

Research Article

Study on Driving Cycle Synthesis Method for City Buses considering Random Passenger Load

Xiaodong Liu ^(b), ¹ Jian Ma ^(b), ¹ Xuan Zhao ^(b), ¹ Juan Du, ² and Yanfeng Xiong¹

¹School of Automobile, Chang'an University, Xi'an, Shaanxi 710064, China
 ²School of Mechanical & Automotive Engineering, Liaocheng University, Liaocheng, Shandong 252059, China

Correspondence should be addressed to Jian Ma; majian@chd.edu.cn

Received 29 July 2019; Accepted 18 January 2020; Published 27 February 2020

Academic Editor: Monica Menendez

Copyright © 2020 Xiaodong Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel driving cycle construction method in consideration of velocity, road slope, and passenger load, based on a real-world bus route with a plug-in hybrid electric bus (PHEB). The main purpose is to address the disadvantage that an inaccurate reflection of the real-world driving characteristics for city buses will be caused when ignoring the passenger load in the course of a driving cycle synthesis. Two contributions are supplemented to distinguish from the previous research. Firstly, a novel station-based method is proposed aiming at developing a driving cycle with high accuracy. The kinematic segments are partitioned according to the distance of adjacent bus stops, while a two-dimensional Markov chain Monte Carlo method is employed to synthesize driving cycle between each interval of adjacent bus stops. Secondly, the random passenger load for different bus stops is treated as a discrete Markov chain model, according to the correlation analysis of the measured passenger data which are distinguished for off-peak and peak hours. Meanwhile, Monte Carlo simulation and maximum likelihood estimation are utilized to determine the most likely number of passengers for each bus stop. At last, the fuel consumption of the PHEB is simulated with the best-synthesized driving cycle and contrasted to the mean fuel consumption of the later measured data which is composed of the velocity, road slope, and the passenger load. The results demonstrate that the synthesized driving cycle has a higher accuracy on fuel consumption estimation.

1. Introduction

The depletion of fossil fuels together with severe environmental pollution has greatly promoted the development of energy-saving technology and advanced hybrid propulsion systems [1, 2]. Plug-in hybrid electric buses (PHEBs) have been considered as one of the most promising solutions to decrease fuel consumption and exhaust emissions for urban public transport, due to its longer pure electric mileage and grid charge capability [3, 4]. Driving cycle is generally described as a velocity-time profile to represent driving characteristics of the real world. It has been widely utilized in the detection of exhaust pollution and energy consumption with the purpose of certification [5]. Moreover, it is also a vital foundation of vehicle design [6], especially for a PHEB with a complicated optimization problem. The component size, the powertrain topology, and the control strategies can be suitably optimized to decrease the fuel consumption and production cost, based on a typical driving cycle [7].

Numerousness standardized driving cycles are existing, which can be classified as transient cycles and modal cycles, such as the well-known Federal Test Procedure-75 (FTP-75) and the New European Driving Cycle (NEDC). The former has frequent velocity variation similar to the real-world driving behavior, whereas the latter is mainly composed of a succession of constant velocity or acceleration [5–7]. Both of them are used as legislative driving cycles for evaluation and certification of the automobile performance (e.g., emissions and energy consumption) by regulatory authorities. However, characteristics of the driving cycle are considerably variational from one region to another, and the legislative driving cycles are not suitable for the vehicle design owing to

their discrepancy with the real-world driving characteristics in different regions [8]. Therefore, nonlegislative driving cycles have been investigated in different countries and cities, such as Edinburgh driving cycle [9], Athens driving cycle [10], Hong Kong driving cycle [11], Pune driving cycle [12], Singapore driving cycle [13], and Mashhad driving cycle [14], as well as some customized driving cycles in typical Chinese cities [15]. All of them have been extensively employed in research from performance estimation to the vehicle design. Furthermore, some typical driving cycles have also been synthesized for city buses, considering the different driving characteristics between the city buses and passenger cars [16–21].

Currently, with the rapid development of electric vehicles (EVs), more and more investigation of driving cycles focused on the battery electric vehicles (BEVs), hybrid electric vehicles (HEVs), and PHEVs. Wang et al. [22] analyzed the energy-saving potential of the EVs in realistic driving conditions and compared to internal combustion engine vehicles (ICEVs). The research confirmed that the energy-saving effect and the energy consumption of the EVs were underestimated under the NEDC cycle suggested by the government. Zhao et al. [23] constructed a representative EV urban driving cycle for passenger cars in Xi'an city and compared the energy consumption between the synthesized driving cycle and the international driving cycles [24]. The comparison revealed that large error was generated when the international driving cycles were adopted to estimate the energy consumption and driving range of the EVs. Patil et al. [25] investigated the influence of the driving cycle on the component sizing design of the PHEVs and simulated the vehicle performance over realistic driving cycles collected in Southeast Michigan. The author found that most of the realworld driving cycles consumed more fuel than standard driving cycles with the same trip distance. Gonder et al. [26] acquired the driving data by Global Positioning System (GPS) from a set of 277 vehicles in Missouri and analyzed the fuel consumption and operating characteristics over the collected driving data. It was found that the simulated fuel consumption of the real-world driving data was lower in comparison with certification cycles, while the electric energy consumption was slightly higher. The above results suggest that the existing driving cycles cannot well capture the characteristics of the real-world driving process, thereby leading to inaccuracy estimation of power demands for EVs. Therefore, it is imperative to develop a dedicated driving cycle for EVs [27]. Gong et al. [28] collected real-world operation data of electric vehicles in Beijing and developed a driving cycle for BEVs on the basis of the data analysis. The developed driving cycle was verified to represent the realworld driving condition well by comparison of the energy consumption between the predicted and the real world. Ashtari et al. [29] logged real-world driving data from a fleet of 76 cars in Winnipeg city and synthesized a driving cycle for PHEVs. The results indicated that the characteristics of the proposed driving cycle were able to represent the aggressive driving of the real-world well compared to existing standard driving cycles. Similar studies were also carried out in Dublin by Brady and O'Mahony [5], in Florence by Berzi

et al. [30], and in Southeast Michigan by Lee et al. [31]. It had also found that the role of a typical driving cycle was increasingly significant, not only for vehicle certification and design but also for analyzing the impact of PHEV on the grid. As plug-in hybrid electric buses (PHEBs) generally operate in a fixed bus route, with frequent stop-and-go behavior, it is hard to employ the existing standard driving cycle to represent their operating characteristics. It is of great significance to construct a typical driving cycle which can accurately reflect the real-world driving characteristics of the bus route, during the PHEB design process.

There are various methodologies to develop a typical driving cycle, which can be substantially classified into four categories: microtrips, segment-based, pattern classification, and Markov chain method [32-35]. Microtrips are usually extracted from the measurement data sequence between two uninterrupted time points at which the vehicle is stopped [12]. In the process of driving cycle construction, candidate cycles are randomly synthesized by the microtrips and then assessed based on acceptable target parameters to determine the preferable driving cycle [11-18]. Comparing with the microtrips, the segment-based method is more flexible for the partition of measured driving data. The kinematic segments are able to be partitioned according to the road characteristics and traffic conditions, where the vehicle may not be stationary or idle. Therefore, additional constraints must be added to the velocity and acceleration for connecting the segments together (i.e., ensuring smoothness between connection segments) when developing a new cycle [21, 35, 36]. In order to reduce the number of microtrips or kinematic segments during driving cycle construction, the pattern classification method is applied to cluster them into heterogeneous classes based on their statistical properties [5, 37]. Abundant studies employ principal component analysis (PCA) together with the clustering method to synthesize the new driving cycles [19, 23, 24, 38].

Since the real-world driving cycles are considerable stochastic and have been confirmed to have the Markov property [39], Markov chain method has been extensively accepted as one of the most accurate approaches for driving cycle construction so far [7, 20, 28, 29, 31, 36]. In previous literature studies, the Markov chain method was usually used to calculate the probability of transition between different types of snippets classified by the maximum likelihood estimation (MLE) method or other clustering methods [20, 29, 32]. To improve the quality of the constructed driving cycle, other methodologies based on the Markov chain have been proposed and exhibited better performance. In Ref. [31], a two-dimensional Markov chain containing the information of velocity and acceleration was introduced and employed to synthesize the driving cycle. Moreover, the road altitude was also considered in Ref. [7], and the new driving cycle was synthesized by a combined three-dimensional Markov chain Monte Carlo (MCMC) method. The author found that a satisfactory driving cycle can be synthesized by employing a two-dimensional Markov chain method including velocity and road altitude, whereas the use of threedimensional Markov chain increased the computational burden. In addition, the random passenger load has been

confirmed to have a considerable impact on both the powertrain design and the control strategies of the vehicles [40, 41]. Nevertheless, there are few literature studies taking into consideration of it in the course of driving cycle construction. It is a significant characteristic of the driving cycle, and the unrealistic vehicle design will be brought out without regard to it, especially for a PHEB with a greatly variational passenger load.

With the motivation of introducing passenger load information in the synthesized driving cycle, we collect the passenger data from a real road bus route, respectively, in off-peak and peak hours and analyze the correlation of the passenger data between adjacent bus stops. Based on the analysis, we propose a method in a combination of Markov chain. Monte Carlo simulation as well as maximum likelihood estimation, to determine the most likely passenger numbers for each bus stop in off-peak and peak hours. In addition, despite the obvious advantage of the Markov chain-based method in driving cycle synthesis, the accuracy and the time efficiency are usually conflicting [7]. Reducing the interval and increasing the number of states can improve the accuracy of the synthesized driving cycle; however, it leads to a computational burden [42]. To overcome this drawback and facilitate the introduction of the passenger load information, we employ a new station-based method to construct a representative driving cycle considering the velocity and acceleration, for a bus route. The kinematic segments are partitioned according to the actual distance of adjacent bus stops, and the velocity profile between the adjacent bus stops is synthesized separately by a two-dimensional MCMC method. Finally, the passenger number information is integrated into the synthesized driving cycle and the effectiveness of the proposed method is verified by comparison of the fuel consumption between the bestsynthesized driving cycle and later measured data, while the different initial state-of-charge (SOC) is also considered.

The remainder of the paper is organized as follows. Section 2 introduces the proposed method for driving cycle construction for a bus route, with no consideration of the passenger load. In Section 3, a Markov chain-based method is proposed to dispose of the random passenger load for different bus stops, on the basis of realistic passenger data analysis, respectively, for off-peak and peak hours. In Section 4, the synthesized driving cycle including passenger load information is validated by comparing the fuel consumption with the mean fuel consumption of the later measured data, followed by the conclusions in Section 5.

2. Construction of the Driving Cycle Based on a Bus Route

2.1. Driving Data Collection. The driving data are collected from a typical bus route in Liaocheng, Shandong province, and the studied PHEB is also running on the route. The bus route is comprised both of the urban and suburban driving conditions with a single trip distance of approximately 20.5 km (see Figure 1).

Considering the passenger load may have a distinct difference for the operation of a round trip, the round-trip

3

route is designed for data acquisition. In other words, the PHEB is driven routinely from the starting bus stop to the terminal bus stop and then returns to the starting bus stop. The bus route consists of 31 bus stops for a single trip, and the trip distances of different bus stops to the starting bus stop are shown in Table 1.

The on-board measurement method is utilized to obtain driving data, and the driving data are collected from both the global position system (GPS) and on-board diagnostics (OBD) [23]. The velocity data obtained from the OBD are mainly employed to identify and amend the abnormal data acquired from the GPS. The collected driving data are mainly constituted by the signal of vehicle velocity, acceleration, traveled distance, and road altitude with the sampling time interval of 5 s. The data collection has continued for approximately half a month including a whole day operation of the weekdays and weekends. There are 285,652 valid data being collected, which traveled for approximately 1,435 km. All measured data are preprocessed to remove the abnormal data and then extended with an interpolation method to transform the sampling time into 1s for which it will facilitate the driving cycle construction.

In addition, the number of passengers getting on and off the bus for the selected bus route is separately collected corresponding to the bus stops for both of the off-peak hours and peak hours. A total of 26,280 valid data for 219 round trips are manually logged according to the actual passenger flow, for which it continues for approximately 2 months.

2.2. Verification of Sampling Saturation. Theoretically, a large number of sample data have a great benefit to the accuracy of the synthesized driving cycle. Since the driving data are acquired from a fixed route with the same test vehicle ignoring the influence of the drivers, the characteristic parameters of the sample data will have a negligible variation when the sample size reaches a certain level. Accordingly, it is of great necessity to determine the appropriate sampling number, aiming at constructing a more accurate driving cycle via less measured data.

In this study, four characteristic parameters are analyzed to determine the saturation of the collected driving data, which are composed of the mean velocity v_m , mean driving velocity (i.e., not including the idle time) v_{md} , maximum acceleration a_m , and minimum acceleration a_d . Moreover, the percentages of acceleration time, deceleration time, idle time, and cruise time are also considered, which are respectively expressed as T_a , T_d , T_b and T_c .

As shown in Figures 2(a)-2(c), the selected 8 characteristic parameters tend to be stabilized as the number of samples is increased. It implies that the selected parameters can be utilized for a further determination of the sampling number. Therefore, an evaluation parameter namely "stability margin" is defined to identify a suitable minimum number of samples for the construction of a driving cycle. The stability margin is expressed as **K**, which can be calculated by the following equation:



FIGURE 1: The selected bus route in Liaocheng, Shandong, China (Map data ©2019 Baidu).

TABLE 1: Bus stop information for the round trip.

| Bus stop | Bus stop name | Trip distance (m) |
|----------|-------------------|-------------------|
| 1 | Starting bus stop | 0 |
| 2 | Bus stop 2 | 528 |
| 3 | Bus stop 3 | 782 |
| | | |
| 30 | Bus stop 30 | 18500 |
| 31 | Terminal bus stop | 20500 |
| 32 | Bus stop 30 | 22500 |
| | | |
| 59 | Bus stop 3 | 40218 |
| 60 | Bus stop 2 | 40472 |
| 61 | Starting bus stop | 41000 |

$$\mathbf{K}(j) = \frac{1}{8} \sum_{i=1}^{8} \frac{M_i(j) - M_i(j-1)}{M_i(j)}, \quad i = 1, 2, \dots, 8; \ j = 1, 2, \dots, n,$$
(1)

where *i* is the index of the selected characteristic parameters, *j* is the sampling number, $\mathbf{K}(j)$ is the stability margin when the sampling number is *j*, and $M_i(j)$ represents the accumulated value of the *i*th characteristic parameter when the sampling number is *j*.

As shown in Figure 2(d), the stability margin **K** is converging to zero with the increase of the sampling number. When the confidence level α is defined as ±0.001, the stability margin **K** fluctuates within the acceptable level with a successive increment of the sampling number from 29 to 35. In other words, at least 29 groups of valid data should be provided by measurements. Here, 35 groups of valid driving data are employed for the construction of the driving cycle to ensure reliability and stability.

2.3. Synthesis of the Driving Cycle Containing Velocity and Road Slope

2.3.1. Partition of Kinematic Segments. In the course of a driving cycle synthesis, the measured driving data are usually divided into the kinematic segments with the microtrip method or other regular methods [21, 23]. However, it may not keep the characteristics of the synthesized driving cycle consistent with the actual driving characteristics of the

measured data, especially for a bus driving cycle. In this study, a novel partition method based on the distribution of the bus stop is proposed to promote the accuracy of the synthesized driving cycle as well as to easily consider the variation of passenger load for different bus stops.

As shown in Figure 3, vehicle velocity and the corresponding acceleration of each measured driving cycle is divided into 60 kinematic segments according to the bus stop distribution in Table 1. Meanwhile, each segment should also satisfy the following restrictions:

$$\begin{cases} v^{(r)} = \left\{ v_1^{(r)}, \dots, v_t^{(r)} \right\}, \\ v_1^{(r+1)} = v_t^{(r)} + a_t^{(r)}, \\ S_r = \int v^{(r)} \cdot dt, \end{cases}$$
(2)

where $v^{(r)}$ represents the velocity collection of the *r*th kinematic segment, $v_1^{(r)}$ and $v_t^{(r)}$ represent the start and end velocity of the *r*th kinematic segment, respectively, $r = 1, 2, \ldots, 60, a_t^{(r)}$ is the corresponding acceleration when the vehicle velocity is $v_t^{(r)}$ for the *r*th kinematic segment, S_r represents the actual distance between the *r*th bus stop and (r+1)th bus stop, and *t* is the driving time of the kinematic segment.

After the segmentation of each group driving data, the kinematic segments which belong to the same interval of the adjacent bus stops are classified into the same dataset. As shown in Figure 4, since the kinematic segment should be satisfied with the restriction of trip distance, the start and end velocity may not be zero. This may cause a velocity fluctuation when the driving cycle of each interval is connected. Thus, the problem is disposed of by adding zeros and simultaneously restricting the starting and ending acceleration of constructed driving segments when synthesizing the driving cycle.

2.3.2. Establishing a Markov Chain Model including Velocity and Acceleration. Many available methodologies have been adopted to construct a representative driving cycle in the previous research studies. Since the Markov chain is a stochastic process for which the future state depends only on the current state and independent of the past state, it has been widely utilized to dispose of the random property of



FIGURE 2: Saturation analysis of driving data sampling. Variation of (a) v_m and v_{md} , (b) a_{max} and a_{min} , (c) T_a , T_d , T_i , and T_c , and (d) **K** with the increase in samples.



FIGURE 3: The partition method for a measured driving cycle.



FIGURE 4: The dataset of a kinematic segment.

driving cycles. A Markov chain can be defined as a discretetime sequence of random variables $\{X_k\}$ that have the Markov property and can be expressed as follows [31]:

$$P\left\{ \left(X_{k+1} \in x_{j} \right) \middle| X_{0}, X_{1}, \dots, X_{k} \right\}$$

$$= P\left\{ \left(X_{k+1} \in x_{j} \right) \middle| X_{k} \right\},$$
(3)

where {*X_k*} represents the sequence of the state; meanwhile, {*X_k*} and *x_i* can take its values in space $\Omega = \{x_1, x_2, x_3, ..., x_r\}$.

For driving cycle synthesis, the state variables of the Markov chain usually consist of the velocity, acceleration, and the road slope or two of them. Once the current state is known, the future state for the next step is determined according to the transition probabilities which can be expressed as follows:

$$P_{ij} = P\left(X_{k+1} = x_j \mid X_k = x_i\right),$$
 (4)

where P_{ij} is the transition probability from state X_k to X_{k+1} , which can be calculated by the following equation:

$$P_{ij} = \frac{Q_{ij}}{\sum_j Q_{ij}},\tag{5}$$

where Q_{ij} represents the number of transition times from state X_k to X_{k+1} in the measured data.

Moreover, the probabilities used in the synthesis process are time-independent, and the sum of all probabilities transferring out from a certain state must satisfy the following restriction:

$$\sum_{j} P_{ij} = \sum_{j} P\left(X_{k+1} = x_j \mid X_k = x_i\right) = 1.$$
(6)

In this paper, a Markov chain model on the basis of the velocity and acceleration is employed to synthesize the driving cycle for each interval between adjacent bus stops. As shown in Figure 5(a), the measured driving data are expressed as a phase diagram of the velocity and acceleration, while segmenting into discrete bins with a constant bin width. The bin containing the driving data is considered as a state space, and the transition probability of each state can be calculated by equation (5). Then, all transition probabilities

are combined to form a transition probability matrix (TPM). When the current state is known, the possible state and its corresponding probability for the next state can be derived according to the TPM. For instance, the current state belongs to a state bin where the velocity $v(t) \in [8.8, 9)$ m/s and the acceleration $a(t) \in [0.08, 0.28)$ m/s², and the TPM, from the current state to the next 40 states, can be obtained according to the distribution of the measured data and the number of transitions between different states (see Figure 5(b)).

2.3.3. Synthesis of the Driving Cycle Based on Monte Carlo Method. When constructing the driving cycle with the Monte Carlo method, the obtained TPM of each kinematic segment dataset is transformed into a vector for each row, and equation (6) should be satisfied for each row vector. Then, the cumulative sum for each row is calculated, and each element in the newly obtained TPM is represented by a new row vector with values starting at zero and ending at one [7].

The Monte Carlo method is applied to repeatedly generate random numbers $\mu \in [0, 1]$. When the random number μ falls in an interval which satisfies equation (7), the *n*th state bin is determined for the transition of the next state, and the velocity is chosen randomly from the corresponding state bin to synthesize a velocity profile:

$$\sum_{j=1}^{n-1} P_{ij} < \mu \le \sum_{j=1}^{n} P_{ij}, \tag{7}$$

where n is the number of the state bins which is determined by each kinematic segment dataset.

The procedure of driving cycle synthesis is shown in Figure 6, and there are three points that should be further explained. Firstly, the number of μ should guarantee the requirement for the trip distance between the adjacent bus stops and also minimize the computation time. Secondly, once the beginning and the ending velocity of the generated kinematic segment are not zero, it should be processed by the method which has been mentioned in Section 2.3.1. Finally, some restrictions need to be added to ensure the higher representativeness of the synthesized driving cycle during the synthesis process, i.e., the relative error (RE) of the statistical properties for the synthesized driving cycle ought to be within 10%, compared to the measured data [7, 23].

Moreover, the road slope information of the selected route is also considered in the course of the driving cycle synthesis. Since the road altitude is an inherent property for a fixed bus route, it is not random and mainly be influenced by the travel position of the vehicle. As shown in Figure 7(a), when the vehicle is driving from the position P_1 to P_2 , the road altitude for a variational position and the actual driving distance L can be acquired by the on-board GPS. Thus, the road slope θ is described as follows:

$$\theta = \arcsin\left(\frac{H_2 - H_1}{L}\right),\tag{8}$$



FIGURE 5: Markov chain model for a kinematic segment dataset. (a) State bins of the driving data and (b) TPM of a certain state.



FIGURE 6: The procedure of a driving cycle synthesis.



FIGURE 7: The formation of the road slope for a bus route. (a) Sketch map of the road slope and (b) road slope of the bus route.

where θ is the road slope from P_1 to P_2 , H_1 and H_2 are the road altitudes of P_1 and P_2 respectively, and *L* is the travel distance from P_1 to P_2 .

Considering the errors caused by GPS measurements, the road altitude of the bus route is repeatedly measured and the mean value of the same position is taken to calculate the slope of the real road. The result of the road slope can be seen in Figure 7(b).

2.3.4. Validation of the Synthesized Driving Cycle. The synthesized driving cycle including the information of velocity and road slope is shown in Figure 8. As the road slope is derived from the repeated measurement data based on the real world, it is of great representativeness for the selected bus route. The validity of the synthesized driving cycle is primarily verified by comparing the statistical properties of the synthesized velocity and the measurement cycles, which can be seen in Table 2. The statistical properties of the measurement cycles used for validation are derived from the mean value of measurement data and most of the evaluation parameters presented in the table are selected according to the previous studies [7, 23, 32].

The relative errors of the synthesized cycle compared with the measured cycles are restricted within a maximum of 10% when constructing a representative driving cycle with the candidate constructed segments. Thereby, an acceptable cycle is obtained with its maximum relative error for -9.23% which occurs at the idle percentage. The comparison indicates that the synthesized cycle can have a good reflection of statistical properties for the measurement cycles.

Moreover, to further validate the matching degree of the synthesized cycle and measured cycles, in Figure 9, the distribution histogram of the velocity and acceleration is shown. The comparison results demonstrate that the synthesized cycle is the great resemblance of the real-world driving cycle, where the average relative errors are 9.6% and 8.3%, respectively, for the distribution of the velocity and acceleration. The results reveal that the synthesized cycle has a high degree of matching with the measured cycle in terms of velocity and acceleration.

Since the synthesized driving cycle is based on the distribution of the bus stops, the trip distance of each interval of adjacent bus stops also needs to be guaranteed. Therefore, the trip distance of each synthesized segment is compared to the real distance of the corresponding adjacent bus stops. As shown in Figure 10, the maximum error point occurs at the 59th interval where the trip distance of the synthesized segment is 263.16 m with a relative error of 3.6% compared to the desired 254 m.

3. Synthesis of the Driving Cycle considering Random Passenger Load

The random passenger load has a significant impact on the fuel economy of PHEBs. However, there are no effective conventional approaches to integrate it into the driving cycle synthesis. To develop a reliable methodology, the variation characteristics of the passenger load for different bus stops are analyzed based on the measured data.

3.1. Property Analysis of the Passenger Load for Different Bus Stops. The variation of the passenger load is an extremely complicated stochastic process, which is influenced by population, transportation, environment, climate as well as the subjectivity, etc. However, the passenger flow of the bus route is approximately stable for a period of time, when ignoring other nonirresistible reasons. In other words, the data-based method is able to reflect the passenger changing characteristics as long as the amount of data is large enough.

When the bus arrives at the *k*th bus stop, the number of passengers on the bus is described as follows:

$$N_{ij} = \sum_{i=1}^{k} (A_{ij} - B_{ij}),$$
(9)

where *i* represents the *i*th bus stop (i = 1, 2, ..., k), *j* represents sampling number of the measured data (j = 1, 2, ..., m), N_{ij} is the number of the passengers on the bus at the *i*th bus stop for the *j*th measured data, and A_{ij} and B_{ij} denote the number of the passengers getting on or off the bus, respectively, at the *i*th bus stop for the *j*th measured data.

Since it has a great imbalance between the number of passengers in off-peak and peak hours, it is difficult to identify the relationship of passengers' number between adjacent bus stops. Thus, the measured data are distinguished according to the off-peak and peak hours, as shown in Figure 11.

The database of the passengers is composed of more than 200 sampling data for off-peak and peak hours, where it has more passengers nearby the 10th and 50th bus stops compared to the fewer passengers at the 30th bus stop. This is mainly because these bus stops are located in the urban area while the 30th bus stop is the terminal located in the suburban.

As shown in Figures 12 and 13, to analyze the relationship of the passengers' number between two bus stops, the Pearson correlation coefficient, expressed as λ , is, respectively, calculated based on the measured data of off-peak and peak hours. There is a distinct difference in the number of passengers for different periods, where the maximum amount is 57 for peak hours compared to 22 for off-peak hours. However, the variation tendency of the correlation coefficient has a good consistency. When the step of the transition is increased from 1 to 10, the correlation coefficient is dramatically decreased. It means that the number of passengers between adjacent bus stops may have a certain possibility to transform; whereas two bus stops are far apart, no evidence can determine there is some kind of relationship between them.

Moreover, with the increase of the measured data, the correlation coefficient of the passengers' number between two bus stops has gradual stability, while it has the same characteristic consistent with the previous analysis. The correlation is also weakened as the step of the transition is increased (shown in Figure 14).





TABLE 2: Comparison of the statistical properties for the synthesized cycle and measured cycles.

| Parameter | Measured cycles | Synthesized cycle | Relative error (%) |
|---------------------------------------|-----------------|-------------------|--------------------|
| Max. velocity (km/h) | 56.6 | 57.56 | 1.7 |
| Mean velocity (km/h) | 18.25 | 18.21 | -0.22 |
| Mean driving velocity (km/h) | 24.73 | 24.6 | -0.53 |
| Std. dev., velocity (km/h) | 16.14 | 15.87 | -1.67 |
| Cycle distance (km) | 41 | 41.13 | 0.32 |
| Max. acceleration (m/s ²) | 1.1 | 1.12 | 1.82 |
| Min. deceleration (m/s^2) | -1.82 | -1.77 | -2.75 |
| Mean acceleration (m/s^2) | 0 | 0 | _ |
| Std. dev., acceleration (m/s^2) | 0.36 | 0.37 | 2.78 |
| Cycle duration (s) | 8162 | 8134 | -0.34 |
| Acceleration percentage (%) | 27.2 | 27.8 | 2.21 |
| Deceleration percentage (%) | 22.1 | 23.3 | 5.43 |
| Idle percentage (%) | 26 | 23.6 | -9.23 |
| Cruise percentage (%) | 24.7 | 25.3 | 2.43 |



FIGURE 9: Validation of the matching degree between the synthesized and measured cycle. Distribution of the (a) velocity and (b) the acceleration.

The analysis results demonstrate that the number of passengers for the next bus stop has a strong dependence on the current while having no significant relationship with previous bus stops. That means the variation of the passengers for different bus stops can be considered as a discrete Markov process for a period of time. Therefore, a method based on a Markov chain is proposed in this study to solve the random variation of passenger load for the bus route.

3.2. Synthesis of the Driving Cycle including Passenger Information. The random variation in the number of













FIGURE 12: Measured passenger data and the correlation analysis for off-peak hours. (a) Measured data. (b) 1-step transition. (c) 5-step transition.



FIGURE 13: Measured passenger data and the correlation analysis for peak hours. (a) Measured data. (b) 1-step transition. (c) 5-step transition.



FIGURE 14: Correlation coefficient analysis with the increase of sampling number. (a) Off-peak hours. (b) Peak hours.

passengers from one station to another is disposed of as a Markov chain model based on the historical measurements, and three steps are needed to synthesize a driving cycle for the passengers (see Figure 15).

- (i) Calculating the TPM of the Markov chain model based on the measured passenger data, respectively, for the off-peak and peak hours
- (ii) Employing the Monte Carlo simulation to generate plenty of random number μ∈ [0, 1], and determining the states corresponding to the generated μ, according to equation (7)
- (iii) The most possible state of each bus stop is derived by the maximum likelihood estimation (MLE), and the number of passengers is determined on the basis of the most possible state for the corresponding bus stop

Similar to the construction of the velocity cycle, the measured passenger data are partitioned into some discrete bins with a constant bin width, where the state bins are distinguished by bus stops and the number of passengers (see Figure 16(a)). The transition probability matrix (TPM) is calculated according to the measured data, respectively, for off-peak and peak hours. It is worth noting that it may



FIGURE 15: The procedure to synthesize a driving cycle for passengers.



FIGURE 16: Markov chain model and TPM for off-peak hours. (a) State bins of the driving data and (b) TPM for the 550th state bin.

have more than one state bin for a certain bus stop, which leads to a state transition occurring within the same bus stop. Therefore, the state bins are coded horizontally to ensure the state transitions for the adjacent state bins can only occur at different bus stops. For instance, when the state bin is at the 7th bus stop with the number of passengers including 25 and 26, the state number is coded as 550, and the next state coded as 551 is located at the 8th bus stop. The TPM, which represents the transitions from the current state to the next 50 states, is shown in Figure 16(b).

When the Markov chain model of the passenger changing is established, the Monte Carlo simulation is utilized to determine the state bins by a large amount of the generated random numbers μ . For example, when the number of μ is set to 20000, 20000 state bins are determined according to the TMP of the established Markov chain model, where some state bins are repetitive (see Figure 17). Accordingly, the most frequently appeared state bin for each bus stop is selected as the most possible state on the basis of the MLE and the corresponding number of passengers can be derived.

The number of passengers for each bus stop is derived according to different numbers of μ , respectively, for the offpeak and peak hours. More than 1,000 cycles for passenger changes are generated, respectively, for the off-peak hour and peak hours. The most representative one utilized for the fuel consumption estimation and optimization of the PHEB is selected based on the candidates according to the following restriction:

$$Ng_{i} = \frac{1}{m} \sum_{j=1}^{m} x_{ij},$$
 (10)

where Ng_i is the average number of the generated passengers for the *i*th bus stop, *m* is the sample number of the generated passenger data, and x_{ij} represents the generated passengers of *i*th bus stop for the *j*th sample.

The generated passengers and an acceptable candidate well matching to the characteristic of random passengers for the off-peak and peak hours are shown in Figure 18. It can be seen that the generated passengers of different times have an obvious difference with each other. The number of passengers during peak hours is much larger than during offpeak hours. Thus, the representative passenger flow of the bus route needs to be determined according to the following equation:

$$Np_i = \omega \cdot Ng_{n,i} + (1 - \omega) \cdot Ng_{r,i}, \qquad (11)$$



FIGURE 17: Monte Carlo simulation for the passengers.



FIGURE 18: Generated passengers of each bus stop.

where Np_{*i*} is the representative number of passengers for the *i*th bus stop, ω is the weight factor between the off-peak and peak hours, and Ng_{*n*,*i*} and Ng_{*r*,*i*} are the generated passengers of the off-peak and peak hours for the *i*th bus stop.

In general, to better reflect the realistic passenger flow of the bus route, the off-peak and peak hours need to be considered eclectically. When the weight factor ω is defined as 0.5, the representative number of the passengers is calculated by the average of the generated passengers for the off-peak and peak hours, and the calculated number should be rounded to the nearest integer. The driving cycle containing the variation information of passengers is shown in Figure 19. Accordingly, the passenger load can be acquired according to the number of passengers for each bus stop, when the mass of a passenger is defined as 68 kg [35].

4. Analysis of the Fuel Consumption Based on the Synthesized Driving Cycle

To verify the representativeness of the synthesized driving cycle, the fuel consumption of the PHEB is simulated on an established model based on the Matlab platform [43]. The flow diagram of the model is shown in Figure 20. The synthesized driving cycle is comprised of velocity, road

slope, and the passenger load. In the course of the simulation, the control commands are determined by the driver model and executed via the controller model. The driving cycle is tracked in accordance with the error between reference velocity v_{ref} and the actual velocity v_x . Meanwhile, the required torque T_{req} is supplied from both of the engine and motor (i.e., T_e and T_m) according to the rule-based control strategy, namely, charge-depletion and charge-sustaining (CD-CS) which is adopted for the researched PHEB [44]. Moreover, the motor can also work as a generator, when the required torque is negative and the state-of-charge (SOC) of the battery meets the charging demand. The main parameters of the simulation model are listed in Table 3.

4.1. Powertrain Model. In this paper, the powertrain of the studied PHEB is mainly composed of the diesel engine, electric motor, AMT, final drive, and power battery pack. The engine model is simplified as a steady-state fuel consumption model neglecting the impact of temperature and dynamic performance. As shown in Figure 21(a), the fuel consumption rate can be derived by the interpolation method from the brake-specific fuel consumption (BSFC) map.



FIGURE 19: The synthesized driving cycle including passengers.



FIGURE 20: The flow diagram of the simulation model.

| Items | Description | | |
|-------------|--|--|--|
| Vehicle | Curb/gross mass: 12,500/16,500 kg | | |
| Passengers | Max./rated: 74/17-44 | | |
| Engine | Max torque: 850 Nm | | |
| Eligilie | Max power: 162 kW | | |
| Motor | Max torque: 850 Nm | | |
| MOIOI | Max power: 130 kW | | |
| AMT | 4-speed, speed ratio: 4.13/2.45/1.49/1 | | |
| Final drive | Ratio: 6.33 | | |
| Battery | Capacity: 50 Ah, voltage: 384 V | | |

Meanwhile, the motor can also be simplified as a look-up table of the efficiency map on a hypothetical stable condition. The efficiency of the motor η_m is expressed as the function of the speed and torque, which can be described as follows:

$$\eta_m = \eta_m (n_m, T_m), \tag{12}$$

where n_m and T_m denote the rotate speed and torque of the motor, respectively. Moreover, the generator torque of the motor is limited, considering the stability and reliability of motor during brake energy recovery (see Figure 21(b)).

The lithium-ion battery is equipped to the PHEB for its higher stability and longer operating distance. The battery pack is simplified as an internal resistance battery model. As shown in Figure 22, the main working area of the battery in the simulation is restricted in the scope of the 0.35 to 0.8 according to the actual situation of the researched PHEB.

When ignoring other dynamic characteristics, the PHEB is regarded as a point-mass model. The torque requirement of the vehicle is governed by the following equation:

$$T_{\rm req} = \frac{T_e + T_m}{\eta_T} = (M_r + m_r)gf_r\cos\theta + (M_r + m_r)g\sin\theta + \frac{1}{2}C_dA\rho_du_a^2 + \frac{\delta(M_r + m_r)\cdot du_a}{dt},$$
(13)

where $T_{\rm req}$ is the torque requirement of the vehicle, T_e and T_m represent the output torque of the engine and motor, M_r is the curb mass of the vehicle, m_r is the payload which is changed according to variation of the passengers, g is the gravity acceleration, f_r is the rolling resistance coefficient, θ is the road slope, C_d is the aerodynamic drag coefficient, A is the windward area, ρ_d is the air density, u_a is the vehicle speed, δ is the correction coefficient of the rotation mass, and η_T is the efficiency of the powertrain.

4.2. Results and Discussion. To examine the impact of the passenger load on fuel consumption (FC) of the PHEB, the simulation under different constant passenger load



FIGURE 21: Characteristic diagrams. (a) Engine. (b) Motor.



FIGURE 22: Characteristic curves of the battery pack.

conditions is carried out. As shown in Figure 23, the FC is simulated, respectively, when the passenger load is predefined as no-load, half-load, and full-load, when the SOC of the battery is limited within 0.35 to 0.8. It is noteworthy that there is a conspicuous difference in the FC for different passenger loads. The FC under no-load, half-load, and full-load are, respectively, 6.23 L, 7.44 L, and 8.75 L, where the maximum deviation can reach up to 40.45% between the full-load and no-load. Therefore, it is of great challenge to give an accurate estimation of the FC in the course of the PHEB design, when the selected driving cycle does not take into account the passenger changes.

Moreover, the passenger load also has a distinct influence on the driving mode, where the PHEB with no-load has longer distance travel under the CD mode compared to the half-load and full-load. The results indicate that the passenger load has a considerable impact on the FC and the control strategy of the PHEB, and it is of great significance to synthesize a driving cycle containing the passengers via an effective methodology. As shown in Figure 24, the FC of the synthesized driving cycle is also analyzed with different initial SOC. It can be seen that, when the SOC is started from 0.8, the travel distance of the PHEB under the CD mode can approximately reach up to 10 km while it is decreased to 6 km with the initial SOC of 0.6. Particularly, the PHEB almost entirely works in the CS mode when the initial SOC is 0.4. Hence, the SOC needs to be considered in the validation.

To verify the effectiveness of the proposed method for driving cycle construction, the FC of the synthesized driving cycle is simulated and compared to the simulated FC of the later measured data. As the synthesized driving cycle needs to be representative to reflect various driving characteristics of the bus route, the FC of the measured data is obtained by the mean of multiple measurement cycle simulation results. More than 100 measured driving cycles containing the information of the velocity, road slope, and random passenger load are utilized to the simulation, while the mean and the standard deviation of the FC are analyzed with the increase of the simulated cycles.



FIGURE 23: Impact of the passenger load on fuel consumption. (a) SOC changing and (b) fuel consumption.



FIGURE 24: FC simulation of the synthesized driving cycle with different initial SOC. (a) SOC changing and (b) fuel consumption.

As shown in Figure 25(a), the mean and the standard deviation of the FC become stable with the increase of the simulated driving cycles. The mean FC is varying within ± 0.01 L/100 km when the number of the simulated cycles is more than 77, and the standard deviation at the end of the simulations is approximately 0.42 L/100 km. Moreover, the distribution of the FC shown in Figure 25(b) also reveals that the simulation results have remarkable characters of centralized distribution within the scope of 16 L/100 km to 17 L/100 km, and the proportion reaches up to 80%. Therefore, the reliable mean FC can be acquired from the simulation results.

The mean FC and standard deviation for different initial SOC are shown in Figures 26 and 27. As similar to the one with the initial SOC of 0.8, both of them gradually tend to

stabilize with the increase of simulated driving cycles. However, the stable values are larger than the case with the SOC started at 0.8. Moreover, the distribution of FC has a slight difference from each other. The smaller the initial SOC is, the lower the concentration of the distribution is. Nevertheless, it has a neglectable impact on the mean FC, and the number of measured cycles is enough to determine the mean FC for the researched bus route.

Besides, the FC based on the average of historical measurement passenger data is also analyzed and compared with the best-synthesized cycle. As shown in Figure 28, the SOC changing of them is close to each other, while the FC is different. The FC of the synthesized cycle is slightly lower than the FC of the average, for different initial battery SOC.



FIGURE 25: Mean FC, std. dev., and FC distribution for the measured data (initial SOC = 0.8). (a) Mean FC and std. dev and (b) distribution of the FC.



FIGURE 26: Mean FC, std. dev., and FC distribution for the measured data (initial SOC = 0.6). (a) Mean FC and std. dev and (b) distribution of the FC.

Table 4 gives the FC comparison between the bestsynthesized cycle and the later measured data. To illustrate the advantages of the best-synthesized cycle, the FC under different constant passenger loads is also listed out, and the relative error versus the mean FC is, respectively, compared. Moreover, the FC based on the average value of the historically measured passengers is also compared.

As shown in Figure 29, the FC under the no-load is lower than the mean, while that under the full-load is much larger. Although the FC under the half-load is closer to the mean compared to the no-load and full-load, it is still inferior to the synthesized cycle. It seems that the FC of the average measurements is better than the constant load; however, it is not as superior as the best-synthesized cycle. When the synthesized cycle is utilized to simulate the FC, the relative error is less than 2.5% for different initial SOC. More importantly, its absolute error is within the allowance of the standard deviation, which means the synthesized cycle can better reflect the variation of the passenger load. In other words, the proposed method is effective for constructing a representative driving cycle considering the random passenger load. Moreover, the initial SOC has an impact on the estimation accuracy of the FC under the same load. It is evidently



FIGURE 27: Mean FC, std. dev., and FC distribution for the measured data (initial SOC = 0.4). (a) Mean FC and std. dev and (b) distribution of the FC.



FIGURE 28: Comparison of SOC and FC between average passengers and synthesized. (a) SOC changing and (b) fuel consumption.

| | | Initial SOC = 0.8 | Initial SOC = 0.6 | Initial SOC = 0.4 |
|-------------------|----------------------|-------------------|-------------------|---------------------|
| Measured data | Mean FC (L/100 km) | 16.53 | 18.75 | 21.32 |
| | Std. dev. (L/100 km) | 0.42 | 0.44 | 0.53 |
| No-load | FC (L/100 km) | 15.15 | 17.47 | 19.99 |
| | Relative error (%) | -14.15 | -6.83 | -6.24 |
| Half-load | FC (L/100 km) | 18.09 | 20.35 | 22.65 |
| | Relative error (%) | 9.44 | 8.53 | 6.24 |
| Full-load | FC (L/100 km) | 21.27 | 23.58 | 26.40 |
| | Relative error (%) | 28.68 | 25.76 | 23.83 |
| Average | FC (L/100 km) | 17.65 | 19.97 | 22.45 |
| | Relative error (%) | 6.77 | 6.51 | 5.3 |
| Synthesized cycle | FC (L/100 km) | 16.94 | 19.09 | 21.66 |
| | Relative error (%) | 2.48 | 1.8 | 1.59 |

TABLE 4: Fuel consumption comparison of the synthesized cycle and measured data.



FIGURE 29: Relative error of the FC for different loads versus the mean FC.

improved relative to the decrease of the initial SOC. The main reason may be that the working time of the PHEB under the CS mode is increased as the initial SOC is decreased.

5. Conclusions

To take into account the random passenger load in a course of the driving cycle synthesis, this paper proposed a method to construct a typical driving cycle based on a bus route, simultaneously considering the information of the velocity, road slope, and the random passenger load. The main findings are concluded as follows:

- (1) To promote the accuracy of the synthesized driving cycle and conveniently for considering the random passenger load, kinematic segments were partitioned according to the distance of the adjacent bus stop. The velocity profile of each station-based interval was synthesized by the Markov chain Monte Carlo (MCMC) method. The validation results demonstrated that there was an acceptable deviation between the statistical properties of the synthesized driving cycle and the measured cycles, as well as the higher matching degree.
- (2) To integrate the information of the passengers into the synthesized driving cycle, the measured passenger data were divided corresponding to the offpeak and peak hours. The variation of passenger numbers between different bus stops was analyzed and discovered to have the Markov property. Therefore, the Markov chain-based method together with Monte Carlo simulation and maximum likelihood estimation was applied to determine the most possible passenger load for each bus stop.
- (3) The fuel consumption (FC) of the PHEB was simulated on the basis of the best-synthesized driving cycle while contrasting to the mean FC of the later measured data. Moreover, the influence of the passenger load and the initial SOC of the battery

were also analyzed. The results demonstrated that it was of great significance to consider the passenger load during the driving cycle synthesis, due to its considerable impact on the FC and driving mode. Moreover, the synthesized driving cycle had an excellent performance on the estimation of the FC compared to the solution that the passenger load was treated as a constant or adopting the average of the historical measured data. The relative error was only 2.48%, 1.8%, and 1.59% when the initial SOC was, respectively, defined as 0.8, 0.6, and 0.4. The synthesized driving cycle had great benefits for the PHEB design or performance estimation.

The proposed methodology is based on a fixed bus route, yet it also can be extended to other bus routes with the information of the historical driving data and the passenger flow is known. In future, the development of telematics and large data technology will greatly facilitate the acquisition of the historical driving data, and more attention will be focused on accurate driving cycle construction with sufficient data, as well as the real-time driving cycle construction.

Data Availability

The data used to support the findings of this study are included within the supplementary information files (available here).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to acknowledge the support of the National Key R&D Program of China (2018YFB1600701) and the Key Research and Development Program of Shaanxi (2019ZDLGY15-01, 2019ZDLGY15-02, and 2018ZDCXL-GY-05-03-01).

Supplementary Materials

The file named "Driving data.xls" contains part of the measured driving data for the studied bus route. The file named "Passenger data.xls" contains the collected data of the passenger numbers for different bus stops. (*Supplementary Materials*)

References

- H. Chen, X. Su, J. Li, and X. Zhong, "Effects of gasoline and polyoxymethylene dimethyl ethers blending in diesel on the combustion and emission of a common rail diesel engine," *Energy*, vol. 171, pp. 981–999, 2019.
- [2] C. M. Martinez, X. Hu, D. Cao, E. Velenis, B. Gao, and M. Wellers, "Energy management in plug-in hybrid electric vehicles: recent progress and a connected vehicles perspective," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4534–4549, 2017.
- [3] S. Stockar, V. Marano, M. Canova, G. Rizzoni, and L. Guzzella, "Energy-optimal control of plug-in hybrid electric vehicles for real-world driving cycles," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 2949–2962, 2011.
- [4] S. Xie, X. Hu, Z. Xin, and J. Brighton, "Pontryagin's minimum principle based model predictive control of energy management for a plug-in hybrid electric bus," *Applied Energy*, vol. 236, pp. 893–905, 2019.
- [5] J. Brady and M. O'Mahony, "Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas," *Applied Energy*, vol. 177, pp. 165–178, 2016.
- [6] Z. Song, J. Hou, S. Xu, M. Ouyang, and J. Li, "The influence of driving cycle characteristics on the integrated optimization of hybrid energy storage system for electric city buses," *Energy*, vol. 135, pp. 91–100, 2017.
- [7] E. Silvas, K. Hereijgers, H. Peng, T. Hofman, and M. Steinbuch, "Synthesis of realistic driving cycles with high accuracy and computational speed, including slope information," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 4118–4128, 2016.
- [8] H. Achour and A. G. Olabi, "Driving cycle developments and their impacts on energy consumption of transportation," *Journal of Cleaner Production*, vol. 112, pp. 1778–1788, 2016.
- [9] A. Esteves-Booth, T. Muneer, H. Kirby, J. Kubie, and J. Hunter, "The measurement of vehicular driving cycle within the city of Edinburgh," *Transportation Research Part D: Transport and Environment*, vol. 6, no. 3, pp. 209–220, 2001.
- [10] E. Tzirakis, K. Pitsas, F. Zannikos, and S. Stournas, "Vehicle emissions and driving cycles: comparison of the Athens driving cycle (ADC) with ECE-15 and European driving cycle (EDC)," *Global NEST Journal*, vol. 8, no. 3, pp. 282–290, 2013.
- [11] W. T. Hung, H. Y. Tong, C. P. Lee, K. Ha, and L. Y. Pao, "Development of a practical driving cycle construction methodology: a case study in Hong Kong," *Transportation Research Part D: Transport and Environment*, vol. 12, no. 2, pp. 115–128, 2007.
- [12] S. Kamble, T. Mathew, and G. Sharma, "Development of realworld driving cycle: case study of Pune, India," *Transportation Research Part D—Transport*, vol. 12, pp. 115–128, 2009.
- [13] S.-H. Ho, Y.-D. Wong, and V. W.-C. Chang, "Developing Singapore driving cycle for passenger cars to estimate fuel consumption and vehicular emissions," *Atmospheric Environment*, vol. 97, pp. 353–362, 2014.

- [14] M. A. Pouresmaeili, I. Aghayan, and S. A. Taghizadeh, "Development of Mashhad driving cycle for passenger car to model vehicle exhaust emissions calibrated using on-board measurements," *Sustainable Cities and Society*, vol. 36, pp. 12–20, 2018.
- [15] Q. Wang, H. Huo, K. He, Z. Yao, and Q. Zhang, "Characterization of vehicle driving patterns and development of driving cycles in Chinese cities," *Transportation Research Part D: Transport and Environment*, vol. 13, no. 5, pp. 289–297, 2008.
- [16] K. S. Nesamani and K. P. Subramanian, "Development of a driving cycle for intra-city buses in Chennai, India," *Atmo-spheric Environment*, vol. 45, no. 31, pp. 5469–5476, 2011.
- [17] J. Lai, L. Yu, G. Song, P. Guo, and X. Chen, "Development of city-specific driving cycles for transit buses based on VSP distributions: case of Beijing," *Journal of Transportation Engineering*, vol. 139, no. 7, pp. 749–757, 2013.
- [18] R. Günther, T. Wenzel, M. Wegner, and R. Rettig, "Big data driven dynamic driving cycle development for busses in urban public transportation," *Transportation Research Part D: Transport and Environment*, vol. 51, pp. 276–289, 2017.
- [19] Y. Peng, Y. Zhuang, and H. Yang, "Development of a representative driving cycle for urban buses based on the K-means cluster method," *Cluster Computing*, vol. 22, no. S3, pp. 6871–6880, 2019.
- [20] Y.-L. T. Nguyen, T.-D. Nghiem, A.-T. Le, and N.-D. Bui, "Development of the typical driving cycle for buses in Hanoi, Vietnam," *Journal of the Air & Waste Management Association*, vol. 69, no. 4, pp. 423–437, 2019.
- [21] P. Shen, Z. Zhao, J. Li, and X. Zhan, "Development of a typical driving cycle for an intra-city hybrid electric bus with a fixed route," *Transportation Research Part D: Transport and Environment*, vol. 59, pp. 346–360, 2018.
- [22] H. Wang, X. Zhang, and M. Ouyang, "Energy consumption of electric vehicles based on real-world driving patterns: a case study of Beijing," *Applied Energy*, vol. 157, pp. 710–719, 2015.
- [23] X. Zhao, Q. Yu, and J. Ma, "Development of a representative EV urban driving cycle based on a k-means and SVM hybrid clustering algorithm," *Journal of Advanced Transportation*, vol. 2018, Article ID 1890753, 18 pages, 2018.
- [24] X. Zhao, J. Ma, S. Wang, Y. Ye, Y. Wu, and M. Yu, "Developing an electric vehicle urban driving cycle to study differences in energy consumption," *Environmental Science* and Pollution Research, vol. 26, no. 14, pp. 13839–13853, 2019.
- [25] R. Patil, B. Adornato, and Z. Filipi, "Impact of naturalistic driving patterns on PHEV performance and system design," in *Proceedings of the SAE Technical Paper*, Detroit, MI, USA, April 2009.
- [26] J. Gonder, T. Markel, M. Thornton, and A. Simpson, "Using global positioning system travel data to assess real-world energy use of plug-in hybrid electric vehicles," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2017, no. 1, pp. 26–32, 2007.
- [27] V. Schwarzer and R. Ghorbani, "Drive cycle generation for design optimization of electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 89–97, 2013.
- [28] H. Gong, Y. Zou, Q. Yang, J. Fan, F. Sun, and D. Goehlich, "Generation of a driving cycle for battery electric vehicles: a case study of Beijing," *Energy*, vol. 150, pp. 901–912, 2018.
- [29] A. Ashtari, E. Bibeau, and S. Shahidinejad, "Using large driving record samples and a stochastic approach for realworld driving cycle construction: Winnipeg driving cycle," *Transportation Science*, vol. 48, no. 2, pp. 170–183, 2014.

- [30] L. Berzi, M. Delogu, and M. Pierini, "Development of driving cycles for electric vehicles in the context of the city of Florence," *Transportation Research Part D: Transport and Environment*, vol. 47, pp. 299–322, 2016.
- [31] T.-K. Lee, B. Adornato, and Z. S. Filipi, "Synthesis of realworld driving cycles and their use for estimating PHEV energy consumption and charging opportunities: case study for midwest/U.S," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4153–4163, 2011.
- [32] U. Galgamuwa, L. Perera, and S. Bandara, "Development of a driving cycle for Colombo, Sri Lanka: an economical approach for developing countries," *Journal of Advanced Transportation*, vol. 50, no. 7, pp. 1520–1530, 2016.
- [33] H. He, J. Guo, J. Peng, H. Tan, and C. Sun, "Real-time global driving cycle construction and the application to economy driving pro system in plug-in hybrid electric vehicles," *Energy*, vol. 152, pp. 95–107, 2018.
- [34] J. Huertas, L. Quirama, M. Giraldo, and J. Díaz, "Comparison of three methods for constructing real driving cycles," *Energies*, vol. 12, no. 4, p. 665, 2019.
- [35] K. Kivekas, J. Vepsalainen, and K. Tammi, "Stochastic driving cycle synthesis for analyzing the energy consumption of a battery electric bus," *IEEE Access*, vol. 6, pp. 55586–55598, 2018.
- [36] S. K. Mayakuntla and A. Verma, "A novel methodology for construction of driving cycles for Indian cities," *Transportation Research Part D: Transport and Environment*, vol. 65, pp. 725–735, 2018.
- [37] H. Xie, G. Tian, H. Chen, J. Wang, and Y. Huang, "A distribution density-based methodology for driving data cluster analysis: a case study for an extended-range electric city bus," *Pattern Recognition*, vol. 73, pp. 131–143, 2018.
- [38] A. Fotouhi and M. Montazeri-Gh, "Tehran driving cycle development using the k-means clustering method," *Scientia Iranica*, vol. 20, no. 2, pp. 286–293, 2013.
- [39] S. Shi, N. Lin, Y. Zhang et al., "Research on Markov property analysis of driving cycles and its application," *Transportation Research Part D: Transport and Environment*, vol. 47, pp. 171–181, 2016.
- [40] K. Kivekäs, A. Lajunen, J. Vepsäläinen, and K. Tammi, "City bus powertrain comparison: driving cycle variation and passenger load sensitivity analysis," *Energies*, vol. 11, no. 7, p. 1755, 2018.
- [41] H. He, M. Yan, C. Sun, J. Peng, M. Li, and H. Jia, "Predictive air-conditioner control for electric buses with passenger amount variation forecast," *Applied Energy*, vol. 227, pp. 249–261, 2018.
- [42] M. Zhang, S. Shi, N. Lin, and B. Yue, "High-efficiency driving cycle generation using a Markov chain evolution algorithm," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1288–1301, 2019.
- [43] X. Liu, J. Ma, X. Zhao, Y. Zhang, K. Zhang, and Y. He, "Integrated component optimization and energy management for plug-in hybrid electric buses," *Processes*, vol. 7, no. 8, p. 477, 2019.
- [44] J. Peng, H. He, and R. Xiong, "Rule based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming," *Applied Energy*, vol. 185, pp. 1633–1643, 2017.