Paper ID APEN-MIT-2020_53
Applied Energy Symposium: MIT A+B
May 17-19, 2020 • Cambridge, USA

Exploring the uncertainty in trip-based eletricity consumption of EBs with a real-world big data from 100% electrification of bus network in Shenzhen, China

Pengshun Li
Tsinghua-Berkeley Shenzhen Institute
Tsinghua University
Shenzhen, PR China
lps19@mails.tsinghua.edu.cn

Yi Zhang
Tsinghua-Berkeley Shenzhen Institute
Tsinghua University
Shenzhen, PR China
zy1214@sz.tsinghua.edu.cn

Yi Zhang
Department of Automation
Tsinghua University
Beijing, PR China
zhyi@mail.tsinghua.edu.cn

Abstract—Currently, prediction of trip-based electricity consumption of electric buses (EBs) has become an important prerequisite for the deployment of large-scale electric bus fleets and the location of the charging infrastructures. Previous state-of-the-art approaches to estimate the electricity consumption focus on making rough electricity consumption assumptions or building physics-based electricity consumption model. This paper constructed a neural network model to predict the trip-based electricity consumption of EBs and six influencing factors were taken as input variables. Further, sensitivity analyses were performed to investigate how these factors influence the consumption results. This model was implemented and validated on real-world electric bus data from a five-month consecutive collection in Shenzhen, China, comprising 1024 EBs. The experiment demonstrated the predictive effectiveness of this model and the results from sensitivity analyses show that trip length is the key factor to determine the consumption, but other factors average travel speed, the number of bus stops and traffic lights, direction, time parameters also have different level of impacts on the results.

Keywords—electric bus, electricity consumption, neural network, big data

I. INTRODUCTION

Currently, in order to encourage the transportation electrification, increasing cities are replacing diesel-powered public buses with electric buses, this is because electric buses (EBs) are classified as zero emission vehicles and being more energy efficient in urban area [1]. A number of advantages of EBs have been listed, including quiet, comfortable due to the lack of motor vibrations, braking energy recovery, being more energy efficient than diesel-powered buses, and being locally emission-free due to the fossil fuel savings and the reduction of green house gases (GHG) emissions [2-4]. However, in order to enlarge the application of electric bus, estimation of the trip-based electricity consumption has become crucial for planning and deployment of large-scale electric bus fleets, calculation of operating costs, and selection of the right battery capacity [5].

In literature, there has been a long line of studies in the calculation of electricity consumption of EBs. Existing studies mostly use two approaches to determine the electricity consumption of EBs for terminus-to-terminus trips. The first is assuming the linear relationship between the electricity consumption of EBs and their trip length. Wang et al. assumed that the electricity consumption of EBs is proportional to the trip length in order to simplify the problem about scheduling the electric bus to be recharged [6]. Xylia et al. considered electricity consumption rate of EBs as a fixed value, which is 1.5 kWh/km in the case study of electric bus network in Stockholm, Sweden in order to simplify the optimization of the distribution of charging infrastructure [7]. Similarly, Paul and Yamada also used a fixed average electricity consumption rate of 1.41 kWh/km to calculate the EB's energy demand [8]. Although this approach is very simple to estimate the electricity consumption, it does not take into account the impact of other influencing factors (e.g. route characteristics) on electricity consumption. Hence, there would be a less accurate electricity consumption estimation through using this approach.

The second is building physics-based electricity consumption models. Marc et al. proposed a longitudinal dynamics model where low-resolution data including the arrival and departure time of the buses at each bus stop were collected [5]. Wu et al. presented an analytical EV power estimation model on the basis of the analysis of the relationship among the EV's power, the velocity, the acceleration, etc. And then the instantaneous power could be obtained and trip electricity consumption could be calculated [9]. Cedric et al. constructed three electricity consumption calculation models through using aggregated values of the kinematic parameters of trips including driving distance, travel time, temperature, acceleration data, and so on, based on the vehicle dynamics physical model and multiple linear regression [10]. Jari et al. developed an equation-based model to predict the real-world electricity consumption of EBs and study the nature and impact of various influencing factors [11]. The latter methods have more accurate prediction results of the electricity consumption of EBs than the former ones, as these models consider the inherent physical characteristics of the electricity consumption of EBs rather than assuming the electricity consumption rate as a fixed value.

However, none of them have considered inherent variability in the trip-based electricity consumption of EB in the real-world operation, which will over or under estimate the electricity consumption. In this study, we develop a trip-based electricity consumption prediction model of EBs based on big data through neural networks (NN). The data-driven method can identify the variability with sufficient data. Moreover, the method can consider the related influencing parameters, including temporal characteristic (e.g. peak hour and off-peak hour), dynamic traffic conditions (e.g. the average travel speed) and route characteristics (e.g. the number of traffic lights and bus stops) and also understand their impacts on the prediction results. Further, as a data-driven method, NN does not need to specify a clear physical relationship between the data in advance [12] and it has the advantage of allowing the approximation of arbitrary non-linear functions of complexity [13], thus, NN can have a good predictive performance in trip-based electricity consumption prediction.

To the best knowledge of the authors, there has been no such model in the existing researches. This study is one of the first attempts to construct a neural network model to predict the trip-based electricity consumption of EBs. The model investigates the uncertainty and evaluate the impacts of influencing parameters under practical operation conditions. The main contributions of this study are as follows: (1) A neural network electricity consumption prediction model has been developed and influencing factors are taken into account, including temporal characteristic (e.g. peak hour and off-peak hour, weekdays and weekends), dynamic traffic conditions (e.g. the average travel speed) and route characteristics (e.g. the trip length, the number of traffic lights and bus stops, the driving direction); (2) The proposed model is validated by real-world electric bus data from a five-month consecutive collection in Shenzhen, China, comprising 1024 EBs. (3) A sensitivity analysis is performed to understand the impacts of influencing factors on the prediction results.

II. Data Collection

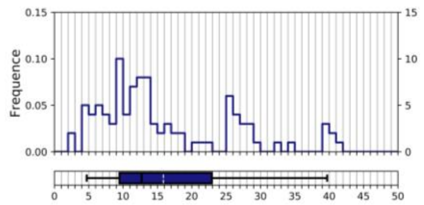
A. Data Source

To protect the environment and reduce public bus exhaust emissions, the government of Shenzhen, China shows a great interest in transportation electrification and all fuel-powered buses has been replaced with EBs in 2017. Currently, Shenzhen's public bus network is very large, with more than 360 bus routes and a fleet of more than 5600 electricity buses. Our proposed neural network electricity consumption prediction model is based on real-world data from this full-scale network.

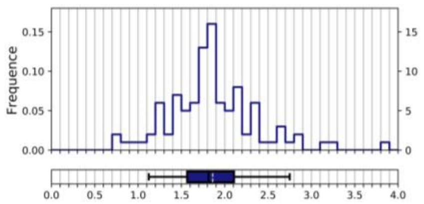
The data in this study originate from the electric bus data from 1024 electric buses consecutively collected from 1 January to 31 May 2019 in Shenzhen, China. The data, including the bus ID, the GPS time, the mileage, the state-of-charge (SoC), and the GPS coordinates, are recorded every 10 seconds.

B. Data Preprocessing

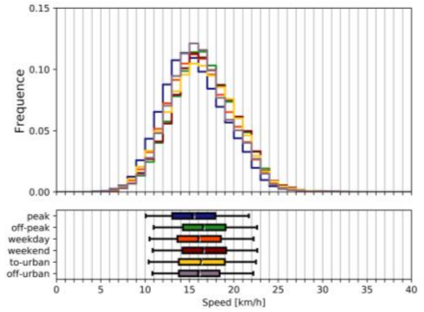
Our study is to research the EB's electricity consumption for a terminus-to-terminus trip. Hence, the raw data need preprocessing to be cut into the data each containing only one trip. The temporal attribute, the trip length, the average travel speed, and the electricity consumption for a terminus-to-terminus trip can be calculated on the basis of the processed data. To find the bus line corresponding to each electric bus, the data of the GPS coordinates are used to do map matching to match the bus lines. Then the characteristics of each line can be obtained, including the number of bus stop, the number of traffic lights needing passing. Further, the driving direction of each trip (to urban area or off urban area) is also obtained on the basis of the data of the GPS coordinates. Besides, the observed electricity consumption is calculated based on the consumed battery SoC (only accurate to two decimal places (e.g. 6%)) and 292 kWh (the battery stored energy corresponding to 100 % SoC).



(a) Trip Length



(b) Number of bus stops per kilometer



(c) Average speed for peak and off-peak hours, weekday and weekend, to and off urban area

Fig. 1. Distribution of some characteristics of the public bus routes per trip. (Box plot whiskers set at 5th and 95th percentiles.)

C. Characteristics of the bus network

Fig. 1 shows the distribution of some key characteristics of the public bus routes in Shenzhen, China. From Fig. 1a, the median bus trip length is 12.7 km, and the mean value is 15.9 km. The distribution is quite wide with the standard deviation (STD) of 9.2 km.

Generally speaking, the increasing number of bus stops would have a negative effect on average travel speed and electricity consumption for each terminus-to-terminus trip as frequent acceleration and deceleration can lead to increased idle time and energy loss. From Fig. 1b, the median number of bus stops reaches up to 1.8 stops per kilometer (in other words, the median distance of two adjacent visited bus stops is 556 m). The STD is 0.51 km.

Fig. 1c presents the differences of average speed for a terminus-to-terminus trip among different times of a day, different driving directions, and different days of a week. The average speed for a terminus-to-terminus trip depends on the number of bus stops and traffic lights and the traffic conditions at different times (e.g. peak and off-peak hours, weekdays and weekends) and traffic conditions for different driving directions (e.g. to urban area and off urban area). As shown in Fig. 1c, the median average speed for peak hour is remarkably different from that for off-peak hour (15.6 km/h and 16.6 km/h respectively). And the figure for weekends (16.2 km/h) is also relatively lower than the figure for weekdays (16.7 km/h). The median average speed difference in the driving direction is not very obvious, but it can also be seen that the speed of driving off urban area is lower than the speed of driving to urban area.

The graphs in the Fig. 1 clearly shows the heterogeneity of characteristics of bus routes in the real world, which can lead to significant changes in electricity consumption. Further, this indicates that these characteristics need to be taken into account when predicting trip-based electricity consumption of EBs.

D. Data Analysis

Based on the analysis of some characteristics of bus network abovementioned, this subsection presents the relationships between energy consumption and the potential influence factors, such as trip length, temporal attribute (time of a day and day of a week), and driving direction.

1) Trip Length. Fig. 2 presents the relationship between the 5th, the median, the 95th percentiles of trip-based electricity consumption observed for each bus line and trip length. The electricity consumption appears to be roughly proportional to the trip length. However, with the trip length over 30 km, the energy consumption rate has a different performance from that of the trip length below 30 km. Hence, when predicting trip-based electricity consumption of EBs, a consideration of trip length alone is not sufficient to have an accurate electricity consumption prediction.

2) Temporal Attributes. Fig. 3 compares the electricity consumption rate of EBs at different departure times of a day. During peak hours (7–8 a.m. for the morning peak hours and 4–6 p.m. for the evening peak hours), the electricity consumption rate (around 1.19 kWh/km) is relatively higher than other time periods, this is because during peak hours, there are more passengers on the buses (increasing the mass of the vehicle) and heavier road congestion (reducing travelling speed and increasing the share of idle time). Conversely, during other time periods, both the number of passengers travelling on buses and traffic on the road are reduced, thus the electricity consumption rate is lower, with the average median value of 1.1 kWh/km. Fig. 4 compares the electricity consumption rate at different days of a week. The electricity consumption rates on weekdays show a similar pattern, so that at weekends are. However, there is a significant difference in the electricity consumption between on weekdays and at weekends. The electricity consumption

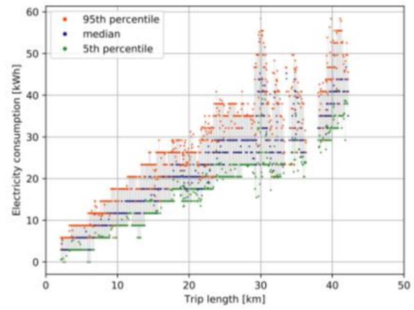


Fig. 2. Trip-based electricity consumption aggregated by bus lines. Each bus line is represented by three types of points (the green, blue, and red points represent the 5th, the median, and the 95th percentile of electricity consumption observed for this line)

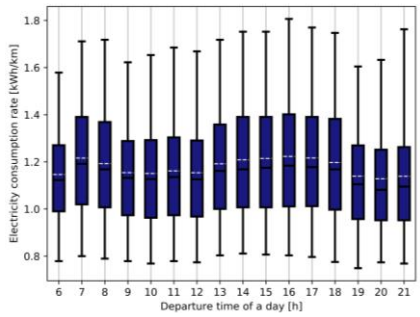


Fig. 3. Trip-based electricity consumption rate at different departure times of a day. (Box plot whiskers set at 5th and 95th percentiles.)

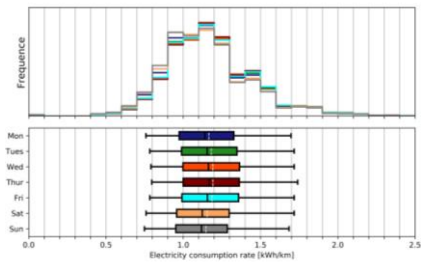


Fig. 4. Trip-based electricity consumption rate at different days of a week. (Box plot whiskers set at 5th and 95th percentiles.)

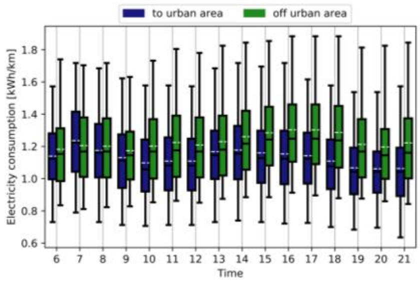


Fig. 5. Trip-based electricity consumption rate for different driving directions, with hourly statistics. (Box plot whiskers set at 5th and 95th percentiles.)

rates on weekdays are relatively higher due to the fact that higher number of people commuting to and off work on buses. In comparison, at weekends, there are not many travel demand and traffic pressure could be eased, so the energy consumption could be lower. Overall, Figs. 3–4 confirm that when predicting trip-based electricity consumption of EBs, temporal attributes should be taken into account to increase the prediction accuracy.

3) Driving Direction. Fig. 5 compares the electricity consumption rates for different driving directions on the basis of each departure time. During 6–9 a.m. the median electricity consumption rates for driving to and off urban area are similar (around 1.2 kWh/km). The electricity consumption rate around 7 a.m. for driving to urban area is slightly higher than that for driving off urban, due to more commuters need to take buses from suburban area to urban area for work. However, during 10 a.m.–9 p.m., the median electricity consumption rates with the driving direction of off-urban area are higher than that with the driving direction of to-urban area, and this difference reaches its maximum during the evening peak hours, with the figure of 0.16 kWh/km. Hence, the driving direction is also an important influencing factor affecting the trip-based electricity consumption of EBs.

III. METHODOLOGY

A. Neural Network Prediction Model

A BP neural network is a type of neural network, which is a multi-layer feedforward neural network based on the error back propagation [14]. The basic concept is to use the network mean square error (MSE) as the objective function. Based on the gradient descent strategy, the parameters are adjusted in the negative gradient direction of the target to minimize the MSE between the expected output value and the observed value.

In this study, a 3-layer neural network is built. The data are randomly divided into a training data set and a testing data set according to a 7:3 ratio. Cross-validation is performed for the selection of parameters. After minimizing the mean absolute percentage error (MAPE) in the process of optimization, the number of hidden layers can be determined, which is set to 10 in this study.

B. Influencing Factors

Based on the data analysis in the Section 2.4, influencing factors, including the departure time of a day of a trip, the departure day of a week of a trip, the driving direction of a trip, the trip length of a trip, the average travel speed, the number of bus stops and traffic lights are taken as input variables, and the trip-based electricity consumption is taken as output variable. Some examples of the input variables and output variables are shown in Table 1. Rows A–F are input variables. Row A records the departure time of each trip; Row B records the departure day of a week of each trip (e.g. Number 1 represents Monday); Row C records the driving direction of each trip (Number 1 represents driving to urban area, and number 2 represents driving off urban area); Rows E–F are the travel distance (in km), the average travel speed (in km/h), the number of bus stops and the number of traffic lights; Row H is the output variable (trip-based electricity consumption of EBs (in kWh).

TABLE I. DATA SAMPLES OF INPUT VARIABLES AND OUTPUT VARIABLE

	Variables	Sample1	Sample2
A	Departure Time of a day (hour)	7	15
B	Departure day of a week	1	6
C	Driving Direction	1	2
D	Trip Length (km)	25.8	17.8
E	Average Travel Speed (km/h)	15.35	16.00
F	Number of Bus Stops & Traffic Lights	91	62
H	Electricity consumption of EBs (kWh)	35.04	20.44

C. Prediction Accuracy Measurement

To better analyse the actual prediction performance of this neural network electricity consumption prediction model, the prediction results need to be evaluated and analysed. In this study, the following performance accuracy indicators are adopted: MAE, MAPE, and RMSE, which are given by:

$$MAE = \frac{1}{N}\sum_i \left|\widehat{D}_i - D_i\right| \quad (1)$$

$$MAPE = \frac{1}{N}\sum_i \frac{|\widehat{D}_i - D_i|}{D_i} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_i [\widehat{D}_i - D_i]^2}{N}} \quad (3)$$

where N is the total sample size, $\widehat{D}_i$ is the predicted value of the ith sample, and D_i is the observed value of the ith sample.

IV. RESULTS AND DISCUSSION

A. Prediction Analysis

The results in this section are based on the processed data set, containing more than 557000 data. Applying the neural network electricity consumption prediction model to this data set, the predicted results are obtained, shown in Fig. 6.

And the relative error on the prediction for this trip-based electricity consumption and its distribution can be seen in Fig. 7. The Fig. 7 shows that average relative error decreases for trips with a higher quantity of energy consumed, and the median and mean relative error is 0.98% and 4.53% respectively. Excluding outliers, the 5th and 95th percentile of relative error are -27.48% and 49.09% respectively. Further, according to prediction accuracy indicators, the values of the MAE, MAPE and RMSE are 2.36, 17.22%, 3.12 respectively, which could indicate that this neural network electricity consumption prediction model performs well.

B. Sensitivity Analysis

To have a deep understanding to identify which parameters have the greatest influence on the prediction results, a sensitivity analysis is performed in our model. Some input variables (departure time of a day, departure day of a

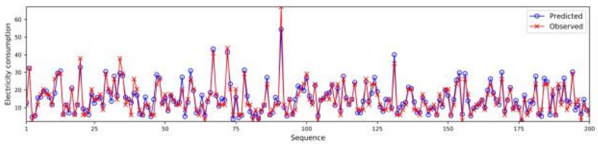


Fig. 6. Predicted and observed results (partial)

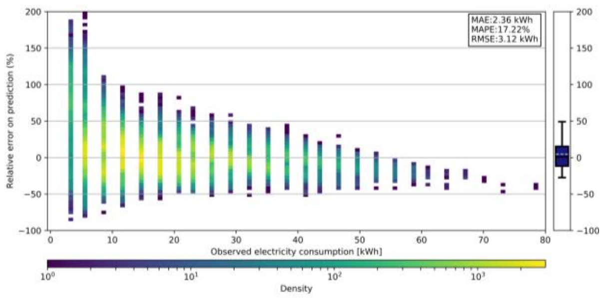


Fig. 7. Relative error on the prediction (%) as a function of the observed electricity consumption (kWh) for the BP neural network. (Box plot whiskers set at 5th and 95th percentiles.)

TABLE II. VARIABLES FOR SENSITIVITY ANALYSES

Variables	Base(a)	Base(b)	Base(c)	Base(d)
Departure Time	peak weekday	off peak weekday	peak weekday	peak weekend
Driving Direction	off urban	off urban	to urban	off urban
Trip Length (km)	12.7	12.7	12.7	12.7
Average Speed	16.12	16.53	14.82	16.42
Number of Bus Stops & Traffic Lights	64	64	64	64

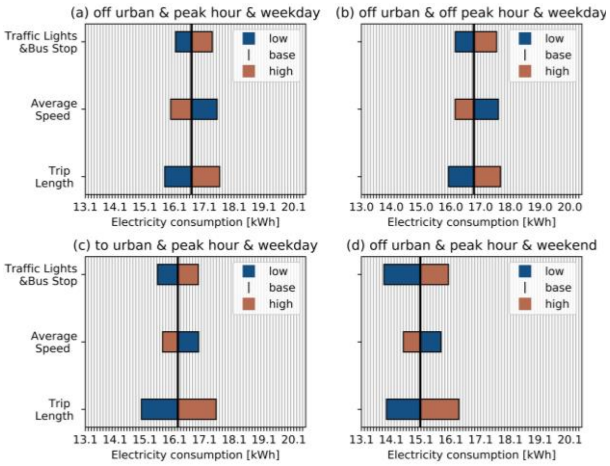


Fig. 8. Sensitivity analyses for electricity consumption under four different conditions. (Parameter variation: ±10% from the "base" value for each parameter)

week, driving direction) could not be quantified, so sensitivity analyses are performed on four different conditions. For the sake of brevity, every other single input variable is fixed to its median value, corresponding to the "base" cases (Table 2).

The "high" and "low" cases correspond to a variation of "+10%" and "−10%" from the "base" value for each variable respectively (Fig. 8).

From the results, it can be observed that the variables with the highest impact on the electricity consumption is the trip length under four different conditions. Under the conditions (c)-(d), its influence is more obvious than other two variables. The impact of average speed is also worth attention. With the average speed increasing, the electricity consumption decreases (only confined to the average on–trip speed of no more than 25 km/h, as the average speeds of terminus-to-terminus trips in our experiment are no more than 25 km/h). Further, the influence of the number of bus stops and traffic lights also cannot be neglected, the electricity consumption also increases with more bus stops and traffic lights.

V. CONCLUSION AND FUTURE WORK

To help facilitate the planning and deployment of large electric bus fleets, the calculation of operating costs, and the selection of the right battery capacity, this paper proposes the neural network electricity consumption prediction model. This model can investigate the uncertainty in trip-based electricity consumption of EBs with consideration of influencing parameters under practical operation conditions. The existing real-world EB data in Shenzhen are taken as a case study in this study. The results indicate that it is effective to use our model to predict trip-based electricity consumption according to the performance accuracy indicators, MAE, MAPE, and RMSE. Moreover, sensitivity analyses have been conducted to figure out which variables have more influence on consumption under four different conditions, and the results show that the trip length has the highest impact, with the average travel speed and the number of bus stops and traffic lights following behind. Overall, we can support that in the case of trip-based electricity consumption prediction, using neural network prediction model can be effective and convenient. Moreover, it can be suggested that the proposed model in this study serve as a basis of scheduling optimization of large-scale electric bus fleets and optimization of the location of charging infrastructures.

ACKNOWLEDGMENT

This research was supported in part by National Natural Science Foundation of China (61673233), Basic Research Program of Shenzhen Science and Technology Innovation Committee (JCYJ20180307123910003), Shenzhen Science and Technology Program (Grant No. KQTD20170810150821146).

REFERENCES

[1] W. Vaz, A. K. Nandi, and U. O. Koylu, "Finding an Optimal Driving Strategy for an Electric Bus Based on Operational Data," in *ASME 2015 9th International Conference on Energy Sustainability collocated with the ASME 2015 Power Conference, the ASME 2015 13th International Conference on Fuel Cell Science, Engineering and Technology, and the ASME 2015 Nuclear Forum*, 2015: American Society of Mechanical Engineers Digital Collection.

[2] R. Kühne, "Electric buses–An energy efficient urban transportation means," *Energy*, vol. 35, no. 12, pp. 4510-4513, 2010.

[3] A. Lajunen and T. Lipman, "Lifecycle cost assessment and carbon dioxide emissions of diesel, natural gas, hybrid electric, fuel cell hybrid and electric transit buses," *Energy*, vol. 106, pp. 329-342, 2016.

[4] B. Zhou, Y. Wu, B. Zhou, R. Wang, W. Ke, S. Zhang, "Real-world performance of battery electric buses and their life-cycle benefits with respect to energy consumption and carbon dioxide emissions," *Energy*, vol. 96, pp. 603-613, 2016.

[5] M. Gallet, T. Massier, and T. Hamacher, "Estimation of the energy demand of electric buses based on real-world data for large-scale public transport networks," *Applied energy*, vol. 230, pp. 344-356, 2018.

[6] Y. Wang, Y. Huang, J. Xu, and N. Barclay, "Optimal recharging scheduling for urban electric buses: A case study in Davis," *Transportation Research Part E: Logistics and Transportation Review*, vol. 100, pp. 115-132, 2017.

[7] M. Xylia, S. Leduc, P. Patrizio, F. Kraxner, and S. Silveira, "Locating charging infrastructure for electric buses in Stockholm," *Transportation Research Part C: Emerging Technologies*, vol. 78, pp. 183-200, 2017.

[8] T. Paul and H. Yamada, "Operation and charging scheduling of electric buses in a city bus route network," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2014: IEEE, pp. 2780-2786.

[9] X. Wu, D. Freese, A. Cabrera, and W. A. Kitch, "Electric vehicles' energy consumption measurement and estimation," *Transportation Research Part D: Transport and Environment*, vol. 34, pp. 52-67, 2015.

[10] C. De Cauwer, J. Van Mierlo, and T. Coosemans, "Energy consumption prediction for electric vehicles based on real-world data," *Energies*, vol. 8, no. 8, pp. 8573-8593, 2015.

[11] J. Vepsäläinen, K. Kivekäs, K. Otto, A. Lajunen, and K. Tammi, "Development and validation of energy demand uncertainty model for electric city buses," *Transportation Research Part D: Transport and Environment*, vol. 63, pp. 347-361, 2018.

[12] F. Guo, J. W. Polak, and R. Krishnamoorthy, "Predictor fusion for short-term traffic forecasting," 2018.

[13] A. D. Alvarez, F. S. Garcia, J. E. Naranjo, J. J. Anaya, and F. Jimenez, "Modeling the driving behavior of electric vehicles using smartphones and neural networks," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 3, pp. 44-53, 2014.

[14] Y. Chauvin and D. E. Rumelhart, *Backpropagation: theory, architectures, and applications*. Psychology press, 201