ARTICLE IN PRESS
IATSSR-00310; No of Pages 10

IATSS Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

IATSS Research

Research Article

Development of *E*-rickshaw driving cycle (ERDC) based on micro-trip segments using random selection and K-means clustering techniques

C. Chandrashekar [a], Prashansa Agrawal [b], Pritha Chatterjee [a], Digvijay S. Pawar [a,*]

[a] *Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi, Sangareddy 502285, India*
[b] *Department of Computer Science and Engineering, Indian Institute of Technology Hyderabad, Kandi, Medak 502285, India*

A R T I C L E I N F O

Article history:
Received 23 July 2020
Received in revised form 26 May 2021
Accepted 8 July 2021
Available online xxxx

Keywords:
E-rickshaw
Driving cycle
Speed profile
Acceleration/deceleration
Urban
Rural

A B S T R A C T

In India, auto rickshaws are the most convenient and cheapest mode of near-to-door transport in both rural and urban areas. Such vehicles powered with internal combustion engines (ICEs) are one of the main sources of pollutants on urban corridors. One way to minimize tail-pipe emissions is to use electric motors in place of ICE. To evaluate the vehicle performance, energy consumption, driving behavior, optimal design and management of such electric vehicles, driving cycle is an important tool. So far, only limited studies exist on the development of a driving cycle for e-rickshaw. Moreover, these studies are concentrated in urban traffic environment and research accounting rural and urban environment together remain unexplored. In this study, real world driving data for 100 trips of e-rickshaw are collected on a road stretch passing through rural and urban setting. A high-end GPS data logger was used to collect vehicle kinematics such as continuous speed profile, acceleration/deceleration, heading, and vehicle position coordinates. Nine different driving characteristics representing actual traffic conditions are identified and used for developing e-rickshaw driving cycle (ERDC). Two approaches, random selection and k-means clustering are explored to arrive at best representative ERDC using micro-trips technique. The analysis results revealed that k-means clustering outperforms the random selection method with additional benefit of accounting traffic conditions systematically. The insights from this study can be used to understand and model the performance of e-rickshaw, in terms of energy consumption and driving characteristics, compared to other fossil-fuel driven automobiles.

1. Introduction

The survey by National Sample Survey Office (NSSO 2016), Ministry of Statistics and Programme Implementation (MoSPI) revealed that auto rickshaws are the most preferred mode of transport after buses in both rural and urban India [1,2]. According to the survey (June 2014–July 2015), auto rickshaws are used by about 38% of rural households and about 47% of urban households (NSSO 2016), and around half a million auto-rickshaws are registered every year (MORTH 2020) [3]. In state of Telangana alone, there are around 435,507 registered auto-rickshaws as of March 2020 (RTA Telangana) [4]. The auto-rickshaws, generally three-wheelers, with both four and two-stroke internal combustion engines (ICE) are fueled by diesel, compressed natural gas (CNG) and liquefied petroleum gas (LPG). Different fuel-engine combination is prevalent in different cities for various reasons. For example, Delhi is dominated by CNG-fueled four stroke auto rickshaws due to government fuel mandate, while Hyderabad, Pune and Bangalore are serviced by LPG-fueled two and four stroke auto rickshaws [5]. According to the recent study by the energy and resources institute (TERI) an average conventional LPG auto rickshaw emits approximately 0.005 t of particulate matter-10 (PM_{10}) and about 3.72 t of carbon dioxide in a year (TERI, 2018) [5]. Due to the rapidly increasing number of auto rickshaws and limited use of emission control strategies, auto-rickshaw are emerging as a major source of vehicular pollution in both urban and rural areas in India. To counteract this issue, researchers are looking for alternative fuel sources as a replacement of conventional fuels along with LPG and CNG driven vehicles.

Electric vehicle is considered to be promising alternative vehicle technology with the potential to lessen the greenhouse gas emission and other air pollutants [6]. Particularly battery electric vehicle such as electric rickshaw (e-rickshaw) has the benefits of high energy efficiency and zero-tailpipe emissions which are suitable for short distance commute within the city. India is home to about 1.5 million battery powered 3-wheeled e-rickshaw with 11,000 new e-rickshaws hitting the streets every month [7]. Also, the annual sales of e-rickshaw are expected to

* Corresponding author.
E-mail addresses: ce18resch11009@iith.ac.in (C. Chandrashekar), cs18mtech11037@iith.ac.in (P. Agrawal), pritha@ce.iith.ac.in (P. Chatterjee), dspawar@ce.iith.ac.in (D.S. Pawar).
Peer review under responsibility of International Association of Traffic and Safety Sciences.

https://doi.org/10.1016/j.iatssr.2021.07.001

Please cite this article as: C. Chandrashekar, P. Agrawal, P. Chatterjee, et al., Development of *E*-rickshaw driving cycle (ERDC) based on micro-trip segments using random selection and K-means clustering techniques..., IATSS Research, https://doi.org/10.1016/j.iatssr.2021.07.001

increase by 9% at the end of 2021 [8]. Since e-rickshaw popularity is rapidly increasing across cities in India. It is necessary to understand the impact of e-rickshaw on the Indian electric grid which depends on design of powertrain, and energy storage management system, in turn, depends on factors like type of road (hilly or plain terrain), traffic conditions and driving behavior [9]. A driving cycle is often used by researchers and practicing engineers to understand and model the driving behavior, emission pattern and energy consumption for different vehicle types [10]. Driving cycle is a representative speed-time profile that reflects the typical driving pattern of a given city or region [11,12,26]. A driving cycle consists of a sequence of vehicle operating conditions (idle, acceleration, cruise, and deceleration) [20]. However, operating conditions may vary across cities due to variations in topography, vehicle type under consideration, vehicle composition, and road type [12]. The main objective of the present study is to develop a driving cycle for e-rickshaw in rural and urban traffic conditions. The developed driving cycle is further used to determine driving characteristics of e-rickshaw and is compared with other typical driving cycles.

2. Literature review

An overview of various methods and techniques used in the previous studies for construction of driving cycles is discussed in this section. Further the significance of various driving cycles developed in different studies is also deliberated in this section.

There are several standard driving cycles such as the federal test procedure (FTP) -75 cycle that represents urban driving including frequent stops for light duty vehicles in the USA [13], highway fuel economy test cycle (HWFET) used to assess fuel economy over a highway in USA [14], Japanese driving cycle used for emission and fuel consumption certification in Japan among many others [15]. Different driving cycles are of varying duration, for example, Sydney cycle has a duration of 637 s [16] whereas, Singapore cycle has a duration of 2344 s [17]. Driving cycle varies with data collection technique. Chase car technique and on-board measurement technique are the most commonly used methods of data collection. In chase car method, the target vehicle is followed externally to collect speed time data, whereas in on-board measurement technique an instrument is fixed in the target vehicle to collect the speed time data. Different methods are used to develop driving cycle from the collected data, such as micro-trip based cycle construction, segment based cycle construction, pattern classification method and modal cycle construction [18,19]. Kamble et al. [20] developed a driving cycle using micro-trips extracted from real-world driving data for Pune, a tier I city in India. In their study, time space-profile, percentage acceleration/deceleration, idle time, cruise time, and the average speed were considered for the better representation of heterogeneous traffic conditions prevalent in developing countries like India. Segment based cycle construction is also similar to micro-trip based cycle construction, however, it is dependent on road type and level of service, and is not suitable for traffic condition with frequent stops [18,19]. Nesamani and Subramanian [19] developed a distance-based driving cycle for intracity buses in Chennai, a metro-city in India, and the developed driving cycle was compared with different international driving cycles. Arun et al. [21] developed a driving cycle for passenger car and motorcycle in urban roads and identified the significant difference between peak and off-peak travel time for passenger cars.

In cycle construction with pattern classification method, the trip is divided, considering trip kinematic situation, into heterogeneous classes using statistical method and is helpful for traffic pattern study [19]. In the modal cycle construction method, driving cycle is formed using markov chain theory. However, this method fails to develop a candidate driving cycle that closely matches with the population parameters [18].

The well-known techniques used for the construction of the driving cycle includes random selection of micro-trips, quasi random selection of micro-trips, clustering techniques and, markov chain theory. In the quasi-random approach, micro-trip selection is an incremental process which continues till it reaches the desired duration of driving cycle. Unlike random selection, the driving cycles whose speed acceleration frequency distribution (SAFD) matches with the SAFD of entire data is selected as the best driving cycle in quasi random selection technique [22]. A major concern related to the quasi random selection of micro-trip is that the micro-trip does not replicate a modal activity, actual frequency, duration and intensity. To avoid this problem, the micro-trips with equal probability should be chosen [22]. Markov chain technique is generally used to calculate the probability of transition between different types of snippets classified by the maximum likelihood estimation or other clustering techniques. A two- dimensional markov chain containing the information of speed and acceleration was introduced and employed to synthesize the driving cycle [23]. The disadvantage of the markov chain is the accuracy and time efficiency are usually conflicting. 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 [24]. Wang et al. [25] developed a driving cycle for Beijing city, stating that the random selection technique outperformed the markov chain. In addition, the k-means technique has flexibility to face different types of driving data [26].

Researchers have recommended the need for dedicated driving cycle for a particular city/region [19] and also for different vehicle types [21]. Though there have been quite a few studies on driving cycles for different vehicle types, limited research is available on the development of driving cycle for e-rickshaw [27]. Berzi et al. [28] developed a driving cycle, for energy consumption and efficiency assessment using an electric vehicle in the city of Florence. A set of electric vehicle driving cycles was presented as final product. Brady and Mahony [29] developed a driving cycle using driving data from a fleet of electric vehicles in Dublin city, Ireland. Markov chain clustering algorithm was used for classification of different driving conditions based on speed and acceleration. Zhao et al. [30] developed an urban driving cycle for Xi'an city (China) using hybrid data collection method. Analysis results revealed that representative electric vehicle driving cycles for each typical city and region is important for accurate estimation of the energy. Pathak et al. [31] developed a driving cycle for battery operated three wheeled rickshaws in Delhi using micro-trip based random selection method. The length of the driving cycle obtained was 4.13 km with a duration of 1125 s.

Most of the studies in India on driving cycles were developed for internal combustion engine vehicles (ICEVs), and very few studies are available on development of driving cycle for e-rickshaw in India. Moreover, no studies have developed driving cycle for e-rickshaws in urban and rural settings. In the present study, random selection and k-means clustering techniques are used to develop driving cycle for e-rickshaw in rural and urban traffic conditions.

3. Research methodology

After considering the prevalent limitation in other techniques, the present study used random selection and k-means clustering techniques to select micro-trips from the entire data pool to arrive the best representative driving cycle for e-rickshaw. The sequence of driving data between two successive stops in the trip is defined as micro-trip [32]. The micro-trips selected, using both the techniques are then compared for relative errors. The relative error is the percentage deviation of all the driving characteristics (V_{ag}, V_{acc}, V_{dcc}, T_i, T_a, T_d, SD_v, SD_{acc}) between the 22 candidate driving cycles and the entire data. The best candidate driving cycle is selected, if the percentage deviation of all the driving characteristics is less than desired error i.e., 10%. The detailed methodology adopted in this study to develop real-world driving cycle for e-rickshaw is shown in Fig. 1

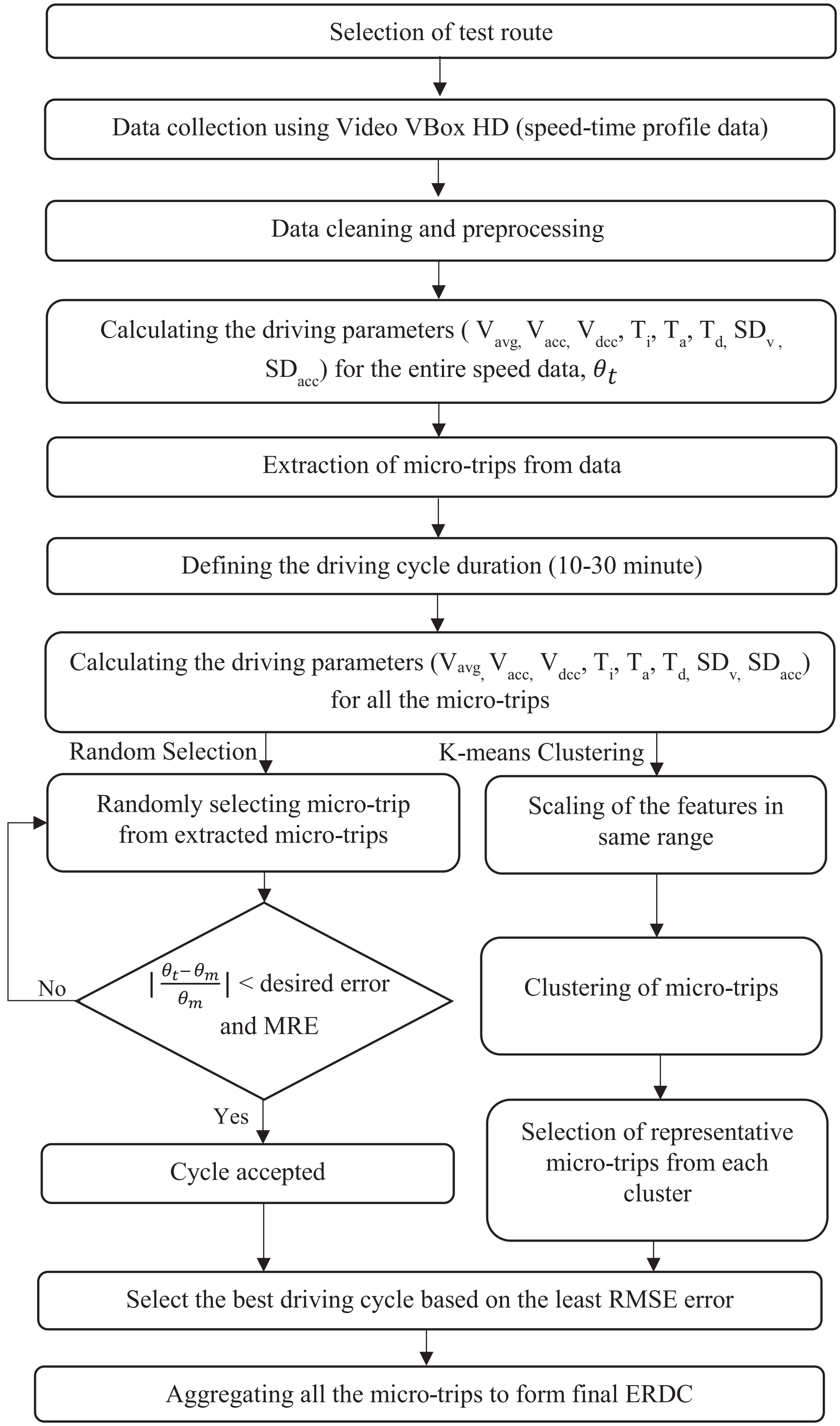

Fig. 1. Flow chart of methodology adopted to develop ERDC.

3.1. Data collection

An electric rickshaw (Mahindra e-Alfa) was used as a test vehicle for data collection (Fig. 2). Specifications of the test vehicle used for the study are shown in Table 1. A GPS data logger (video Vbox HD2) was mounted on the dashboard of the test vehicle. Speed-time profile and acceleration/deceleration characteristics were recorded at a frequency of 10 Hz. Data was collected in morning and evening peak hours, for a total 100 trips, in the months of November and December 2019. The speed-time data collected during peak period influences driving characteristics due to traffic congestion, critical driving conditions (frequent interaction with the vehicle), and some important driving situations (sharp acceleration and deceleration). Therefore, the data collected during the peak hours will capture the critical driving conditions that may have high impact on energy consumption. In total five professional drivers having rich driving experience, proper knowledge of e-rickshaws, familiarity with road conditions, and stable driving were selected for data collection through convenience sampling.

Table 1
Specifications of e-rickshaw used in the study.

Parameters	Specification
Model and make	Mahindra e-Alfa
Seating capacity	Driver +3 person +50 kg luggage
Battery type and consumption	Lead-acid, 5–6 units
Battery charging time	10–12 h
Overall height (m)	1.8
Gross vehicle weight (kg)	758
Max speed	30 kmph
Battery capacity	48 V, 120 Ah
Range per charge (10*h*)	80 km
Charger	15 Amp, Copper
Maximum power of the motor	1000 W (DC motor)
Front and rear brake	Drum
Parking brake	Mechanical cable connected with rear-wheel brake

3.2. Route selection

A 14 km long study stretch, located in Sangareddy district, state of Telangana, India, was selected for the driving cycle data collection. The study stretch comprised of feeder, rural and urban road characterized by heterogeneous traffic typical of developing countries like India Fig. 3 depicts the location of different geometric elements such as mid-block openings, uncontrolled intersections and signalized intersections on the study stretch. The detailed route description of the study stretch is shown in Table 2.

4. Development of driving cycle

4.1. Random selection technique

Random selection technique is widely used for constructing candidate driving cycles [11,33]. This technique is popular for its simplicity and for its effective results. Random selection of micro-trip methods also ensures that the speed-time profiles of the cycle constructed from the real-world data reflects the proper proportions of broad range of vehicle operation [33,34]. The micro-trips that are grouped and arranged to form a representative cycle is stated as candidate driving cycle, in which several micro-trips are randomly selected to form the best representative driving cycle [31]. Each micro-trip in the entire data set has an equal probability of being selected, and they are joined in series to form a driving cycle until target travel time is reached. The candidate driving cycle with minimum error is selected as best e-rickshaw driving cycle (ERDC). The following steps were followed to arrive at the final ERDC using random selection technique.

Step 1: Selection of driving characteristics

The characteristics that describe actual driving pattern are termed as driving characteristics [35]. Several speed time-based variables such as average speed, percentage of idling time, average acceleration/deceleration, etc., are found to describe driving pattern and influence the vehicle performance [36]. Moreover, average speed alone is insufficient to assess and compare driving cycles, hence other driving characteristics such as maximum/minimum speed and acceleration, mean and standard deviation of speed and acceleration, occurrence of stop (idle time) etc. are introduced [18]. The driving characteristics considered in the study are used as selection criteria to minimize the difference between the candidate driving cycles and the entire collected data. Considering the various driving characteristics used by different countries, the selected driving characteristics for the present study are listed in Table 3. The selected driving characteristics can be correlated to vehicle energy consumption, emission factors, traffic conditions and driving behavior [11,19,21,28,32,37].

Fig. 2. Instrumented test vehicle used for the study.

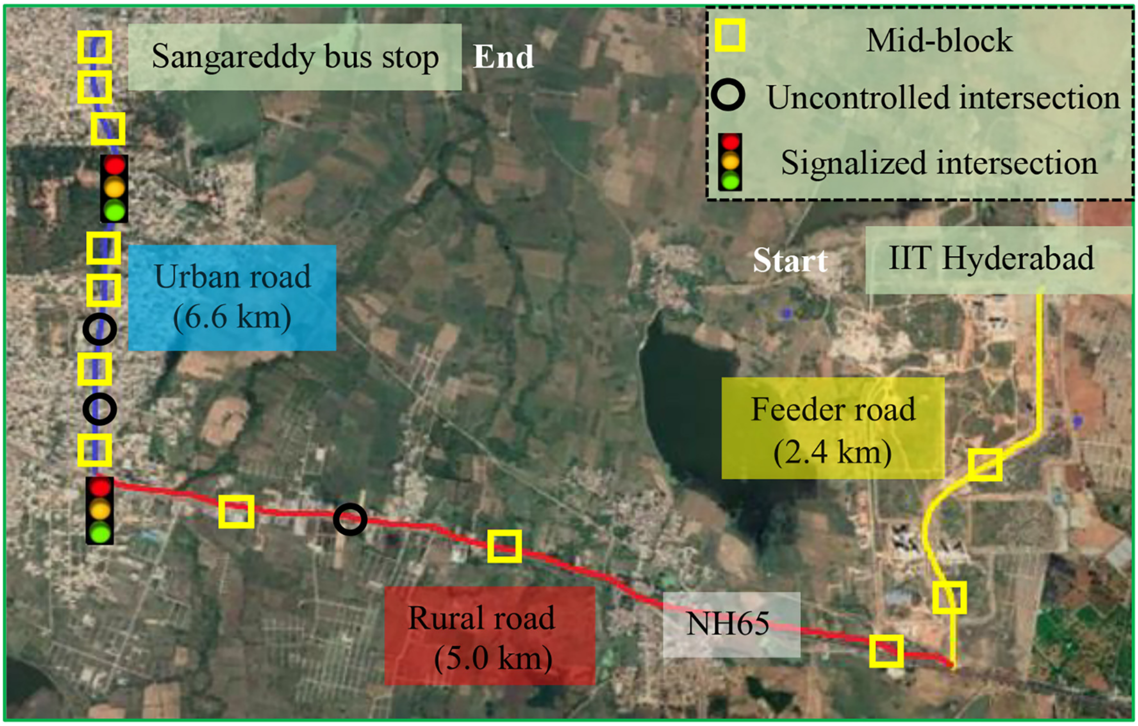

Fig. 3. Google map view of the test route.

Table 2
Route description.

Sr. No	Road type	Road length (km)	Posted Speed limit (km/h)	No. of signalized intersections	No. of unsignalized intersections	No. of Mid-block openings
1	Feeder road	2.4	30	–	2	2
2	Urban road	6.6	40	2	2	7
3	Rural road	5.0	60 to 80	–	1	3

Step 2: Data cleaning

The on-board measurement device (video V-box HD 2) was used for collecting vehicle trajectory data and speed related parameters. The trajectory data from GPS device are often prone to errors mainly due to error in user equivalent range [38]. The spikes in speed data, irregular data and negative speed data dropouts are removed by interpolation [39]. To achieve this, each speed value in the data-set is processed through a filter and compared individually to chosen high or low speed limits. If the data point is found to lie outside the chosen limits, the Savitzky-Golay filter replaces these data points with the speed information derived from the neighboring data. The Savitzky-Golay technique is based on the least squares polynomial fitting across a moving window within the data in the time domain [40]. Fig. 4 shows comparison of filtered and non-filtered speed data samples.

Step 3: Synthesis of the driving cycle

The mean values of the driving characteristics shown in Table 3 were calculated for the entire data. All possible micro-trips (idle to idle speed) were identified for the entire data. A target travel time of 600–1800 s was considered in this study so that a) the driving cycle duration is not too short and has sufficient micro-trips to reflect the real-world driving pattern and b) it is not too long making it impractical for dynamometer experiment and for emission testing standards [12].

An 'R' program was developed to randomly select the micro-trips until it reached the target travel time. A total of twenty-two driving

Table 3
Driving characteristics considered for the development of driving cycle.

Sr·No	Diving characteristics	Unit
1	Average speed, (V_{avg})	kmph
2	Maximum speed, (V_{max})	kmph
3	Average acceleration, (V_{acc})	m/s^2
4	Average deceleration, (V_{dcc})	m/s^2
5	Standard deviation of velocity, (SD_v)	kmph
6	Standard deviation of acceleration, (SD_{acc})	m/s^2
7	Time proportion of driving mode in idling (T_i), (when speed is zero)	%
8	Time proportion of driving mode in acceleration (T_a), (when speed >3 kmph and acceleration > 0.1 m/s^2)	%
9	Time proportion of driving mode in deceleration (T_d), same as acceleration mode except that acceleration should be negative	%

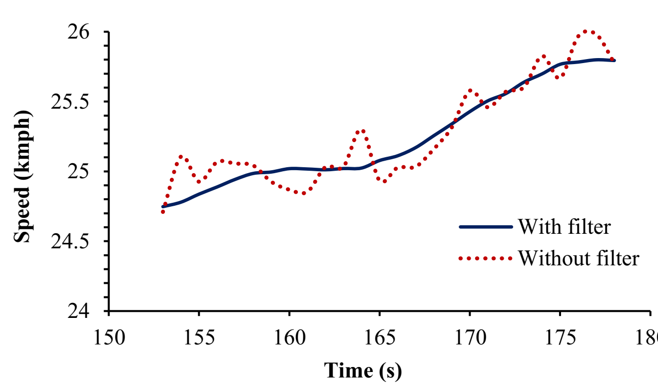

Fig. 4. Comparison of speed time data with and without filter using Savitzy-Golay filter.

cycles were obtained with different combination of micro-trips, which are considered as candidate driving cycles. Relative error (RE) of the driving characteristics between the candidate driving cycles and entire speed data was calculated using Eq. (1). The candidate driving cycles with RE of all the driving characteristics less than 10% were selected [19]. Finally, mean relative error (MRE) was calculated using Eq. (2) and the final driving cycle was selected based on the least MRE. If there were more than one driving cycles with the same MRE [41], root mean square error (RMSE) (Eq. (3)) [19] was calculated for the speed acceleration frequency distribution (SAFD) between candidate driving cycles and entire data [41]. Finally, a driving cycle with minimum RMSE value is selected as best driving cycle.

$$RE = \left|\frac{\theta_t - \theta_m}{\theta_t}\right| * 100\ (\%) \tag{1}$$

where

θ_t= Driving characteristics for the entire data; θ_m= Driving characteristics of candidate driving cycles

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\theta_t - \theta_m}{\theta_t}\right) \tag{2}$$

where

n = Number of all driving characteristics

$$RMSE = \sqrt{\frac{1}{P*Q}\sum_{j=1}^{P}\sum_{i=1}^{Q}\left(m_{ij} - n_{ij}\right)^2} \tag{3}$$

where

Q and P are the number of speed bins and acceleration bins respectively, and m_{ij} and n_{ij} are the frequency values of the ijth bin of the candidate driving cycle entire data, respectively.

4.2. K-means clustering technique

The k-means clustering technique is one of the most widely used techniques in solving the clustering problems because of its simplicity, ease in implementation, and good interpretability [43]. It is essentially a technique to group the samples according to the data similarity without a given classification category. K-means clustering is unsupervised learning techniques used to get intuition of the structure of the data. It is centroid based approach which tries to cluster data points on the basis of similarity in terms of Euclidean. The algorithm assigns the points to the cluster such that the sum of squared distance between data point and the cluster centroid is minimum. k-means clustering algorithm works as follows:

a) Select k points randomly from the data points (without replacement) and call them centroid of the clusters, where k is number of clusters specified
b) Assign each data point to the closest centroid
c) Compute the new centroids by taking average of each cluster
d) Keep iterating steps b and c until no change in the centroids

Here, extracted micro-trips are clustered based upon their driving characteristics such as average speed, average acceleration, idle time percentage, acceleration and deceleration time percentage. Each cluster represents a particular traffic condition. Unlike random selection technique, where micro-trips are randomly selected, here, representative micro-trips are selected from each cluster based upon their closeness to the cluster centers. Number of representative micro-trips may differ according to the cluster size with respect to the entire data. Fig. 5 shows the stepwise procedure used to arrive at final ERDC using k-means clustering.

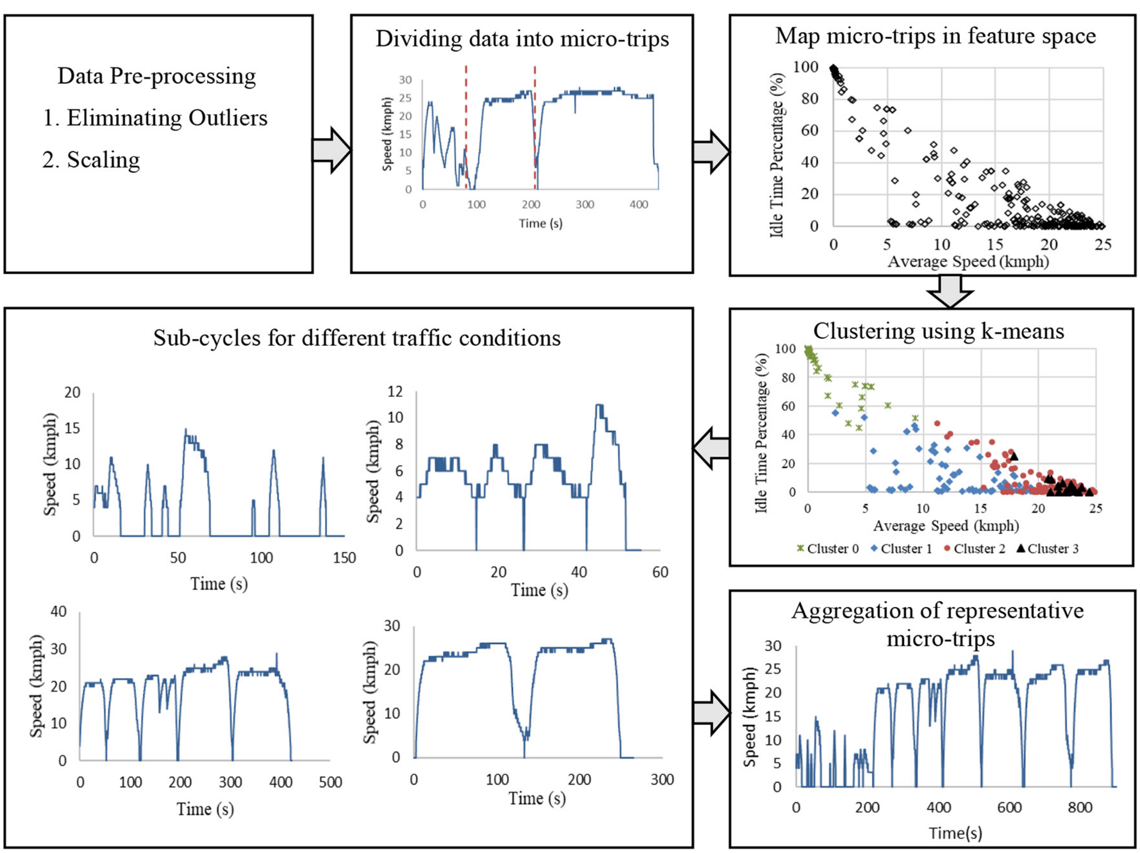

Fig. 5. Driving cycle development using k-means cluster technique.

Step 1: Preprocessing of data

As mentioned earlier in Section 3, the outliers and irregular spikes in GPS speed-time profile data were removed using the Savitzy-Golay filter technique. The second major concern while clustering the micro-trips is the effect of scale of magnitude in the feature map. For instance, the speed of an e rickshaw ranges from 0 to 30 kmph, but the acceleration may range from -3.5 to $+3.5$ m/s^2. Thus, while clustering, speed may get more importance compared to acceleration due to difference in the range of magnitudes. Also, while estimating errors such as MRE, we give equal importance to the magnitude. So, in order to give equal importance to all the features while clustering, they are scaled in the same range.

Step 2: Selection of optimal number of clusters

One of the most common heuristic approach to select optimal number of clusters is "Elbow Method". The method consists of plotting intra cluster variation as the function of number of clusters. The basic idea is to choose the number of clusters so that total within cluster sum of square (WSS) is minimized and adding new cluster doesn't improve WSS further. Fig. 6 shows the plot for Elbow method to determine optimal number of clusters.

Here the scoring parameter is set to distortion, which computes the sum of squared distances from each point to its centroid. From Fig. 6, it can be observed that the knee point is occurring at score = 258, where number of clusters are 4. Therefore, number of clusters in k-means clustering is specified as 4.

Step 3: Micro-trips extraction and clustering of micro-trips

Entire driving data of 100 trips were divided into small fragments called micro-trips. Driving characteristics/features of all the micro-trips are extracted and mapped as points in the feature space. As shown in Fig. 7, each data point represents a micro-trip. Data points are then clustered using k-means clustering technique. Parameters used for clustering the micro-trips are average speed, average acceleration, idle time percentage, percent of acceleration and deceleration time.

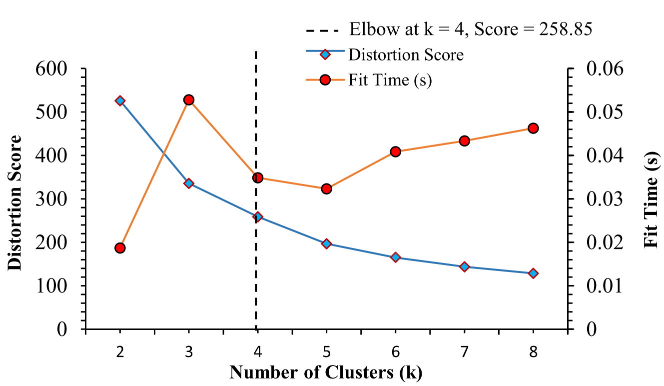

Fig. 6. Selection of number of clusters using Elbow Method.

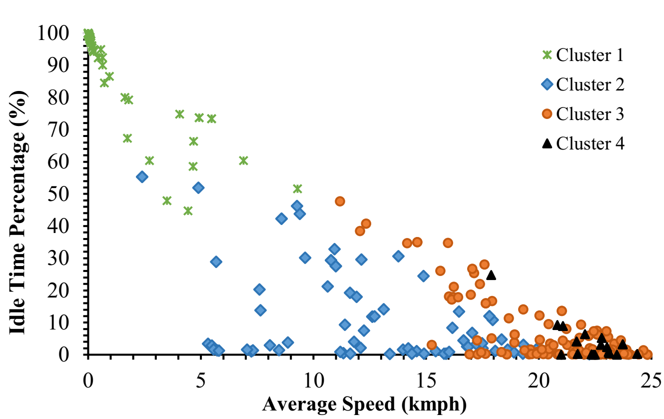

Fig. 7. Clustering of the micro-trips.

Each cluster has its own driving characteristics and represents different traffic conditions as follows: a) Congested traffic condition with relatively zero flow, because of frequent stops and high idle time with low average speed indicated by cluster 1 which was observed near signalized intersections in this study. b) Urban traffic condition characterized by high idle time with low average speed indicated as cluster 2, c) Rural traffic condition with relatively free flows and moderate idle time and moderate average speed indicated by cluster 3, c) Feeder traffic condition with complete free flows with low idle time and high average speed indicated by cluster 4.

Step 4: Selection of representative micro-trips and formation of final driving cycle

The driving cycle is combination of different traffic conditions, so we need to select representative micro-trips from each traffic condition. The share of each traffic condition in the final driving cycle should be proportional to its total time duration in the entire data. To calculate the total duration of each cluster in the final driving cycle, we use the ratio of total duration of that particular cluster to the duration of total data. This was formulated using Eq. (4) [27].

$$K_i = \frac{K_{driving\ cycle}}{K_{entire\ data}} \sum_{j=1}^{n_i} K_{ij} \tag{4}$$

$K_i = duration\ of\ the\ cluster\ number\ i$

$i = 1\ to\ total\ number\ of\ clusters\ in\ the\ final\ driving\ cycle$

$K_{driving\ cycle} = duration\ of\ the\ final\ driving\ cycle$

$K_{entire\ data} = duration\ of\ the\ entire\ data$

$K_{ij} = time\ of\ the\ micro\ trip\ number\ j\ in\ the\ cluster\ number\ i\ and\ n_i$ $is\ the\ total\ number\ of\ microtrips\ in\ the\ cluster\ number\ i$

To select micro-trips from each cluster, it is important to choose the micro-trips that best represent the cluster i.e. in the feature space, we need to select the data points that best define the cluster. The simplest approach to implement this is to pick the data points closest to the cluster centers based on their Euclidian distance. So, to select the representative micro-trips, cluster centers are calculated and the micro-trips closer to the cluster center are included until the total share of that cluster, calculated from Eq. (4), is reached.

5. Results and discussion

5.1. ERDC from random selection technique

In this technique, several micro-trips were randomly selected to form a candidate driving cycle. Out of 22 candidate driving cycles, ERDC which satisfies the minimum error criteria is selected as the final driving cycle. Fig. 8 shows the final ERDC which has a cycle time of 980 s with 5.84 km of distance covered. In Fig. 8, the sudden drop in the speed at 207 to 223 s is due to vehicle waiting to merge on to rural highway from the feeder road, whereas the sudden drop from 511 to 566 s is due to presence of a signalized intersection It can be concluded that the major driving patterns in the route selected are short driving trips with low speed and frequent fluctuation in speed which is typical behavior on urban and rural roads for para transit mode like auto-rickshaws. Such frequent stops and high fluctuation in instantaneous speed can cause increased energy and power consumption [42].

Table 4 shows the comparison of driving characteristics between entire data and developed ERDC using random selection technique. The values of driving characteristics of ERDC were found to represent the

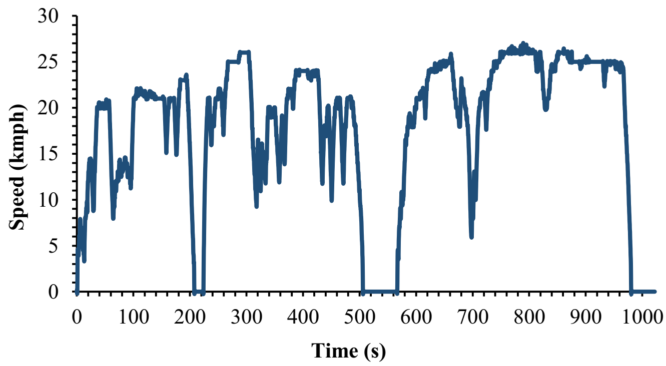

Fig. 8. ERDC developed using random selection technique.

Table 4
Comparison of driving characteristics between entire data and developed ERDC.

Driving characteristics	Unit	Entire data	ERDC	RE	MRE	RMSE
Average speed, (V_{avg})	kmph	18.30	17.40	4.91		
Maximum speed, (V_{max})	kmph	28.40	27.00	4.92		
Average acceleration (V_a)	m/s^2	1.14	1.11	2.63		
Average deceleration, (V_d)	m/s^2	1.16	1.14	1.72		
Std. dev. of speed, (SD_v)	m/s^2	9.39	8.60	8.41		
Std. dev. of acceleration, (SD_a)	m/s^2	0.490	0.460	5.91	4.84%	1.068%
Time proportion of driving modes in idling, (T_i)	%	16.31	15.59	4.41		
Time proportion of driving mode in acceleration, (T_a)	%	22.43	21.68	3.34		
Time proportion of driving mode in deceleration, (T_d)	%	21.92	20.73	5.42		
Time proportion of driving mode in cruise, (T_c)	%	39.34	42.00	6.76		

actual data with relative error ranging from approximately 1 to 8% for different driving characteristics. The candidate driving cycle with minimum MRE of 4.84 and minimum RMSE of 1.068 was selected as the ERDC.

Fig. 9 (a) and (b) shows the speed acceleration probability distribution for entire data and developed ERDC respectively, with the frequency on the vertical axis and the speed and acceleration in the horizontal plane. It is observed that SAFD of ERDC represents the SAFD of the entire data closely. The analysis of the micro-trips shows that about 28.62% of driving situations in the selected route are trips with average speed less than 15 kmph. In particular, those with speed less than 10 kmph alone account for 19% of total trips. When combined with the time spent at idle speed, 16% of the total time of micro-trips was spent at lower average speed intervals (less than 5 kmph), whereas the cruising time was 39% for the selected study stretch. It was also observed that the percentage of time spent by electric rickshaw in acceleration (22.43%) is almost equal to the percentage of time spent in deceleration (21.92%).

5.2. ERDC from k-means clustering technique

The final driving cycle developed using k-means clustering is shown in Fig. 10. The developed driving cycle has a duration of 940 s and a cycle length of 5.26 km. The analysis of the clustering based micro-trips shows that about 35% of driving situations in the selected route are trips with average speed less than 15 kmph. In particular, those with speed less than 10 kmph alone account for 29% of total trips. When combined with the time spent at idle speed, 16% of the total time of micro-trips was spent at low average speed intervals (less than 2 kmph).

Table 5 gives the comparison of driving characteristics between the entire data and ERDC developed using k-means clustering technique. The relative error for all the driving characteristics was found to vary between 1 and 8% with candidate driving cycle having MRE of 3.617 and the RMSE of 1.068. Driving characteristics such as maximum speed, average acceleration, average deceleration, standard deviation of speed, standard deviation of acceleration was found to have low relative error compared those from random selection technique. Fig. 11 (a) and (b) depicts the SAFD of entire data and SAFD of ERDC obtained from the k-mean clustering technique. The SAFD using k-means is found to represent SAFD of entire data closely with peaks at a speed of 25 kmph and 0 kmph at an acceleration range of (-1 to 0 m/s^2). The percentage of idle time was found to be 16% and the cruising time was 42% for the selected study stretch. It was also observed that the percentage of time spent by electric rickshaw in acceleration (21.47%) is almost equal to the percentage of time spent on deceleration time (20.40%), similar to the ERDC obtained from random selection of micro-trips.

6. Comparison of driving cycles

The ERDC obtained from the k-means cluster technique which has less error (REs and MRE) compared to random selection technique is used for comparison with other typical driving cycles. The parameters stated in Table 6 were selected to compare ERDC with other cities driving cycle. The driving cycle considered for the comparison are Electric Vehicle Indian driving cycle (EVIDC) (Battery operated rickshaw with a Max power of motor with 1140 W, and curb weight of 361 kg) [31], Xi'an driving cycle (BYD E6 pure EV with a Max power of 120 kW and a curb weight of 2380 kg) [41] and Florence driving cycle (Flr DC, small size M1 EV with curb weight of 1100 kg and power of 50 kW) [28]. The average acceleration and deceleration of ERDC is 33% higher than Xi'an and Flr DC and 82% higher than EVIDC. Proportion of time spent in acceleration mode for ERDC was found to be 11% lower than EVIDC, 35% lower than Xi'an and 44.8% lower than Flr DC. Proportion of time spent in deceleration mode for ERDC was found to be almost

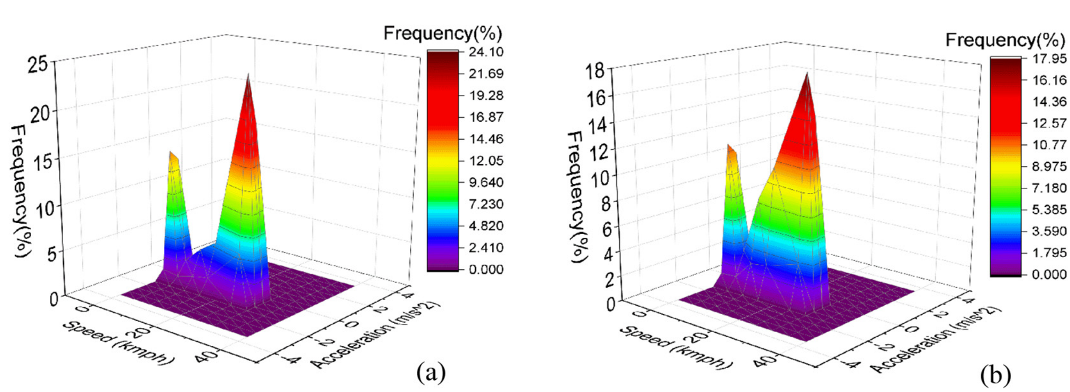

Fig. 9. (a) Speed acceleration frequency distribution (SAFD) plot for entire data (b) SAFD plot for ERDC using random selection technique.

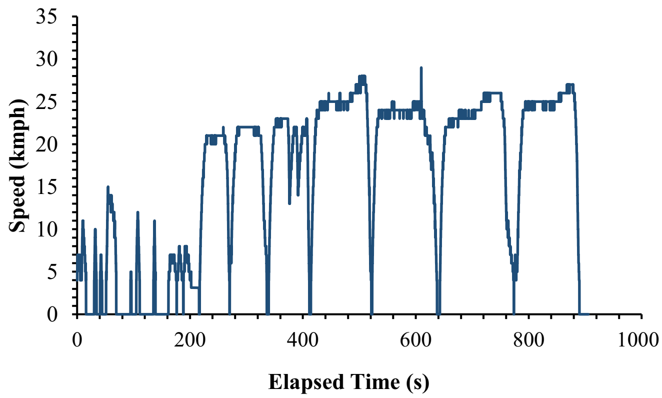

Fig. 10. ERDC developed using k-means clustering technique.

Table 5
Comparison of driving characteristics between entire data and ERDC.

Driving characteristics	Unit	Entire data	ERDC	RE	MRE	RMSE
Average speed, (V_{avg})	kmph	18.30	17.40	5.43		
Maximum speed, (V_{max})	kmph	28.40	28.00	1.42		
Average acceleration, (V_a)	m/s^2	1.14	1.12	1.75		
Average deceleration, (V_d)	m/s^2	1.16	1.15	0.86		
Std. dev. of speed, (SD_v)	kmph	9.39	9.57	1.91		
Std. dev. of acceleration, (SD_a)	m/s^2	0.490	0.481	1.83	3.617%	1.068%
Time proportion of driving mode in idle, (T_i)	%	16.31	15.7	3.74		
Time proportion of driving mode in acceleration, (T_a)	%	22.43	21.47	4.27		
Time proportion of driving mode for deceleration, (T_d)	%	21.92	20.40	6.93		
Time proportion of driving mode for cruise, (T_c)	%	39.34	42.5	8.03		

same with EVIDC, 31.74% lower than Xi'an and 42.74% lower than Flr DC. Similarly, proportion of time spent in cruise mode for ERDC was 22.3% higher than EVIDC, 60% higher than Xi'an and 82% higher than Flr DC. Percentage time spent in acceleration, deceleration, idle and cruise helps in understanding the driver behavior and associated road and traffic condition. The higher proportion of acceleration/deceleration time and lower proportion of cruising time implies frequent vehicle-to-vehicle interaction resulting in higher energy consumption.

7. Conclusions

Local driving cycles are important and are required to describe the actual driving patterns of a region of interest. Moreover, with possible increase in the penetration of e-vehicles it is important to analyze and model the driving cycles for such vehicles. In this study, a driving cycle for e-rickshaw is developed for urban and rural traffic conditions using real world driving data. Two approaches, namely random selection technique and k-means clustering algorithm were used to arrive at best representative driving cycle using micro-trips technique. Both the approaches gave similar results with k-means clustering producing a slightly more representative driving cycle. The developed driving cycle was shown to consist of similar proportions of driving conditions in terms of road types and traffic conditions to those observed in real-world operating conditions. The driving cycle developed for urban and rural driving conditions differ significantly from urban driving cycles commonly used in India thus proving the necessity of city/location specific driving cycles. The driving cycle developed in this study can be used to compute energy consumption by e-rickshaw. In addition, the insights from the study can be used for electricity grid analysis and economic and lifecycle analysis of e-rickshaws.

Table 6
Comparison of ERDC with other typical driving cycles.

Driving cycle	ERDC	EVIDC	Xi'an	Flr .DC
Country	India	India	China	USA
Road type	Composite	Urban	Urban	Composite
Vehicle type	*E*-rickshaw	E-rickshaw	EV	EV
Average speed (kmph)	17.4	50	20.74	27.0
Average acceleration (m/s^2)	1.14	0.204	0.78	0.60
Average deceleration (m/s^2)	1.16	0.206	0.77	0.64
Time proportion of driving mode for acceleration (%)	22.43	25.2	34.5	40.6
Time proportion of driving modes for deceleration (%)	21.92	22.0	32.1	38.3
Time proportion of driving modes for idle (%)	16.31	21.5	17.5	13.7
Time proportion of driving modes for cruise (%)	39.34	30.57	15.9	7.4

7.1. Limitations and future scope

This study has a few limitations: (a) The current study considered limited number of drivers; future studies may investigate the accuracy of representative driving cycle by considering additional number of drivers. (b) The driving style (normal/aggressive/conservative) of the drivers was not considered in the present study. However, driving style may influence the developed driving cycle and therefore future studies may evaluate and consider driving behavior in developing driving cycles. (c) The driving cycle developed in this study accounts for urban and rural traffic setting, and therefore it may not be transferable to other regions such as city traffic. Future studies can use representative driving cycles in chassis dynamometer simulation to determine the power and energy consumption for a particular city/region to

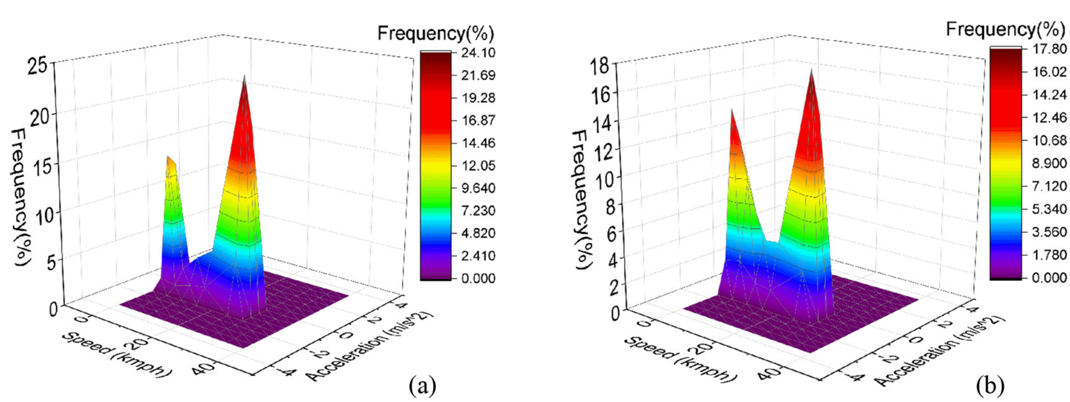

Fig. 11. (a) Speed acceleration frequency distribution plot for entire data (b) Speed acceleration frequency distribution plot for ERDC using k-means clustering technique.

understand the energy consumption, driving range prediction, and charging demand prediction.

Declaration of Competing Interest

None.

Acknowledgment

This work was supported by the project "M2Smart: Smart Cities for Emerging Countries based on Sensing Network and Big Data Analysis of Multimodal Regional Transport System", JST/JICA SATREPS, Japan.

References

[1] NSSO, Motor Vehicles – Statistical Year Book India 2016, Ministry of Statistics and Programme Implementation, 2016.
[2] MOSPI, Press Note on Household Expenditure on Services and Durable Goods, Ministry of Statistics and Programme Implementation, 2016.
[3] MORTH, Road Accidents in India 2020, Transport Research Wing of Ministry of Road Transport and Highways of India, 2020.
[4] RTA, Growth of Vehicles, Government of Telangana State Transport Department, 2020.
[5] T.P. Brief, Estimating Vehicular Emissions from auto Rickshaws Plying in Bengaluru City, IJSER, 2018.
[6] G. Hill, O. Heidrich, F. Creutzig, P. Blythe, The role of electric vehicles in near-term mitigation pathways and achieving the UK's carbon budget, Appl. Energy 251 (2019) 113111.
[7] Bloomberg, Electric Rickshaw Taking off in India, Oct 26, 2018.
[8] Rahul Mishra, India overtakes China with E-rickshaw revolution, Economic Times, Oct 27, 2018.
[9] D. Majumdar, T. Jash, Merits and challenges of e-rickshaw as an alternative form of public road transport system: a case study in the state of West Bengal in India, Energy Procedia 79 (2015) 307–314.
[10] X. Yuan, C. Zhang, G. Hong, X. Huang, L. Li, Method for evaluating the real-world driving energy consumptions of electric vehicles, Energy 141 (2017) 1955–1968.
[11] H.Y. Tong, W.T. Hung, A framework for developing driving cycles with on-road driving data, Transp. Rev. 30 (5) (2010) 589–615.
[12] Q. Wang, H. Huo, K. He, Z. Yao, Q. Zhang, Characterization of vehicle driving patterns and development of driving cycles in Chinese cities, Transp. Res. Part D: Transp. Environ. 13 (5) (2008) 289–297.
[13] K.J. Kelly, B.K. Bailey, T.C. Coburn, W. Clark, L. Eudy, P. Lissiuk, FTP emissions test results from flexible-fuel methanol dodge spirits and ford econoline vans, SAE Trans. (1996) 656–680.
[14] A.C. Mersky, C. Samaras, Fuel economy testing of autonomous vehicles, Trans. Res. Part C: Emerg. Technol. 65 (2016) 31–48.
[15] J.H. Tsai, H.L. Chiang, Y.C. Hsu, B.J. Peng, R.F. Hung, Development of a local real world driving cycle for motorcycles for emission factor measurements, Atmos. Environ. 39 (35) (2005) 6631–6641.
[16] J.H. Kent, G.H. Allen, G. Rule, A driving cycle for Sydney, Transp. Res. 12 (3) (1978) 147–152.
[17] S.H. Ho, Y.D. Wong, V.W.C. Chang, Developing Singapore driving cycle for passenger cars to estimate fuel consumption and vehicular emissions, Atmos. Environ. 97 (2014) 353–362.
[18] U. Galgamuwa, L. Perera, S. Bandara, Developing a general methodology for driving cycle construction: comparison of various established driving cycles in the world to propose a general approach, J. Transp. Technol. 5 (04) (2015) 191.
[19] K.S. Nesamani, K.P. Subramanian, Development of a driving cycle for intra-city buses in Chennai, India, Atmos. Environ. 45 (31) (2011) 5469–5476.
[20] S.H. Kamble, T.V. Mathew, G.K. Sharma, Development of real-world driving cycle: case study of Pune, India, Transp. Res. Part D: Transp. Environ. 14 (2) (2009) 132–140.
[21] N.H. Arun, S. Mahesh, G. Ramadurai, S.S. Nagendra, Development of driving cycles for passenger cars and motorcycles in Chennai, India, Sustain. Cities Soc. 32 (2017) 508–512.
[22] J. Lin, D.A. Niemeier, An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle, Atmos. Environ. 36 (38) (2002) 5759–5770.
[23] X. Liu, J. Ma, X. Zhao, J. Du, Y. Xiong, Study on driving cycle synthesis method for city buses considering random passenger load, J. Adv. Transp. (2020) 2020.
[24] M. Zhang, S. Shi, N. Lin, B. Yue, High-efficiency driving cycle generation using a Markov chain evolution algorithm, IEEE Trans. Veh. Technol. 68 (2) (2018) 1288–1301.
[25] Z. Wang, J. Zhang, P. Liu, C. Qu, X. Li, Driving cycle construction for electric vehicles based on Markov chain and Monte Carlo method: a case study in Beijing, Energy Procedia 158 (2019) 2494–2499.
[26] A. Fotouhi, M.J.S.I. Montazeri-Gh, Tehran driving cycle development using the k-means clustering method, Sci. Iran. 20 (2) (2013) 286–293.
[27] H.S. Kim, K.K. Jeon, S.J. Choi, A study on city driving cycle for performance evaluation of electric comer module of compact EV, KSAE Annual Conference Proceedings 2012, pp. 2305–2309.
[28] L. Berzi, M. Delogu, M. Pierini, Development of driving cycles for electric vehicles in the context of the city of Florence, Transp. Res. Part D: Transp. Environ. 47 (2016) 299–322.
[29] J. Brady, M. O'Mahony, Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas, Appl. Energy 177 (2016) 165–178.
[30] W. Zhou, K. Xu, Y. Yang, J. Lu, Driving cycle development for electric vehicle application using principal component analysis and K-means cluster: with the case of Shenyang, China, Energy Procedia 105 (2017) 2831–2836.
[31] S. Kumar Pathak, Y. Singh, V. Sood, S.A. Channiwala, Drive Cycle Development for Electrical Three Wheelers (No. 2017-01-1593), SAE Technical Paper, 2017.
[32] H.C. Watson, E.E. Milkins, P.A. Holyoake, E.T. Khatib, S. Kumar, Modelling emissions from cars, Australasian Road Research Board (ARRB), The 10 th Australian Transport Research Forum, Vol. 1, 1985, pp. 88–109.
[33] C. Lin, L. Zhao, X. Cheng, W. Wang, A DCT-based driving cycle generation method and its application for electric vehicles, Math. Probl. Eng. (2015) 2015.
[34] S. Tamsanya, S. Chungpaibulpatana, B. Limmeechokchai, Development of a driving cycle for the measurement of fuel consumption and exhaust emissions of automobiles in Bangkok during peak periods, Int. J. Automot. Technol. 10 (2) (2009) 251–264.
[35] J.I. Huertas, M. Giraldo, L.F. Quirama, J. Díaz, Driving cycles based on fuel consumption, Energies 11 (11) (2018) 3064.
[36] E. Ericsson, Independent driving pattern factors and their influence on fuel-use and exhaust emission factors, Transp. Res. Part D: Transp. Environ. 6 (5) (2001) 325–345.
[37] G. Amirjamshidi, M.J. Roorda, Development of simulated driving cycles for light, medium, and heavy duty trucks: case of the Toronto waterfront area, Transp. Res. Part D: Transp. Environ. 34 (2015) 255–266.
[38] L. Della Ragione, G. Meccariello, GPS signal correction to improve vehicle location and related emission evaluation, J. Stat. Sci. Appl. 3 (3–4) (2015) 50–61.
[39] E.D. Jackson, L. Aultman-Hall, Accuracy of GPS-Based Acceleration for Vehicle Emissions Modeling (No. 07-2202), 2007.
[40] J. Jun, R. Guensler, J. Ogle, Smoothing methods designed to minimize the impact of GPS random error on travel distance, speed, and acceleration profile estimates, Transp. Res. Rec. 1972 (1) (2005) 141–150.
[41] X. Zhao, Q. Yu, J. Ma, Y. Wu, M. Yu, Y. Ye, Development of a representative EV urban driving cycle based on a k-means and SVM hybrid clustering algorithm, J. Adv. Transp. (2018) 2018.
[42] S. Buggaveeti, Dynamic Modeling and Parameter Identification of a Plug-in Hybrid Electric Vehicle, Master's thesis, University of Waterloo, 2017.