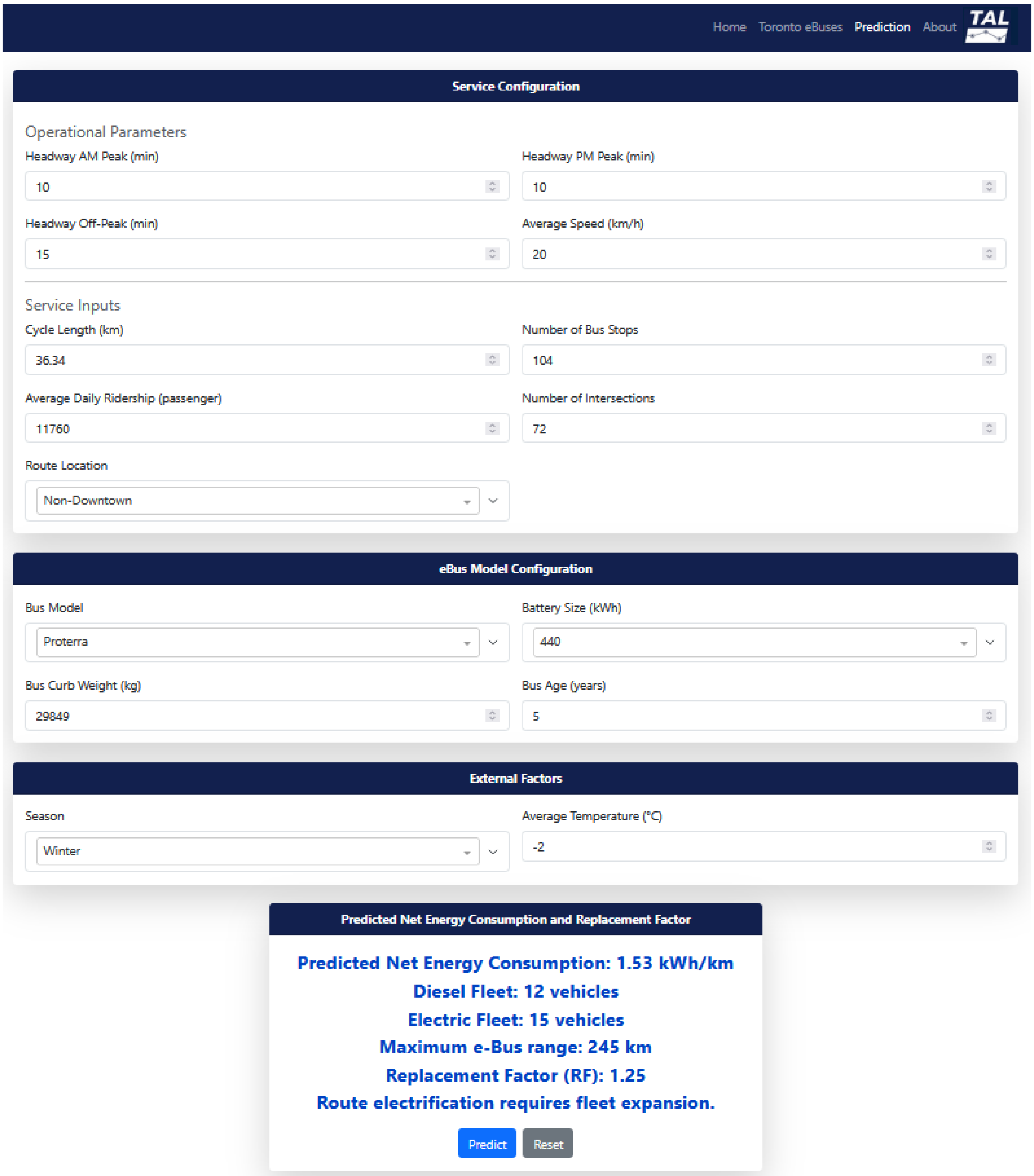

Figure 8. Screenshot of the Prediction module from the TALe-Bus dashboard, showing the inputs and outputs.

9. Discussion of the TALe-Bus outputs and their benefits

The results generated by the TALe-Bus Dashboard provide valuable insights for transit agencies during the transition to electric bus fleets by offering data-driven estimates of energy consumption, fleet requirements, replacement factors, and maximum bus range under different operational and environmental conditions. Each of these components provides distinct benefits that collectively strengthen transit agencies planning, operational efficiency, and risk mitigation strategies during the transition from diesel to electric buses. The energy consumption prediction model demonstrated good performance, capturing the influence of operational, route, vehicular, and environmental factors on electric bus energy use, which offers several practical benefits. First, the ability of route-specific and seasonally adjusted energy consumption rates enables transit agencies to more accurately forecast operational energy demands. Unlike static or manufacturer-reported values (representing the maximum bus range at favorable conditions), which often underestimate real-world variability, TALe-Bus predictions incorporate dynamic inputs such as bus speeds and ambient temperature. This ensures that planners have a realistic understanding of daily and seasonal energy requirements, which is particularly critical in cold-weather environments like Toronto, where heating energy demands substantially impact the energy consumption rate. Second, these predictions support effective energy management and charging infrastructure planning. Knowing the expected daily energy consumption rates allows agencies to better design depot charging schedules, size charging equipment appropriately, and forecast electricity demand peaks.

Fleet sizing estimation provides agencies with a critical tool for evaluating the operational feasibility of electric buses in maintaining current service levels. Accurate fleet size predictions help agencies avoid two major risks: fleet underestimation, which could result in missed service targets and public dissatisfaction, and fleet overestimation, which could unnecessarily inflate capital and operational costs. By explicitly integrating energy consumption constraints into fleet sizing calculations, the TALe-Bus Dashboard moves beyond the traditional diesel-based planning methodologies. The dashboard accounts for the reduced daily range under conditions of high auxiliary load or poor weather, ensuring that the fleet plans are resilient to real-world performance limitations. Moreover, the fleet sizing tool also supports strategic asset management. Agencies can identify routes where single diesel-engine buses can be replaced by a single electric bus (routes with replacement factor = 1) and prioritize them for early conversion. Conversely, routes requiring additional vehicles can be flagged for further infrastructure investment (e.g., opportunity charging stations) or vehicle technology upgrades (e.g., buses with larger battery capacities or different battery technology). Finally, the ability to estimate fleet requirements under different scenarios (changing bus models, changing bus routes, or seasonal conditions) enables planners to conduct sensitivity analyses and identify which bus models fit the deployment at certain bus routes.

The replacement factor (RF) offers a concise, intuitive metric that captures the operational implications of fleet electrification in a manner that is easily understood by both technical staff and decision-makers. The RF metric simplifies complex technical analyses into a single value indicating whether fleet expansion is necessary. This interpretable value supports budgeting justification and public reporting, where accessible metrics are critical for securing funding and community support. Replacement factor predictions also enable route prioritization based on electrification feasibility. Routes with RF = 1 can be prioritized during the transitioning phase to minimize transitioning costs, as these routes do not require fleet expansion. In addition, the maximum bus range is one of the outputs provided by the TALe-Bus dashboard. Understanding the maximum usable range under varying conditions is crucial for multiple aspects of electrification planning. First, estimating the maximum bus range enables route-to-vehicle matching. Agencies can ensure that assigned buses have sufficient range to complete assigned routes without the need for midday recharging, reducing operational complexity and avoiding service disruptions. Second, maximum range insights guide procurement strategies. For instance, if certain routes demand higher energy consumption due to operational characteristics, the procurement process can prioritize buses with larger battery capacities or more efficient thermal management systems.

The integration of energy consumption prediction, fleet size and replacement factor estimation, and maximum bus range assessment within a single dashboard offers multiple intertwined benefits. Each component informs and reinforces the others, creating a holistic decision-support environment. For instance, understanding that a route's winter energy consumption rate rises sharply enables planners to anticipate fleet expansion and plan accordingly. Similarly, maximum range insights derived from energy consumption estimates feed directly into fleet sizing calculations and charging infrastructure planning. By unifying these analytics within an interactive, user-friendly interface, TALe-Bus empowers transit agencies to move beyond reactive planning and toward proactive, data-driven strategic management of their electrification transitions. Users can explore multiple scenarios, adjust assumptions dynamically, and evaluate impacts, all of which strengthen the transition planning to electric buses.

The demonstrated capabilities of the TALe-Bus Dashboard extend beyond immediate operational benefits to broader strategic and societal implications. By reducing uncertainty around key electrification challenges, the tool accelerates the transition to low-carbon public transit systems. Accurate fleet planning minimizes the risks of underperformance, service disruptions, and financial inefficiencies that could otherwise hinder public support for electrification initiatives. Additionally, the dashboard offers access to complex predictive analytics. Thus, the dashboard allows transit agencies the ability to conduct sophisticated analyses traditionally requiring specialized consulting support. This empowerment promotes equity across transit agencies of varying sizes and resource levels, ensuring that sustainable transit transformation is accessible to a wider range of communities.

While the TALe-Bus dashboard aids transit agencies in the transition from diesel fleet to electric fleet, it should be noted that the dashboard was built based on certain assumptions. Thus, the following limitations should be considered when applying the tool in transit planning contexts:

- The modeling framework assumes that all electric buses rely exclusively on overnight depot charging and does not incorporate on-route fast-charging strategies. Consequently, buses are required to complete their scheduled service before returning to the depot for charging. This assumption may influence fleet size estimates, particularly in systems where opportunity charging is available and can extend operational range or reduce fleet requirements.
- Temperature is explicitly incorporated as a model input to capture seasonal variability, reflecting Toronto's four-season climate. While this enhances prediction reliability under similar seasonal patterns, the model may require recalibration for locations experiencing extreme climates, limited seasonal variation, or substantially different temperature ranges.
- The training of the energy model was done based on data from three electric bus models deployed in Toronto: BYD K9M, New Flyer XE40, and Proterra E2. Differences in battery technology, vehicle design, auxiliary loads, and heating systems across other bus types may influence energy consumption and operational performance. Therefore, predictions may vary when applied to fleets composed of different electric bus technologies.
- In addition, the feature-selection results obtained in this study are based on the characteristics of the Toronto dataset and should not be interpreted as universally applicable to all transit systems. The relative importance of explanatory variables may vary across cities due to differences in operating practices, service scheduling, passenger demand patterns, route characteristics, climatic conditions, and fleet technologies. For example, variables identified as highly influential in the Toronto case study may have a different level of importance in transit systems operating under different conditions. Therefore, when applying the framework to other cities, both model recalibration and feature re-evaluation should be

considered to ensure that the selected predictors appropriately reflect local operational realities.

Given the above constraints, the dashboard outputs should be validated using local operational and environmental data prior to implementation. Transit agencies operating under different climatic conditions, charging strategies, fleet compositions, or service configurations may require model recalibration (retraining or fine-tuning) to ensure reliable and context-specific predictions.

10. Conclusions

This study presented the development, validation, and application of the TALe-Bus Dashboard, an interactive decision-support tool designed to support transit agencies in planning the transition from diesel-engine buses to battery-electric buses. The dashboard integrates machine learning–based energy consumption prediction, fleet sizing analysis, replacement factor estimation, and interactive visualization within a unified web-based platform. Using real-world operational data collected from 60 battery-electric buses operating on 48 routes in Toronto, several statistical and machine learning techniques were evaluated for energy consumption prediction. Among the tested models, the random forest model achieved the best performance, yielding an RMSE of 0.13 kWh/km, a mean absolute error of 0.11 kWh/km, and an R^2 value of 0.70 on the testing dataset. The analysis also demonstrated the significant influence of seasonal conditions on electric bus performance, with winter operations exhibiting noticeably higher energy consumption for bus technologies relying on battery-powered heating systems, while buses equipped with auxiliary diesel heaters exhibited more stable energy consumption patterns across seasons.

Through the integration of predictive machine learning models, dynamic visualization, and user-centered design principles, the dashboard addresses key operational and planning challenges associated with fleet electrification and provides valuable insights to transit agencies. The energy consumption prediction model demonstrated strong capability in capturing the influence of operational characteristics, route characteristics, vehicle attributes, and environmental conditions on electric bus energy consumption. This capability enables transit agencies to develop realistic estimates of daily energy requirements, which are critical for charging infrastructure planning, vehicle procurement, and operational scheduling. By providing route-specific energy consumption estimates that explicitly account for seasonal variability, the dashboard empowers agencies to make informed decisions that reflect the realities of real-world operating conditions. Moreover, the fleet sizing and replacement factor estimation modules translate energy consumption predictions into practical fleet-planning metrics. By quantifying the additional fleet required to maintain service levels under different seasonal and technological scenarios, the dashboard supports risk mitigation, budget optimization, and service reliability. The replacement factor framework enables transit agencies to directly assess whether electrification of a particular route would require fleet expansion, while the interactive route visualization supports the identification of routes that are most suitable for early electrification deployment. The scenario-based prediction module further enables planners to evaluate the impacts of changes in route characteristics, operating conditions, vehicle technologies, and seasonal factors on fleet requirements and operational performance.

Overall, TALe-Bus provides transit agencies with an accessible and data-driven planning tool that supports vehicle procurement decisions, charging infrastructure planning, route prioritization, and long-term electrification strategies. By combining robust predictive analytics with a user-friendly and interactive interface, the dashboard bridges the gap between data, prediction, and decision-making,

making sophisticated planning tools accessible to a broader range of transit professionals. As the electrification of transit fleets continues to advance, tools such as TALe-Bus can play an important role in supporting resilient, efficient, and evidence-based fleet electrification planning. Future research could extend the framework by incorporating opportunity charging and mixed charging strategies, quantifying uncertainty in fleet size and replacement factor estimates, integrating battery degradation indicators, and validating the methodology using datasets from additional transit agencies operating under different climatic, operational, and service conditions. Such enhancements would further improve the transferability, robustness, and planning value of the TALe-Bus framework.

Use of AI tools declaration

AI tools were used for the aim of polishing and improving the paper writeup.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

Conceptualization, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; methodology, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; software, Kareem Othman and Diego Da Silva; validation, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; formal analysis, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; investigation, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; resources, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; data curation, Kareem Othman and Diego Da Silva; writing—original draft preparation, Kareem Othman and Diego Da Silva; writing—review and editing, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; visualization, Kareem Othman, Diego Da Silva, Amer Shalaby and Baher Abdulhai; supervision, Amer Shalaby and Baher Abdulhai; project administration, Amer Shalaby and Baher Abdulhai; funding acquisition, Amer Shalaby and Baher Abdulhai. All authors reviewed and approved the final version of the manuscript.

References

1. Blynn KKM (2018) Accelerating bus electrification: Enabling a sustainable transition to low carbon transportation systems (Doctoral dissertation, Massachusetts Institute of Technology). Available from: https://dspace.mit.edu/entities/publication/a079231a-22db-4fd9-92f1-3540dfd2517b.
2. Logan KG, Hastings A, Nelson JD (2022) Transportation in a net zero world: Transitioning towards low carbon public transport. *Springer Nature*. https://doi.org/10.1007/978-3-030-96674-4
3. Choma EF, Robinson LA, Nadeau KC (2024) Adopting electric school buses in the United States: Health and climate benefits. *Proc Natl Acad Sci* 121: e2320338121. https://doi.org/10.1073/pnas.2320338121

4. Saini H (2024) Environmental benefits of transition to electric school buses. *Int J Multidiscip Res Growth Eval* 5: 7–12. https://doi.org/10.54660/.IJMRGE.2024.5.6.07-12
5. Bursey V, Kalisa E (2025) Health, economic, and environmental impacts of electric school bus adoption: A scoping review. *Sci Total Environ* 988: 179843. https://doi.org/10.1016/j.scitotenv.2025.179843
6. Avenali A, Catalano G, Giagnorio M, et al. (2024) Factors influencing the adoption of zero-emission buses: A review-based framework. *Renewable Sustainable Energy Rev* 197: 114388. https://doi.org/10.1016/j.rser.2024.114388
7. Deliali A, Chhan D, Oliver J, et al. (2021) Transitioning to zero-emission bus fleets: State of practice of implementations in the United States. *Transp Rev* 41: 164–191. https://doi.org/10.1080/01441647.2020.1800132
8. Othman K, Da Silva D, Shalaby A, et al. (2025) Interpretable machine learning models for predicting Ebus battery consumption rates in cold climates with and without diesel auxiliary heating. *Green Energy and Intell Transp* 4: 100250. https://doi.org/10.1016/j.geits.2024.100250
9. Alanazi F (2023) Electric vehicles: Benefits, challenges, and potential solutions for widespread adaptation. *Appl Sci* 13: 6016. https://doi.org/10.3390/app13106016
10. Sclar R, Gorguinpour C, Castellanos S, et al. (2019) Barriers to adopting electric buses. Available from: https://www.sustainable-bus.com/wp-content/uploads/2019/05/barriers-to-adopting-electric-buses.pdf.
11. Zhou B, Wu Y, Zhou B, et al. (2016) Real-world performance of battery electric buses and their life-cycle benefits with respect to energy consumption and carbon dioxide emissions. *Energy* 96: 603–613. https://doi.org/10.1016/j.energy.2015.12.041
12. Ekblom V (2024) Comprehensive sustainability in urban transport: The interconnected impacts of electrifying Stockholm's bus fleet. Available from: https://www.diva-portal.org/smash/get/diva2:1910276/FULLTEXT01.pdf.
13. Othman K, Hamed S, Da Silva D, et al. (2024). Decision support tools for effective bus fleet electrification: Replacement factors and fleet size prediction. *Transp Res Interdiscip Perspect* 28: 101267. https://doi.org/10.1016/j.trip.2024.101267
14. Apribowo CH, Sarjiya S, Hadi SP, et al. (2022) Optimal planning of battery energy storage systems by considering battery degradation due to ambient temperature: A review, challenges, and new perspective. *Batteries* 8: 290. https://doi.org/10.3390/batteries8120290
15. Rahman T, Alharbi T (2024) Exploring lithium-ion battery degradation: A concise review of critical factors, impacts, data-driven degradation estimation techniques, and sustainable directions for energy storage systems. *Batteries* 10: 220. https://doi.org/10.3390/batteries10070220
16. Lipu MS, Mamun AA, Ansari S, et al. (2022) Battery management, key technologies, methods, issues, and future trends of electric vehicles: A pathway toward achieving sustainable development goals. *Batteries* 8: 119. https://doi.org/10.3390/batteries8090119
17. Rogge M, Van der Hurk E, Larsen A, et al. (2018) Electric bus fleet size and mix problem with optimization of charging infrastructure. *Appl Energy* 211: 282–295. https://doi.org/10.1016/j.apenergy.2017.11.051
18. Gairola P, Nezamuddin N (2023) Design of battery electric bus system considering waiting time limitations. *Transp Res Rec: J Transp Res Board* 2677: 1415–1430. https://doi.org/10.1177/03611981221113321

19. Li JQ (2016) Battery-electric transit bus developments and operations: A review. *Int J Sustainable Transp* 10: 157–169. https://doi.org/10.1080/15568318.2013.872737
20. Pencheva V, Asenov A, Georgiev A, et al. (2026) An integrated model for the electrification of urban bus fleets in public transport systems. *Eng Proc* 121: 28. https://doi.org/10.3390/engproc2025121028
21. Topić J, Soldo J, Maletić F, et al. (2020) Virtual simulation of electric bus fleets for city bus transport electrification planning. *Energies* 13: 3410. https://doi.org/10.3390/en13133410
22. Booysen MJ, Abraham CJ, Pretorius BG (2025) A system-of-systems framework for planning the electrification of paratransit. *Energy Strategy Rev* 62: 101892. https://doi.org/10.1016/j.esr.2025.101892
23. Heide L, Guo S, Göhlich D (2025) From simulation to implementation: A systems model for electric bus fleet deployment in metropolitan areas. *World Electric Vehicle Journal* 16: 378. https://doi.org/10.3390/wevj16070378
24. Sennefelder RM, Martín-Clemente R, González-Carvajal R (2023) Energy consumption prediction of electric city buses using multiple linear regression. *Energies* 16: 4365. https://doi.org/10.3390/en16114365
25. Li P, Zhang Y, Zhang K, et al. (2021). The effects of dynamic traffic conditions, route characteristics and environmental conditions on trip-based electricity consumption prediction of electric bus. *Energy* 218: 119437. https://doi.org/10.1016/j.energy.2020.119437
26. Dong C, Xiong Z, Li N, et al. (2025) A real-time prediction framework for energy consumption of electric buses using integrated machine learning algorithms. *Transp Res Part E: Logistics Trans Rev* 194: 103884. https://doi.org/10.1016/j.tre.2024.103884
27. Morlock F, Rolle B, Bauer M, et al. (2019) Forecasts of electric vehicle energy consumption based on characteristic speed profiles and real-time traffic data. *IEEE Trans Veh Technol* 69: 1404–1418. https://doi.org/10.1109/TVT.2019.2957536
28. Ekici YE, Akdağ O, Aydin AA, et al. (2023) A novel energy consumption prediction model of electric buses using real-time big data from route, environment, and vehicle parameters. *IEEE Access* 11: 104305–104322. https://doi.org/10.1109/ACCESS.2023.3316362
29. Nan S, Tu R, Li T, et al. (2022) From driving behavior to energy consumption: A novel method to predict the energy consumption of electric bus. *Energy* 261: 125188. https://doi.org/10.1016/j.energy.2022.125188
30. Beckers CJ, Besselink IJ, Frints JJ, et al. (2019) Energy consumption prediction for electric city buses. *In 13th ITS European Congress: Fulfilling its promises*. Available from: https://pure.tue.nl/ws/files/127995147/postprint_2019_04_11_Energy_Consumption_Prediction_for_Electric_City_Buses_CJJBeckers_IJMBesselink_JJMFrints_and_HNijmijer.pdf.
31. Abdelwahed A, van den Berg PL, Brandt T, et al. (2020) Evaluating and optimizing opportunity fast-charging schedules in transit battery electric bus networks. *Trans Sci* 54: 1439–1731. https://doi.org/10.1287/trsc.2020.0982
32. Lajunen A (2018) Lifecycle costs and charging requirements of electric buses with different charging methods. *J Cleaner Prod* 172: 56–67. https://doi.org/10.1016/j.jclepro.2017.10.066
33. Hasan MM, Avramis N, Ranta M, et al. (2021) Multi-objective energy management and charging strategy for electric bus fleets in cities using various ECO strategies. *Sustainability* 13: 7865. https://doi.org/10.3390/su13147865

34. Hasan MM, Ranta M, El Baghdadi M, et al. (2020, November 18). Charging management strategy using ECO-charging for electric bus fleets in cities. *In 2020 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 1–8. https://doi.org/10.1109/VPPC49601.2020.9330970
35. Zeng B, Wu W, Ma C (2023) Electric bus scheduling and charging infrastructure planning considering bus replacement strategies at charging stations. *IEEE Access* 11: 125328–125345. https://doi.org/10.1109/ACCESS.2023.3330369
36. Houbbadi A, Trigui R, Pelissier S, et al. (2019). Optimal scheduling to manage an electric bus fleet overnight charging. *Energies* 12: 2727. https://doi.org/10.3390/en12142727
37. Wang Y, Liao F, Lu C (2022). Integrated optimization of charger deployment and fleet scheduling for battery electric buses. *Transp Res Part D: Transp Environ* 109: 103382. https://doi.org/10.1016/j.trd.2022.103382
38. Zhou GJ, Xie DF, Zhao XM, et al. (2020). Collaborative optimization of vehicle and charging scheduling for a bus fleet mixed with electric and traditional buses. *IEEE Access* 8: 8056–8072. https://doi.org/10.1109/ACCESS.2020.2964391
39. Wang J, Kang L, Liu Y (2020). Optimal scheduling for electric bus fleets based on dynamic programming approach by considering battery capacity fade. *Renewable Sustainable Energy Rev* 130: 109978. https://doi.org/10.1016/j.rser.2020.109978
40. Çölbay K, Engin O (2025). Simulation-based analysis of bus scheduling problem for mixed fleets: Effects of flexibility strategies on delay performance. *In 2025 9th International Artificial Intelligence and Data Processing Symposium (IDAP)*, 1–6. https://doi.org/10.1109/IDAP68205.2025.11222375
41. Kharouf M (2022) A novel MILP model for the decarbonization of transit bus system considering emissions reduction targets and distributed energy resources. *Doctoral dissertation, University of British Columbia*. Available from: https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0415784.
42. Federico S, Miraftabzadeh SM, Longo M, et al. (2025) Bus fleet electrification: A comprehensive review of charging systems, operations, planning and sustainability. *SSRN, Preprint*. https://doi.org/10.2139/ssrn.5929724
43. Akaber P, Hughes T, Sobolev S (2021) MILP-based customer-oriented e-fleet charging scheduling platform. *IET Smart Grid* 4: 215–223. https://doi.org/10.1049/stg2.12034
44. Chiu CC, Huang H, Chen CF (2024) A simulation-based optimization approach for the recharging scheduling problem of electric buses. *Transp Res Part E: Logistics Transp Rev* 192: 103835. https://doi.org/10.1016/j.tre.2024.103835
45. Chen Y, Zhang Y, Sun R (2021) Data-driven estimation of energy consumption for electric bus under real-world driving conditions. *Transp Res Part D: Transp Environ,* 98: 102969. https://doi.org/10.1016/j.trd.2021.102969
46. Zhang Z, Wang S, Ye B, et al. (2025). A feature prediction-based method for energy consumption prediction of electric buses. *Energy* 314: 134345. https://doi.org/10.1016/j.energy.2024.134345
47. Li P, Zhang Y, Zhang Y, et al. (2021) Prediction of electric bus energy consumption with stochastic speed profile generation modelling and data-driven method based on real-world big data. *Appl Energy* 298: 117204. https://doi.org/10.1016/j.apenergy.2021.117204
48. Jiang J, Yu Y, Min H, et al. (2023) Trip-level energy consumption prediction model for electric bus combining Markov-based speed profile generation and Gaussian process regression. *Energy* 263: 125866. https://doi.org/10.1016/j.energy.2022.125866

49. Pamuła T, Pamuła D (2022) Prediction of electric buses energy consumption from trip parameters using deep learning. *Energies* 15: 1747. https://doi.org/10.3390/en15051747
50. Zhao L, Ke H, Huo W (2023) A frequency item mining-based energy consumption prediction method for electric bus. *Energy* 263: 125915. https://doi.org/10.1016/j.energy.2022.125915
51. Zhang Z, Ye B, Wang S, et al. (2024) Analysis and estimation of energy consumption of electric buses using real-world data. *Transp Res Part D: Transp Environ* 126: 104017. https://doi.org/10.1016/j.trd.2023.104017
52. Fang Y, Yang WH, Ihara Y, et al. (2025) Developing a simple electricity consumption prediction formula for the pre-introduction prediction for electric buses. *World Electr Veh J* 16: 67. https://doi.org/10.3390/wevj16020067
53. Lyu A, Zhang H, Zhang Y, et al. (2025) A study on energy consumption analysis and prediction of electric bus at intersections considering driving behavior. *Sci Rep* 15: 44755. https://doi.org/10.1038/s41598-025-28835-4
54. Dong C, Xiong Z, Zhang C, et al. (2025) A transformer-based approach for deep feature extraction and energy consumption prediction of electric buses based on driving distances. *Appl Energy* 380: 123941. https://doi.org/10.1016/j.apenergy.2024.123941
55. Ma Y, Ye B, Wang S, et al. (2025) Accurate prediction of energy consumption of electric buses based on traffic condition, vehicle status, driving behaviour and environmental condition. *IEEE Trans Transp Electrification* 4: 9778–9792. https://doi.org/10.1109/TTE.2025.3556033
56. Zhao J, He J, Wang J, et al. (2025) Energy consumption prediction for electric buses based on traction modeling and LightGBM. *World Electr Veh J* 16: 159. https://doi.org/10.3390/wevj16030159
57. Kang Y, Wei J, Liu Z, et al. (2025). An energy consumption prediction model for electric buses based on extreme gradient boosting fusion algorithm. *Int J Green Energy* 22: 2504–2517. https://doi.org/10.1080/15435075.2025.2464155
58. Wang Z, Xu G, Sun R, et al. (2025) Online energy consumption forecast for battery electric buses using a learning-free algebraic method. *Sci Rep* 15: 1931. https://doi.org/10.1038/s41598-024-82432-5
59. Vehviläinen M, Lavikka R, Rantala S, et al. (2022) Setting up and operating electric city buses in harsh winter conditions. *Appl Sci* 12: 2762. https://doi.org/10.3390/app12062762
60. Tian X, Wang B, Wang Z, et al. (2025). Unraveling energy demand in battery electric bus operations through an explainable machine learning approach using real-world cold-climate data. *Energy*, 139256. https://doi.org/10.1016/j.energy.2025.139256
61. Pettinen R, Anttila J, Muona T, et al. (2023) Testing method for electric bus auxiliary heater emissions. *Energies* 16: 3578. https://doi.org/10.3390/en16083578
62. Vuchic VR (2005) *Urban transit: Operations, planning, and economics*. John Wiley & Sons. Available from: https://www.wiley.com/en-us/shop/general-introductory-civil-engineering-construction/urban-transit-operations-planning-and-economics-p-9780471632658.