A Bottom-Up Clustering Approach to Identify Bus Driving Patterns and to Develop Bus Driving Cycles for Hong Kong

Hing Yan TONG* and Ka Wai NG

College of Professional and Continuing Education,

The Hong Kong Polytechnic University,

8 Hung Lok Road, Hung Hom, Kowloon, Hong Kong SAR

ccftong@cpce-polyu.edu.hk

Tel: (852) 3746 0486 Fax: 852 2625 4259

Abstract

Bus transport has been an important mode taken up a significant share of urban travel demand and thus the corresponding impacts on the environment are of great concerns. Use of driving cycles to evaluate the environmental impacts of buses has attracted much attention in recent years worldwide. The franchised bus service is currently playing important roles in the public transport system in Hong Kong, however there is no driving cycle developed specifically for them. A set of bus driving cycle was therefore developed using a bottom-up approach where driving data on the bus network with mixed characteristics were collected. Using the Ward's method for clustering, the collected data were then categorised into three clusters representing distinct franchised bus route patterns in Hong Kong. Driving cycles were then developed for each route pattern including (i) congested urban routes with closely spaced bus stops and traffic junctions; (ii) inter-district routes containing a number of stop-and-go activities and a significant

portion of smoother high speed driving; and (iii) early morning express routes and mid-night routes connecting remote residential areas and urban areas. These cycles highlighted the unique low speed and aggressive driving characteristics of bus transport in Hong Kong with frequent stop-and-go activities. The findings from this study would definitely be helpful in assessing the exhaust emissions, fuel consumptions as well as energy consumptions of bus transport. The bottom-up clustering approach adopted in this study would also be useful in identifying specific driving patterns based on vehicle speed trip data with mixed driving characteristics. It is believed that this approach is especially suitable for assessing fixed route public transport modes with mixed driving characteristics.

Keywords: Bus Driving Patterns, Driving Cycles, Vehicle Specific Power (VSP), GPS Data Collection, Cluster Analysis, Vehicle Emissions and Energy Consumption

Declarations

- **Ethics approval and consent to participate** – Not Applicable
- **Consent for publication** – Not Applicable
- **Availability of data and materials** – The datasets generated during the current study are not publicly available but are available from the corresponding author on reasonable request.
- **Competing interests** – The authors declare that they have no competing interests
- **Funding** – The work described in this paper was fully supported by a grant from the College of Professional and Continuing Education, an affiliate of The Hong Kong Polytechnic University

- **Authors' contributions** – **TONG Hing yan** was responsible for the conceptualization, methodology, data collection and analysis, further investigation as well as writing, reviewing and final editing of the manuscript. **NG Ka wai** performed data analysis as well as writing, reviewing and editing of the manuscript. Both authors read and approved the final manuscript.

1. Introduction

Research on the development and applications of driving cycles has been increasingly getting more attention in recent years. A number of large scale studies on driving cycles have been undergoing in different part of the world in the past decade such as the Worldwide Harmonised Light Duty Driving Test Cycles (WLTC) for light duty vehicles (Tutuian, et al., 2014) as well as in different cities in Europe, Brazil, Thailand, India and different parts of the mainland China. These driving cycles have also been developed for different uses ranging from traditional purposes such as vehicle emissions and fuel consumption estimation, vehicle certification to accident analysis (Mongkonlerdmanee, and Koetniyom, 2019) and electric vehicle performance evaluations (Feroldi, and Carignano, 2016; Gong, et al., 2018; Zhang, et al., 2017). In order to achieve these objectives, individual studies have been focusing on collecting data for a specific vehicle type or a specific driving condition (Han, 2012, Seers, et al. 2014; Zamboni, et al., 2015). For these types of studies, test routes would be selected based on specific criteria so as to achieve the objectives of individual studies. Data analysis would be relatively straight forward as the collected data would generally be able to reflect the driving characteristics of the specific driving conditions being investigated. However, for studies involving more complicated driving conditions, the collected data may exhibit mixed driving characteristics. Classification techniques such as Principal Component Analysis (PCA) and k-means clustering analysis might therefore be necessary. In the literature, these techniques were mainly employed for classifying the micro-trips at the driving cycle constructions stage. This approach is useful to analyse data collected from specific test route(s). For driving cycles developed for vehicles with fixed route or network (e.g. public transport modes), each route would have different structures and specific operational considerations. This approach of data analysis (i.e. using clustering techniques to analyse speed data with mixed driving characteristics) may not be appropriate as it completely ignored the structure of individual routes (e.g. express routes with limited stops, routes running along different districts (or across different districts) with different characteristics during different periods of a day).

In Hong Kong, driving cycles developed were all for light duty vehicles (Tong, et al., 1999; Hung, et al., 2007), so the purpose of this paper is to develop a set of driving cycles for the franchised bus transport sector in Hong Kong. Since the Hong Kong bus transport system is responsible for 31.5% (i.e. about 4.05 millions) of the total daily passenger trips in 2018 (TD, 2019), the bus network is very complicated and comprehensive. Therefore, the previous approaches of selecting just a single or a few test routes to collect speed data would probably be insufficient to reflect the overall driving conditions of the whole network. As such, this study would adopt another approach to collect bus speed data which is capable of covering a wide range of bus driving conditions. A bottom-up approach of speed data analysis would then be devised in identifying and defining the bus driving patterns in Hong Kong. This method would make use of statistical classification techniques to classify the trip data and then develop the corresponding bus driving cycles.

The remainder of this paper is organised as follows. Section 2 starts by reviewing relevant driving cycle studies, and then based on it to introduce the proposed "bottom-up" approach for driving cycle development for a complicated bus route network. In particular, the proposed "bottom-up" approach is compared with the traditional "top-down" approach with the major differences highlighted. In Section 3, details of the franchised bus route network are described. Section 4 then depicts the data collection campaign which is capable of collecting bus speed data covering a wide range of bus routes. In Section 5, a clustering method is proposed to identify the bus route driving patterns using the collected bus speed data. According to the identified clusters, corresponding bus driving cycles is then synthesized in Section 6. The synthesized bus driving cycles in this study are then compared with other international bus cycles in Section 7 and eventually come up with a conclusion in Section 8 to highlight the unique bus driving characteristics in Hong Kong.

2. Literature Review and the Proposed Methodology

As mentioned earlier, one of the largest driving cycle studies worldwide was the WLTC in 2012 with an aim to represent typical driving characteristics round the world for light duty vehicles (Tutuian, et al., 2014). This project collected data from 14 countries in five different regions namely USA, ~~UE~~ EU and Switzerland, Korea, Japan and India covering a wide range of traditional vehicle categories under different driving conditions (e.g. urban vs highways, peak vs off-peaks, etc.). Eventually, separate sets of driving cycles were developed for 3 vehicle categories according to the vehicle's power mass ratio and maximum speed for type-approval applications. In different parts of the mainland China, different driving cycles have also been developed for different cities (including Beijing, Fuzhou, Hefei, Jinan, Nanjing, Shanghai, Tianjing, Xi'an, etc.) for different vehicles types (e.g. passenger cars, electric vehicles, traditional buses, BRT buses, etc.). The major purposes were to understand local driving characteristics and to develop corresponding driving cycles for local uses. In India, there were also a series of studies in Bangalore (Mayakuntla and Verma, 2018), Chennai (Arun, 2017; Desineedi et al., 2020), Delhi (Kumar, et al., 2013), Dhanbad (Adak, et al., 2016) and New Delhi (Maurya and Bokare, 2012) to develop local driving cycles for different purposes. There were also some other driving cycles developed in other locations for different uses including traditional purposes such as vehicle emissions and fuel consumption estimation, vehicle certification, accident analysis (Mongkonlerdmanee, and Koetniyom, 2019) as well as electric vehicle performance evaluations (Feroldi, and Carignano, 2016; Gong, et al., 2018; Zhang, et al., 2017; Zhao, et al., 2020).

Based on the purposes of developing the driving cycles, individual studies would then be focusing on collecting data for a specific vehicle type or a specific driving condition, normally known as stratification (Tong and Hung, 2010a; 2010b). Test route(s) would then be selected according to the stratification for data collection. For example, Han (2012) developed a driving cycle for the military areas in Korea. Seers et al. (2014) and Zamboni et al. (2015) produced driving cycles for a utility vehicle at an airport in Montreal and for container vehicles at port areas respectively. For these types of studies, test routes would be selected based on specific criteria so as to achieve the objectives of individual studies (e.g. for urban driving during peak periods, for highway driving, etc.). For example, Maurya and Bokare (2012) collected data for a specific bus route travelling through a specific highway in Delhi. Analysis of the collected speed data would then be relatively straight forward as the data collected should be representative of the corresponding driving conditions. The resultant driving cycles developed would also be able to reflect the driving characteristics of the specific driving conditions being investigated.

However, for studies with more complicated driving conditions under multiple factors (e.g. peak versus non-peak, weekdays versus weekends, urban versus highway, etc.), the collected driving data would definitely exhibit mixed driving characteristics which is not easily separable or identifiable. Therefore, composite driving cycles would normally be developed to represent the overall driving characteristics of the region or city concerned. Should driving cycles for specific driving conditions be developed, classification techniques would be necessary to clearly classify the collected data and thus to define specific driving conditions based on their statistical properties. In the literature, statistical classification techniques have been commonly adopted in the development of driving cycles (Abas, et al., 2018; Berzi, et al., 2016; Desineedi, et al., 2020; Fotouhi and Montazeri, 2013; Jing, et al., 2017; Li, et al., 2016; Ma, et al., 2019; Peng, et al., 2019; Shen, et al., 2018; Tong and Hung, 2008, 2010a, 2010b; Yan, et al., 2020; Yang, et al., 2019; Zhang, et al., 2017; Zhao, et al., 2018). However, these techniques are mainly employed for classifying the micro-trips into different categories for candidate driving cycle constructions. None of them has used them to classify the originally collected trip data and to define the corresponding driving conditions. For example, Yang, et al., (2019) used PCA and cluster analysis methods to classify the micro-trips into three clusters representing three different driving conditions. Representative micro-trips were then selected to construct candidate driving cycles to determine the final cycle. Peng et lal., (2019) classified the micro-trips into three clusters based on PCA and k-means cluster analysis. The results were then incorporated into a Markov model to develop sub-cycles for each cluster and then combined together to become the final cycle. Desineedi, et al., (2020) categorised micro-trips into 6 clusters using k-means clustering and then applied to a one-chain Markov modelling process to develop driving cycles for peaks and off-peak periods. Yan, et al., (2020) developed a k-MPSO (modified particle swarm optimization) clustering algorithm to cluster the micro-trips and then constructed a typical driving cycle. This analysis approach decomposes the whole trip into small segments (i.e. the micro-trips) and then categorises those with common characteristics. It is useful to analyse the driving characteristics for speed data collected from specifically selected test route(s) as the structure of the test routes are intentionally selected to cover driving conditions concerned. For driving cycles developed for vehicles with fixed route or network (e.g. public transport modes such as buses and minibuses), each route would have different structures and is designed for different operational purpose in fulfilling the passenger travel demand but not for speed data collection. Therefore, some routes might only travel within the urban areas while some others might involve cross districts travel via highways. This approach of data analysis (i.e. clustering micro-trips) by decomposing the original structure of the routes may not be appropriate as it completely ignored the structure of individual routes.

In the literatures, there are relatively less studies concerning driving cycles for buses or bus routes. Among these studies, bus speed data were collected for a single bus route or a small number of selected bus routes. Studies with data collected for only a single bus route (Gunther, et al., 2017; Kivekas, et al., 2018a; 2018b; Kumar, et al., 2013; Li, et al., 2016; Liu, et al., 2020; Maurya and Bokare, 2012; Mongkonlerdmanee and Koetniyom, 2019; Nguyen, et al., 2016; Shen, et al., 2018) usually aimed at investigating the specific driving conditions along those particular test routes.

Therefore, use of just one route to collect data was deemed satisfactory. For studies that intended to examine the composite driving characteristics of the whole bus network, more test routes were employed for on-road data collection (Lai, et al., 2013; Nesamania and Subramanian, 2011; Nguyen, et al., 2018; Peng, et al., 2019; Quirama, et al., 2020;). The selection of test routes in these studies was based on the researchers' local understanding of the driving conditions and structures of the bus routes. This approach might be appropriate for cities or regions with a relatively simple bus network. For a city like Hong Kong, with a very complicated bus network, selection of just a few bus routes might not be able to cover different driving characteristics and structures of the bus route network. Data collection has to be conducted on a much wider coverage of bus routes. Therefore, statistical classification techniques are necessary to identify the driving patterns concerned and thus to develop the corresponding driving cycles.

To sum up the review, the traditional approach of driving cycle development can be viewed as a "Top-Down" approach as the whole process starts from pre-defined stratification criteria (Tong and Hung, 2010b). Test routes selection and data collection are then based on the stratification. Speed data analysis and type of cycles to be developed are also according to the stratification. Clustering technique may be used in classifying the micro-trips or trip segments during the cycle construction stage. However, this study has adopted a "Bottom-Up" approach in which data containing mixed driving characteristics are to be collected without prior stratification criteria.

The method adopted to collect bus speed data is capable of covering a wide range of bus routes across the whole territory of Hong Kong as well as their corresponding driving conditions during different time periods of a day on both weekdays and weekends. Undoubtedly, the data collected using this approach would exhibit mixed driving characteristics that may not be easily identifiable through simple and direct groupings. Specific bus route driving patterns and types of driving cycles to be developed would be determined using cluster analysis and then be defined on the basis of a list of quantitative assessment parameters. A comparison of the two approaches is shown in Figure 1 and Table 1. Therefore, the main objectives of this paper are: (1) to use statistical classification techniques to classify the collected trip data (instead of classifying micro-trips only) into clusters exhibiting similar statistical driving properties; (2) to define different driving patterns of franchised bus services in Hong Kong on the basis of the identified clusters; and (3) to develop a set of bus driving cycles representative of and corresponding to the identified bus driving patterns (i.e. the identified clusters).

[Figure 1]

[Table 1]

3. Bus Network in Hong Kong

Hong Kong is a city having the largest citywide public transport modal share around the world. Everyday, over 12 million passenger trips are made by different transport modes in Hong Kong which exceeds 90% of the total daily passenger trips (THB, 2017). Apart from the backbone railway system, the franchised bus transport system plays a very important role in enabling the extremely high public transport modal share in Hong Kong. Of all the daily public transport passenger trips, 31.5% (i.e. about 4.05 millions) are ~~taken up~~ made by the franchised bus transport system (TD, 2019). There are 5 franchised bus operators in Hong Kong operating 6 franchises regulated by the Transport Bureau. The franchised bus network contains over 600 routes and more than 6000 buses in which a majority of them are double deckers (TD, 2019). Detailed information of the franchised bus route network (such as the distribution of different route types across different bus operators and their geographical coverage) is also shown in Table 2. Kowloon Motor Bus (1933) Co. Ltd. (KMB), as the operator with the longest history in providing public bus services in Hong Kong, is mainly operating routes in Kowloon Peninsula (Kln) and New Territories (NT) while City Bus (CTB) and New World First Bus (NWFB) focus more on developing their bus networks on Hong Kong Island (HKI). For cross-harbour routes connecting Kln and HKI, KMB would collaborate with either CTB or NWFB to provide crossharbour bus services. On the other hand, New Lantau Bus Co. (1973) Ltd. (NLB), Long Win Bus Company (LW) and the second franchise of CTB are operating Tung Chung and airport routes. Tung Chung is a newly developed district on the Lantau Island. Theise information shows that the franchised bus network in Hong Kong is very comprehensive and complicated. To develop a set of driving cycles capable of reflecting the characteristics of the whole bus network requires data covering different route structures.

[Table 2]

4. Data Collection

Speed data in this study were collected using the GPS positioning capability of smartphones. The accuracy and sensitivity of the GPS device within smartphones have been well researched and have adopted in many transportation applications to track the vehicle dynamics. In this study, an iPhone 6 was the major mobile device used for the surveyors to conduct on-board bus speed data collection. Previous assessment of different versions of iPhones indicated that the relative spatial accuracy of its internal GPS devices was up to 99% (Garnett and Stewart 2015; Menard, et al., 2011) and was identified as reliable when compared to other more expensive vehicle tracking devices (Menard and Miller, 2010, 2011). In another study, the GPS sensor of iPhones also performed satisfactorily for transport navigation and safety applications (Gikas and Perakis, 2016). These results showed that the positioning capability of iPhones were very sensitive, reliable and suitable for collecting vehicle movement data.

Voluntary surveyors were employed to use their own smartphones with a data collection APP "MyTrack" to collect bus speed data during their daily commuting activities for a period of 4 months in 2017. The surveyors were carefully selected with an aim to cover travel patterns at and across different districts in Hong Kong. The plan was to let the surveyors did their daily travelling activities as usual so as to capture different activity patterns across different time periods on weekdays and weekends. During the data collection process, the surveyors were only required to turn on the APP while they were travelling on buses. Bus position and speed data would then be collected at one second intervals. Time for every stop and start activity was also recorded by the surveyors for cross-checking purpose in confirming the exact period of idling in the trip data. A total of 91 bus trips data were finally collected covering more than 30 bus routes across the whole territory of Hong Kong (Table 3). Among all the data collected 58% and 42% of the trips were for weekdays and weekends respectively. At the same time, around 8%, 12%, 43% and 37% of the trips were collected during the AM Peak, PM Peak, Inter-Peaks, and After Peaks periods respectively. It is understood that this dataset did not provide a complete coverage of the franchised bus routes in Hong Kong, however, the above statistics showed that the collected dataset did cover different time periods of a day, different days of a week (i.e. weekdays vs weekends) and more importantly across different districts (Table 1). It represents an improvement over similar bus driving cycle studies in the literature.

[Table 3]

As part of this large scale bus driving cycle development study, a driving cycle for a supercapacitor bus route in Hong Kong was developed using a separate set of 16 bus trip data obtained using the same data collection method. More details of the data collection process and the validity of this data collection approach can be found from a separate paper (Tong, 2019). It is believed that this approach of data collection could guarantee a good coverage of various kinds of bus routes. A wide range of mixed driving characteristics should have been captured in this dataset. Direct groupings of the collected bus trip data could be done based on the background information of the collected bus trips, however, use of statistical classification technique could also be another quantitative approach to identify the special driving patterns concerned.

To make sure that the speed dataset are ready for detailed analysis, the collected trip speed-time profiles were first screened for any abnormality such as sudden high or low speeds, occasional short time gaps, exceptional accelerations / decelerations exceeding the physical limits of buses, as well as unreasonably long zero speed period (through the double-checking with the stop and go records of by the surveyors on-board of the bus). The exact procedures of the data cleaning process are not going to be described here but further details can be found in a separate paper (Tong, 2019).

5. Identification of Bus Driving Patterns

After the data set were cleaned up, summary statistics were then derived for each of the 91 trip data. The set of commonly used assessment parameters are summarised in Table 4. It is important to note that this list of assessment parameters is consistent with those parameters being used elsewhere to analyse vehicle driving characteristics.

[Table 4]

As mentioned earlier, the collected dataset should contains mixed characteristics for a wide range of bus routes structures in terms of different time periods of a day, different days of a week as well as districts covered. Therefore, a cluster analysis technique was employed to classify the 91 trip datasets into clusters based on the 13 assessment parameters. Cluster analysis is a statistical technique to group a pool of subjects into significantly different groups on the basis of a list of variables. The groupings are built to be as statistically different as possible between groups, and as statistically homogeneous as possible within a group. The method begins with an $N \times k$ database. Then, an $N \times N$ matrix is generated using the k variables to indicate the level of similarity (or dissimilarity) of every group to every other group. There are a number of measures of similarity or difference, such as squared Euclidean distance, Euclidean distance, and cosine of vector variables. The subjects are then sorted into significantly different groups by one of the several clustering methods available so that (1) within the group subjects are as homogeneous as possible, and (2) across the groups are as different as possible. It is important to note that different clustering methods can come up with different cluster solutions. There is a number of clustering methods available depending upon two factors: (1) the metrics used to measure the similarity or distance between subjects; and (2) the clustering algorithm used. The method employed in this study is called the Ward's Method, which is an agglomeration method that combines two clusters at each stage until all subjects are finally combined into clusters. The procedures adopted to classify the 91 cleaned trip speed time data are summarised as follows:

<u>Step 1</u> Normalise the 13 assessment parameters for each trip data set.

<u>Step 2</u> Construct a dissimilarity matrix indicating the differences among all pairs of subjects using Squared Euclidean Distances of the 13 assessment parameters:

$$\text{Squared Euclidean Distance} = \sum_{k=1}^{n} (x_{ik} - x_{jk})^2$$

where: x_{ik} = value (normalised) of parameter k of subject i

x_{jk} = value (normalised) of parameter k of subject j

n = number of parameters

Step 3 Calculate the agglomeration schedule by the Ward's Method. The Ward's Method forms cluster by selecting the subject which minimises the within cluster sum of squares (i.e. Sum Squares Error, SSE).

$$SSE = \sum_{k=1}^{n} (x_{ik} - \bar{x}_k)^2$$

where: x_{ik} = value (normalised) of parameter k of subject i

$\bar{x}_k$ = mean value (normalised) of parameter k

n = number of parameters

It includes the following procedures:

(1) Starts with each subject in its own cluster (i.e. N subjects will start with N clusters);

(2) Next, locates the two subjects that are closest to each other and a cluster is built with two subjects (i.e. the N subjects now become N-1 clusters, one with two subjects, and N-2 with one subject each);

(3) Then, locates the next two closest subjects and a two-subject cluster is built. (i.e. the N-1 clusters now become N-2 clusters, two with two subjects each, and N-4 with one subject each);

(4) As Ward's algorithm progresses, it will start to combine a single subject with a pre-existing cluster or to combine one pre-existing cluster with another. This process is continued until all N subjects are eventually combined into one cluster.

As a result of the cluster analysis, the trips were classified into 3 distinct clusters, in which 54%, 20% and 26% of the trips were fallen into Clusters 1 to 3 respectively. The detailed clustering results are shown in Tables 5 to 8. The summary statistics in Table 8 indicated obvious differences in the driving patterns represented by these three clusters. Cluster 1 showed the lowest average speed and proportion of cruising, the highest proportion of idling, and the shortest average micro-trip length. Among the three clusters, it also had the highest values for all the acceleration related parameters (including average acceleration and deceleration rates, RMS and PKE) and the highest proportion of creeping. These implied that this driving pattern was characterised by the slowest speed and the most aggressive driving with frequent stop-and-go activities. Therefore, Cluster 1 resembled typical congested driving conditions at the urban areas. Bus speeds were constrained by the closely spaced bus stops and traffic junctions as well as the congested traffic. From the perspective of the bus services, Cluster 1 should represent routes running within one district during peak and congested periods. Tables 6 to 7 showed that Cluster 1 comprised of nearly all the trips collected during peak periods and routes running within only one district. This further ~~affirms~~ affirmed the characteristics of urban congested driving conditions during the peak periods.

For Cluster 2, all the 13 assessment parameters were located at the median position among the three clusters. It implied that Cluster 2 represented the conditions between very congested driving and relatively smooth highway driving. This condition was signified by some stop-andgo activities due to the bus stops at individual districts but the traffic conditions were not as congested as in the urban areas. These characteristics were similar to the structure of interdistrict bus routes which was composed of a section with multiple short micro-trips (i.e. representing typical urban driving with frequent stop-and-go due to closely spaced bus stops and traffic junctions at individual district) and another section containing very long high speed driving (i.e. representing driving along high speed road sections with just a few stops).

Cluster 3 was the fastest and least aggressive among the three identified clusters. It was characterised by the highest proportion of cruising, the smallest proportion of idling (only about 5.7%) and all other acceleration parameters, as well as the longest mean micro-trip length. These implied that this driving pattern was much smoother and relatively stable. Thus, Cluster 3 represented bus routes with significant portions of travelling on smoother (i.e. possibly highway sections) and/or uncongested road sections (i.e. possibly during early morning or late night traffic conditions) with fewer stops (including stops due to bus stops, traffic junctions and congestions). These characteristics were similar to the structure of two special bus route types in Hong Kong. The first type is special express routes operating only a few trips during early morning and they are mainly taken by home-to-work (or home-to-school) travellers from remote residential areas to the urban areas. Characteristics of express routes are having only a few stops at the residential areas and the urban areas to pick-up and drop-off passengers respectively. In between the two areas, normally there would be a significant portion of relatively smooth and high speed travelling with limited stops (normally a highway section). The second type is midnight bus routes running after regular bus service hours. This type of route normally connects the urban areas with the residential areas as well, but under much smoother driving conditions at mid-night. Tables 5 to 7 also indicated that Cluster 3 was constituted by significantly higher percentages of trips under smoother driving conditions (i.e. (1) nearly 60% of Cluster 3 trips were less congested weekend trips (Table 5); (2) had the highest proportion of night time after the peak trips (Table 6); and (3) more than half of the inter-district trips were classified under Cluster 3; (Table 7)).

For visual investigation of the clustering results, scatter plot between two well-known vehicle driving variables (i.e. the average speed and positive kinetic energy) of the trip data are shown in Figure 2. Trip average speed is an important variable for emission and fuel consumption estimations while PKE is related to acceleration and energy consumption. The scatter plot between these two variables can help identifying the characteristics of specific driving patterns.

The scatter plot shows a reasonable inversely proportional pattern between these two variables. Higher trip average speed implies smoother driving and thus smaller PKE. Distinct groups of trip data are clearly shown as three clusters on the scatter plot.

[Figure 2]

[Table 5]

[Table 6]

[Table 7]

[Table 8]

On the whole, three distinct driving patterns were identified representing three types of bus routes in Hong Kong and were summarised below. In the next section, corresponding driving cycles would be developed for each of these three patterns.

Cluster 1: Routes running within typical congested urban areas with closely spaced bus stops and traffic junctions.

Cluster 2: Inter-district routes comprising a number of stop-and-go activities at individual district and a significant portion of smoother high speed driving.

Cluster 3: Early morning express routes and mid-night routes connecting remote residential areas and urban areas.

6. Development of Bus Driving Cycles in Hong Kong

For each of the three bus driving patterns represented by the three clusters, separate driving cycles would be developed (Figure 3). For Clusters 1 and 2, the trip data contains considerable amount of stop-and-go driving as well as a series of micro-trips. Therefore, the following microtrip random selection method would be directly applied to generate driving cycles for these two clusters.

[Figure 3]

To begin with, the mean values of the 13 assessment parameters for each cluster were set as target statistics. Candidate driving cycles were constructed by first identifying the micro-trips which ~~are~~ were defined as speed profiles bounded by two consecutive idling periods. The microtrips together with the preceding idling portion ~~came immediately before it~~ were then selected at random to constitute a driving cycle until the required cycle length was achieved. The number of micro-trips constituting the candidate cycle depended on different driving patterns. The assessment parameters of the candidate cycle were then matched with the target statistics. Cycles with all the assessment parameters within 5% of the target statistics became an acceptable cycle. If not, another candidate cycle would be constructed and matched in the same manner. Eventually, 10 acceptable cycles would be developed for the selecting the best cycle. The acceptable cycles would then be ranked to determine the best cycle according to the average absolute percentage error (AAPE) across all the 13 assessment parameters. Speed acceleration probability distributions (SAPD) and vehicle specific power (VSP) distributions were also derived for the acceptable cycles and for the whole data set under each cluster. Comparisons between those distributions would also be referenced for determining the best cycle for each cluster. The formula used for calculating VSP was as follows (Chen, et al., 2019).

$$VSP = v\ (\ 1.1 \times a + 0.132\ ) + 0.0000745\ v^3$$

where v is the speed (in m/s) and a is the acceleration (m/s^2)

For Cluster 3, most of the trips comprised at least one relatively long micro-trip reflecting highway driving conditions. Therefore, the corresponding driving cycle would be constructed by selecting the single trip with values of the 13 assessment parameters closest to the target statistics. This approach was also consistent with some other studies (Achour and Olabi, 2016; Atiq et al., 2017; Knez et al., 2014; Kumar et al., 2013; Liu et al., 2016; Mansour et al., 2018; Zomboni et al., 2015).

The resultant driving cycles for each cluster are shown in Figure ~~3~~ 4 and the comparison of summary statistics between the synthesized cycle and the target statistics are also summarised in Table 9. Cycles for the three clusters were characterised by cycle durations of 1889, 1907 and 2074 seconds and average speeds of 13.8 km/h, 38.8 km/h and 56.8 km/h respectively. This was consistent with the characteristics of these three driving patterns. Figure 3 also showed that the Cluster 1 cycle consisted of a series of relatively low speed and short micro-trips separated by idling periods. For the Cluster 2 cycle, it was composed of high speed and long micro-trip together with a series of low speed and short micro-trips. These characteristics were consistent with inter-district routes described earlier. Cluster 3 contained two very long and high speed micro-trips reflecting the corresponding high speed and smooth driving conditions.

[Table 9]

[Figure 34]

The SAPD and VSP distributions derived for the synthesized cycles and the whole dataset are compared in Figures 4 5 to 56. The resolutions of the SAPDs are 5 km/h and 0.2 m/s^2. Both the SAPD and VSP distributions also presented very good agreements between the whole dataset and the synthesized cycle.

[Figure 45]

[Figure 56]

7. Comparison with Worldwide Bus Driving Cycles

The major parameters of the synthesized driving cycles and other worldwide bus driving cycles are summarised in Table 10. The comparisons between the synthesized cycles and the corresponding international cycles are basically consistent. Cluster 1 cycle can be matched with international cycles developed for urban areas while Cluster 2 cycle (i.e. inter-district routes) generally agrees with international cycles involving highways. When looking at the acceleration parameters, the driving characteristics for bus routes in Hong Kong are quite different from other bus driving cycles. Cycles developed in this study exhibit significantly higher acceleration and deceleration rates, as well as much shorter cruising periods than most of the other international bus cycles. These driving characteristics have highlighted a very unique bus driving pattern in Hong Kong.

[Table 10]

These results reflect some unique bus driving characteristics in Hong Kong. First, the extremely high demand of franchised bus services in Hong Kong significantly increases the dwell times at bus stops for picking up and dropping off passengers, and thus exhibit long idle proportions in the driving pattern. On the same token, bus routes are also designed to have closely spaced bus stops at the urban and residential areas. Together with the closely spaced junctions, there would be frequent stop-and-go driving. The identified patterns also indicate that bus drivers in Hong Kong are more aggressive (in terms of the relatively higher acceleration and deceleration rates) than elsewhere possibly because of the tight bus schedules, short bus stops spacing and frequent stop-and-go operations.

8. Conclusions

In this paper, a set of bus driving cycles representative of different bus route driving conditions in Hong Kong have been developed. On-road speed-time data were collected using a cost effective data collection approach utilising the GPS capability of smartphones. This approach is particularly useful for developing economies in which significant amount of bus speed data covering different types of bus routes and driving conditions could be collected within a relatively low budget. Cluster analysis was employed to classify the collected trip data into 3 distinct clusters reflecting the driving patterns of (i) congested urban bus routes with closely spaced bus stops and junctions; (ii) inter-district bus routes with a relatively smoother and high speed driving section; and (iii) early morning express routes and mid-night routes with much smoother traffic conditions. Separate synthesized driving cycles were then developed by using the micro-trips random selection approach to match the overall summary statistics as well as the SAPD and VSP distributions. The three developed driving cycles showed characteristics that matched well with the corresponding bus route patterns. The findings from this study would definitely be helpful in assessing the exhaust emissions, fuel consumptions as well as energy consumptions of bus transport. The bottom-up clustering approach adopted in this study would also be useful in identifying specific driving patterns based on vehicle speed trip data with mixed driving characteristics. It is believed that this approach is especially suitable for assessing fixed route public transport modes with mixed driving characteristics.

Acknowledgement

The work described in this paper was fully supported by a grant from the College of Professional and Continuing Education, an affiliate of The Hong Kong Polytechnic University.

References

Abas, M.A., Rajoo, S. and Abidin S.F.Z. (2018) Development of Malaysian Urban Drive Cycle using Vehicle and Engine Parameters. *Transportation Research D* 63, pp.388-403.

Achour, H. and Olabi, A.G. (2016) Driving Cycle Development and Their Impacts on Energy Consumption of Transportation. *Journal of Cleaner Production* 112, pp. 1778-1788.

Adak, P., Sahu, R. and Elumalai, S.P. (2016) Development of Emission Factors for Motorcycles and Shared Auto-Rickshaws Using Real-World Driving Cycle for a Typical Indian City. *Science of the Total Environment* 544, pp. 299-308.

Arun, N.H., Mahesh, S., Ramadurai, G. and Nagendra, S.M.S. (2017) Development of Driving Cycles for Passenger Cars and Motorcycles in Chennai, India. *Sustainable Cities and Society* 32, pp. 508-512.

Atiq, W.H., Rabiha, S.R., Norbakyah, J.S. and Salisa, A.R. (2017) Development and Characterisation of UMT PHEM Driving Cycle. *APRN Journal of Engineering and Applied Sciences* 12, pp. 1926-1931.

Berzi, L., Delogu, M. and Pierini. M. (2016) Development of Driving Cycles for Electric Vehicles in the context of the City of Florence. *Transportation Research D* 47, pp. 299-322.

Chen, T., Li, M.X, Feng, H.J. Chen B. and Gao, Y. (2019) Statistical Vehicle Specific Power Profiling of Heavy-Duty Vehicles for Mountainous Highways. Green Intelligent Transportation Systems: Proceedings of the 8th International Conference on Green Intelligent Transportation Systems and Safety. P. 229-236.

Desineedi, R.M., Mahesh, S. and Ramadurai, G. (2020) Developing Driving Cycles using kmeans Clustering and Determining Their Optimal Duration. *Transportation Research Procedia* 48, pp.2083-2095.

Duran, A. and Earleywine, M. (2012) GPS Data Filtration Method for Drive Cycle Analysis Applications. *SEA International*.

Feroldi, D. and Carignano, M. (2016) Sizing for Fuel Cell / Supercapacitor Hybrid Vehicles Based on Stochastic Driving Cycles. *Applied Energy* 183, pp.645-658.

Fotouhi, A. and Montazeri, M. (2013) Tehran Driving Cycle Development using the k-means Clustering Method. *Scientia Iranica A* 20 (2), pp. 286-293.

Garnett, R. and Steward, R. (2015) Comparison of GPS Units and Mobile Apple GPS Capabilities in an Urban Landscape. *Cartography and Geographic Information Science*, 42, 1-8.

Gikas, V. and Perakis, H. (2016) Rigorous Performance Evaluation of Smartphone GNSS/IMU Sensors for ITS Applications. *Sensors*, 16, 1240. doi:10.3390/s16081240.

Gong, H.M., Zhou, Y., Yang, Q.K., Fan, J., Sun, F.G. and Goehlich, D. (2018) Generation of a Driving Cycle for Battery Electric Vehicles: A Case Study of Beijing. *Energy* 150, pp.901912.

Gunther, R., Wenzel, T., Wegner, M. and Rettig, R. (2017) Big Data Driven Dynamic Driving Cycle Development for busses in Urban Public Transportation. *Transportation Research D* 51, 276-289.

Han D.S., Choi, N.W., Cho, S.L., Yang, J.S., Kim, K.S., Yoo, W.S. and Jeon, C.H. (2012) Characterisation of Driving Patterns and Development of a Driving Cycle in a Military Area. *Transportation Research D* 17, pp. 519-524.

Hung, W.T., Tong, H.Y., Lee, C.P., Ha, K. and Pao, L.Y. (2007) Development of a practical driving cycle construction methodology: A case study in Hong Kong, *Transportation Research D*, 12, 115-128.

Jing, Z.C., Wang, G.L., Zhang, S.P. and Qiu, C.Q. (2017) Building Tianjin Driving Cycle Based on Linear Discriminant Analysis. *Transportation Research D* (53), p.78-87.

Knez,, M., Muneer, T., Jereb, B. and Cullinane, K. (2014) The Estimation of a Driving Cycle for Celje and a Comparison to other European Cities. *Sustainable Cities and Society~~ies~~* 11, pp. 56-60.

Kivekas, K., Vepsalainen, J. and Tammi, K. (2018a) Stochastic Driving Cycle Synthesis for Analysing the Energy Consumption of a Battery Electric Bus. *IEEE Access* PP(99);1-1 (10.1109/ACCESS.2018.2871574)

Kivekas, K., Lajunen, A., Vepsalainen, J. and Tammi, K. (2018b) City Bus Powertrain Comparison: Driving Cycle Variation and Passenger Load Sensitivity Analysis. *Energies*, 11(7), 1755; (doi:10.3390/en11071755).

Kumar, R., Parida, P., Durai, B.K. and Saleh, W. (2013) How Real-World Driving Cycle Differs in Heterogeneous Traffic Conditions: A Case Study in Delhi. *World Journal Science, Technology and Sustainable Development*, 10, 66-80.

Lai, J.X., Yu, L., Song, G.H., Guo, P. and Chen, X.M. (2013) Development of City-Specific Driving Cycles for Transit Buses Based on VSP Distributions: Case of Beijing. *ASCE Journal of Transportation Engineering*, 139(7), 749-757.

Li, Y.C., Liu, P., Yang, W. and Liu, J.Y. (2016) Research on Xi'an Pure Electric Bus Driving Cycle. *China Science Paper*, 11(7), 754-759.

Liu, J., Wang, X. and Khattak, A. (2016) Customising Driving Cycles to Support Vehicle Purchase and Use Decisions: Fuel Economy Estimation for Alternative Fuel Vehicle Users. *Transportation Research C* 67, pp.280-298.

Liu, X.D., Ma, J., Zhao, X., Du, J. and Xiong, Y.F. (2020) Study on Driving Cycle Synthesis Method for City Buses considering Random Passenger Load. *Journal of Advanced Transportation 2020*, (Article ID: 3871703) (https://doi.org/10.1155/2020/3871703)

Ma, R.Y., He, X.O., Zheng, Y.L., Zhou, B.Y., Lu, S. and Wu, Y. (2019) Real-World Driving Cycles and Energy Consumption Informed by Large-Sized Vehicle Trajectory Data. *Journal of Cleaner Production* 223, pp.564-574.

Mansour, C., Haddad, M. and Zgheib, E. (2018) Assessing Consumption, Emissions and Costs of Electrified Vehicles under Real Driving Conditions in a Developing Country with an Inadequate Road Transport System. *Transportation Research D* 63, pp. 498-513.

Maurya, A.K. and Bokare, P.S. (2012) Bus Driving Cycle on State Highway in Maharashtra, India – A Case Study. National Workshop on Understanding Real World Indian Driving Cycle and Its Impacts, Central Road Research Institute, New Delhi, India. 4th December, 2012.

Mayakuntla, S.K. and Verma, A. (2018) A Novel Methodology for Construction of Driving Cycles for Indian Cities. *Transportation Research D* 65, pp. 725-735.

Menard, T. and Miller, J. (2010) FreeSim_Mobile: a Novel Approach to Real-Time Traffic

Gathering using the Apple iPhone™. 2010 IEEE Vehicular Networking Conference.

Menard, T. and Miller, J. (2011) Comparing the GPS Capabilities of the iPhone 4 and iPhone 3G for Vehicle Tracking using FreeSim_Mobile. 2011 IEEE Intelligent Vehicles Symposimu (IV). Baden-Baden, Germany, June 5-9, 2011.

Mongkonlerdmanee, S and Koetniyom, S. (2019) Development of a Realistic Driving Cycle Using Time Series Clustering Technique for Buses: Thailand Case Study. *Engineering Journal* 23 (4), pp. 49-63.

Nesamania, K.S. and Subramanian, K.P. (2011) Development of a Driving Cycle for Intra-city Buses in Chennai, India. *Atmospheric Environment* 45, 5469-5476.

Nguyen, T.Y.L., Nghiem, T.D. and Cao, M.Q. (2016) Impact of the Driving Cycle on Exhaust Emissions of Buses in Hanoi. *Journal of Vietnamese Environment* 8(4), pp. 274-251.

Nguyen, T.Y.L., Nghiem, T.D., Le, A.T. and Bui, N.D. (2018) Development of the Typical Driving Cycle for Buses in Hanoi, Vietnam. *Journal of the Air & Waste Management Association*, DOI: 10.1080/10962247.2018.1543736.

Peng, Y.H., Zhuang, Y. and Yang, H.B. (2020a) Development of a representative Driving Cycle for Urban Buses Based on the K-means Cluster Method. *Cluster Computer* 22:S6871–S6880 (https://doi.org/10.1007/s10586-017-1673-y).

Quirama, L.F., Giraldo, M., Huertas, J.I. and Jaller, M. (2020) Driving Cycles that Reproduce Driving Patterns, Energy Consumptions and Tailpipe Emissions. *Transportation Research D* 82, 102294.

Seers, P., Nachin, G. and Glaus, M. (2014) Development of two Driving Cycles for Utility Vehicles. *Transportation Research D* 41, pp. 377-385

Shen, P.H., Zhao, Z.G., Li, J.W. and Zhan, X.W. (2018) Development of a Typical Driving Cycle for an Intracity Hybrid Electric Bus with a Fixed Route. *Transportation Research D* 59, pp.346-360.

Tong, H.Y., Hung, W.T., and Cheung, C.S. (1999) Development of a Driving Cycle in Hong Kong. *Atmospheric Environment*, 33, 2323-2335.

Tong, H.Y. and Hung, W.T. (2008) Driving Cycle Approaches in Characterising Driving Patterns. Proceedings of The 13th Conference of The Hong Kong Society for Transportation Studies, Hong Kong, 13-15 December 2008, pp. 137-145.

Tong, H.Y. and Hung, W.T. (2010a) Methods for Developing Driving Cycles. The 4th National Symposium on Air Resource Management in Sri Lanka – "Air That We Breathe". 18 Mar, 2010.

Tong, H.Y. and Hung, W.T. (2010b) A Framework for Developing Driving Cycles with On-Road Driving Data. *Transport Review*, 30(5), 589-615.

Tong, H.Y. (2019) Development of a Driving Cycle for a Supercapacitor Bus Route in Hong Kong. *Sustainable Cities and Society* 48, (Article ID: 101588) (https://doi.org/10.1016/j.scs.2019.101588)

Transport and Housing Bureau (2017) Public Transport Strategy Study: Final Report.

Transport Department (2019) September Monthly Transport Digests.

Tutuianu, M., Marotta, A., Steven, H., Ericsson, E., Haniu, T., Ichikawa, N. and Ishii, H. (2014)

Development of a World-wide Worldwide Harmonized Light Duty Driving Test Cycle (WLTC): Technical Report (UN/ECE/WP.29/GRPE/WLTP-IG).

Yan, W., Li, M.J., Zhong, Y.C., Qu, C.Y. and Li, G.X. (2020) A Novel k-MPSO Clustering Algorithm for the Construction of Typical Driving Cycles. *IEEE Access* 8, pp.64028-64036.

Yang, Y., Li, Y.Z., Hu, H.F., Zhang, T., Cai, X.H., Chen, S.Y. and Qiao, F.X. (2019) Development and Emissions Performance Analysis of Local Driving Cycle for Small-sized Passenger Cars in Nanjing, China. *Atmospheric Pollution Research* 10, pp.1514-1523.

Zamboni, G., Andre, M., Reveda, A. and Capobianco, M. (2015) Experimental Evaluation of Heavy ~~z~~Duty Vehicles Speed Patterns in Urban and Port Areas and Estimation of Their Fuel Consumption and Exhaust Emissions. *Transportation Research D* (35), pp. 1-10.

Zhang, F., Guo, F. and Huang, H. (2017) A Study of Driving Cycle for Electric Special-Purpose Vehicle in Beijing. *Energy Procedia* 105, p. 4884 – 4889.

Zhao, X., Ye, Y., Ma, J., Shi, P. and Chen, H. (2020) Construction of Electric Vehicle Driving Cycle for Studying Electric Vehicle Energy Consumption and Equivalent Emissions. *Environmental Science and Polluton Research* 27, pp.37395-37409.

Zhao, X, Yu, Q., Ma, J., Wu, Y., Yu, M. and Ye, Y. (2018) Development of a Representative EV Driving Cycle Based on a k-Means and SVM Hybrid Clustering Algorithm. *Journal of Advanced Transportation* 52, (Article ID: 1890753) (https://doi.org/10.1155/2018/1890753)

Figure Captions

Figure 1 Comparison of the Top-Down and Bottom-Up Approaches of Driving Cycle Development

Figure 2 Trip Average Speed versus Trip Positive Kinetic Energy of the three Clusters

Figure 3 Driving Cycle Development Process

Figure 4a Synthesized Bus Driving Cycle for Cluster 1

Figure 4b Synthesized Bus Driving Cycle for Cluster 2

Figure 4c Synthesized Bus Driving Cycle for Cluster 3

Figure 5 Comparisons of SAPDs for the Synthesized Cycles and the Whole Clusters

Figure 6 Comparisons of VSP Distributions for the Synthesized Cycles and the Whole Clusters

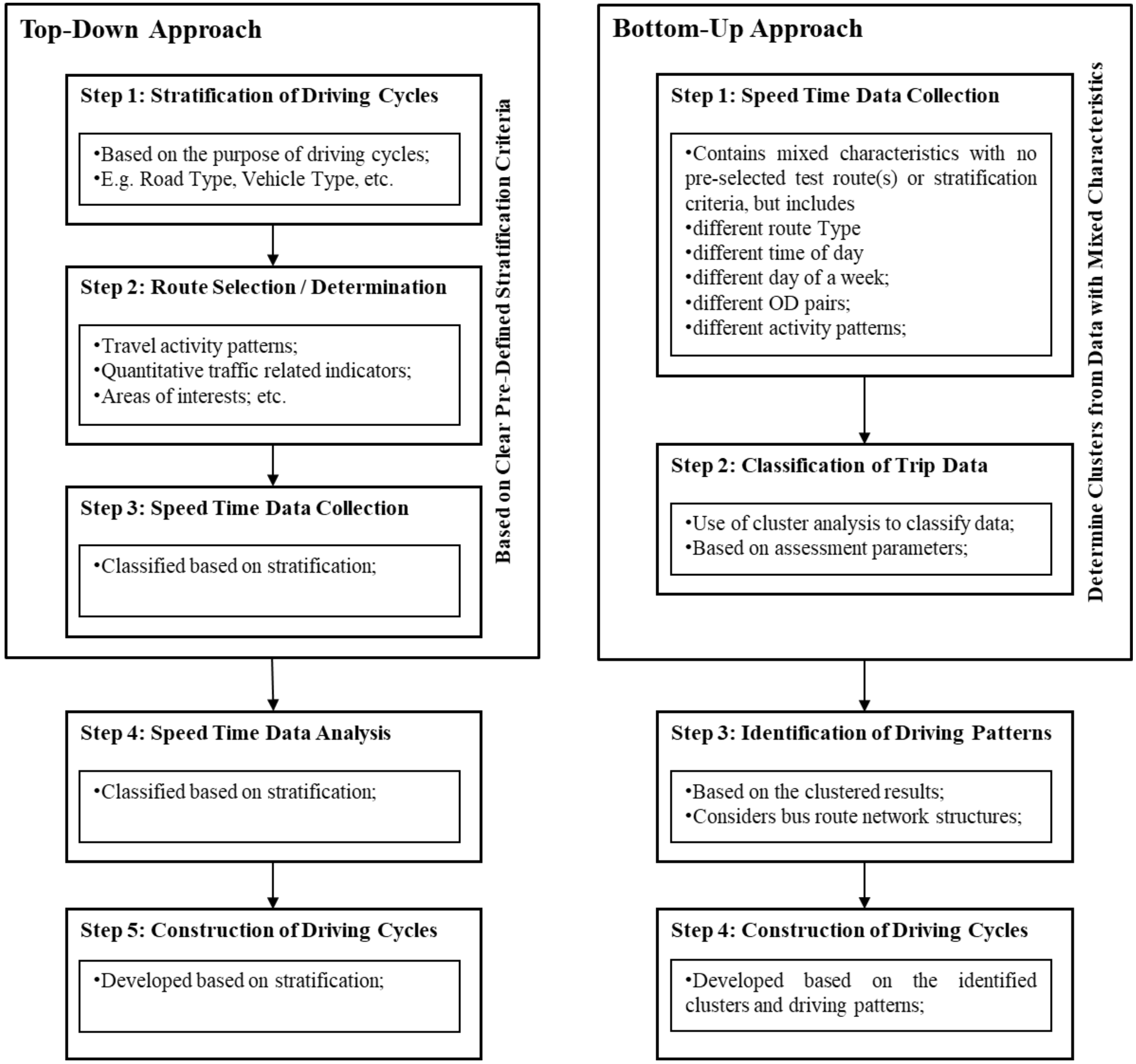


Figure 1 Comparison of the Top-Down and Bottom-Up Approaches of Driving Cycle Development

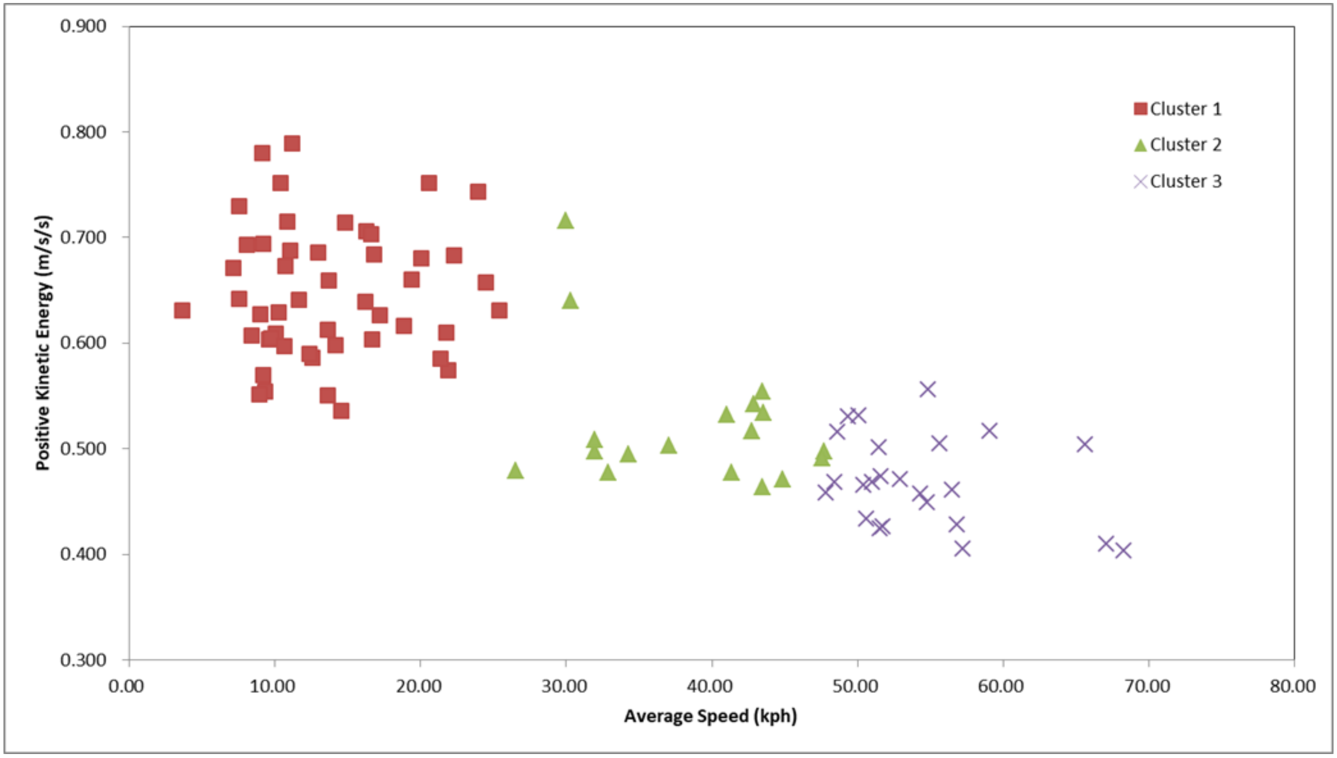


Figure 2 Trip Trip Average Speed vs Trip Positive Kinetic Energy of the three Clusters

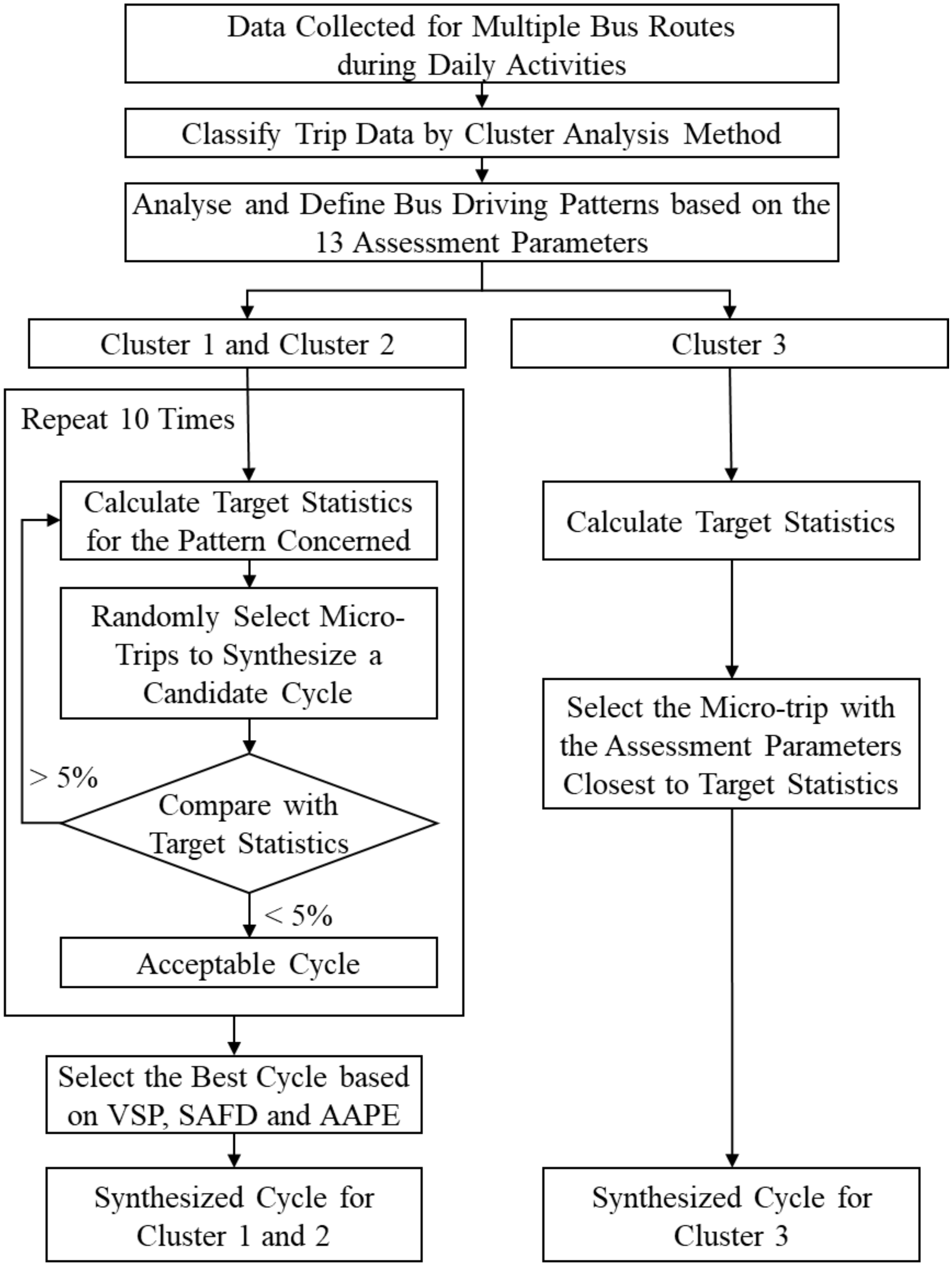


Figure 3 Driving Cycle Development Process

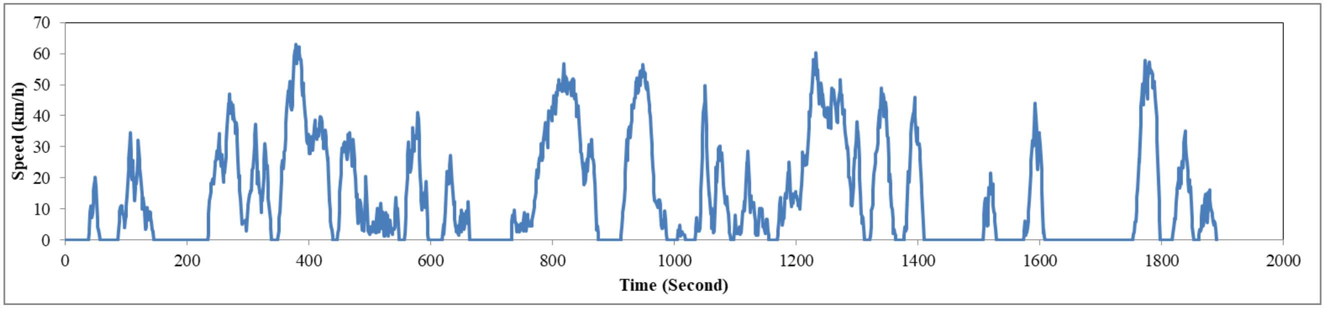


Figure 4a Synthesized Bus Driving Cycle for Cluster 1

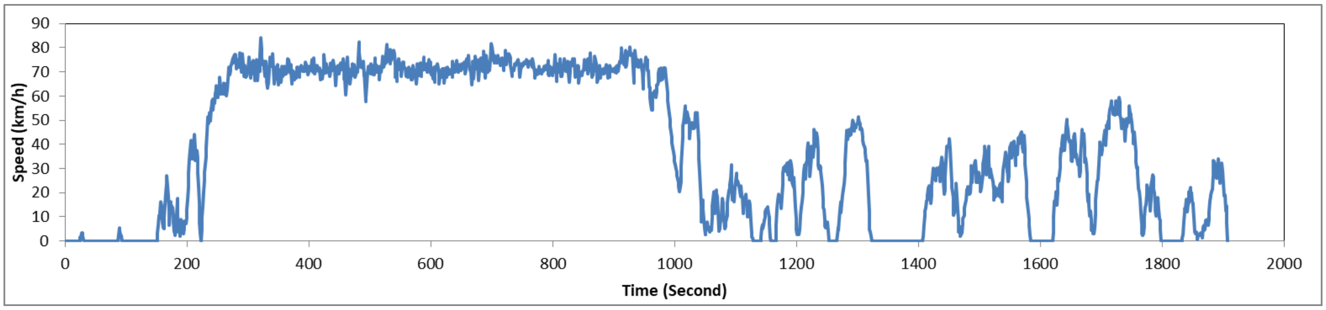


Figure 4b Synthesized Bus Driving Cycle for Cluster 2

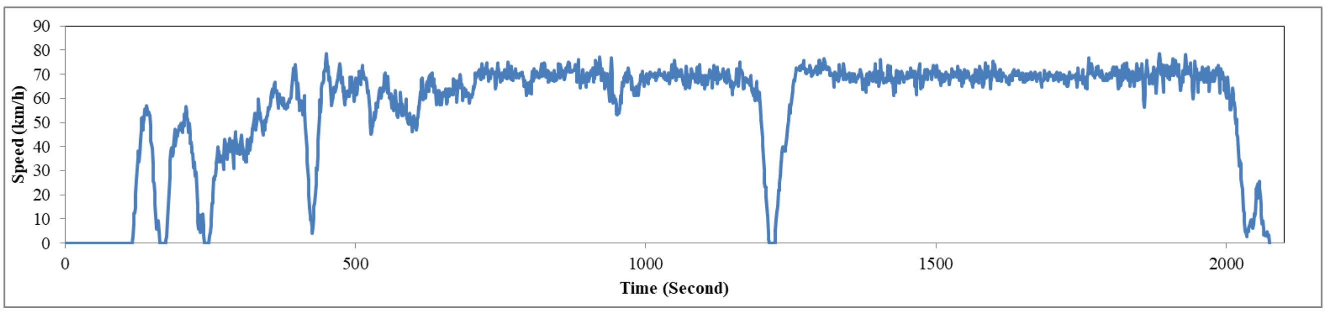
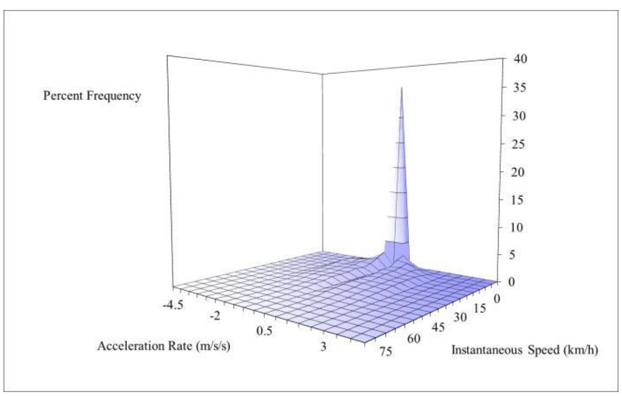
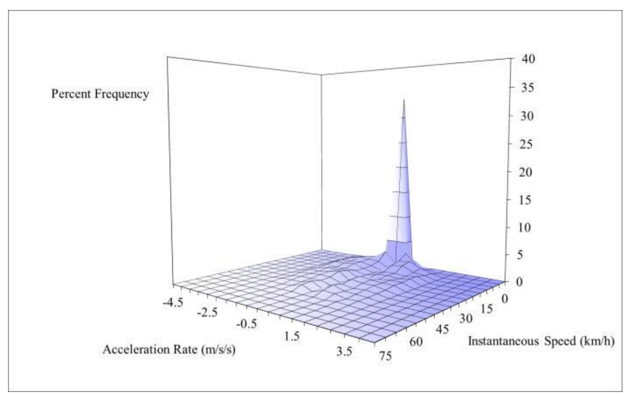
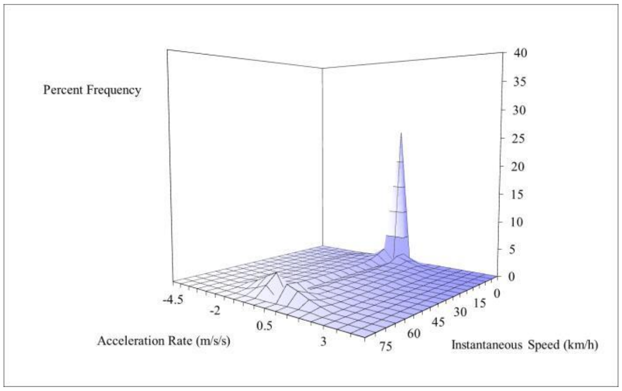


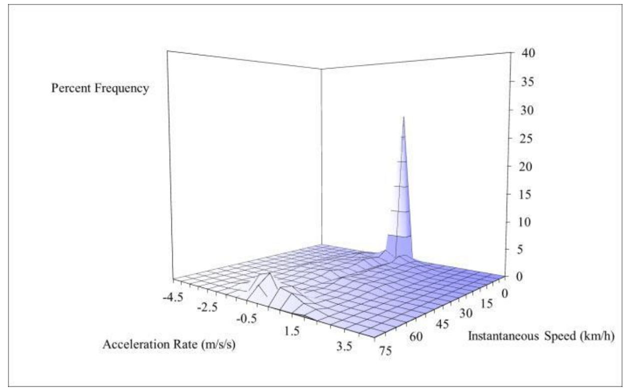
Figure 4c Synthesized Bus Driving Cycle for Cluster 3



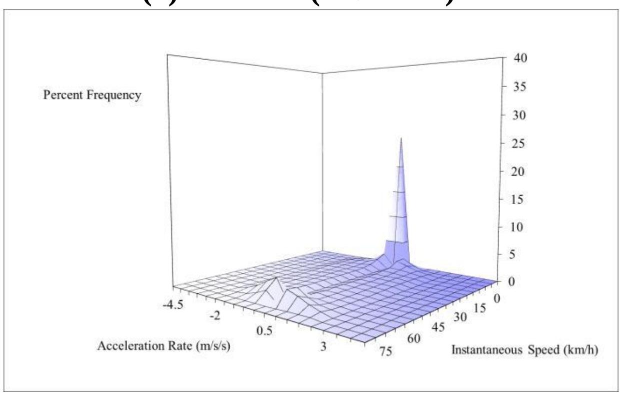


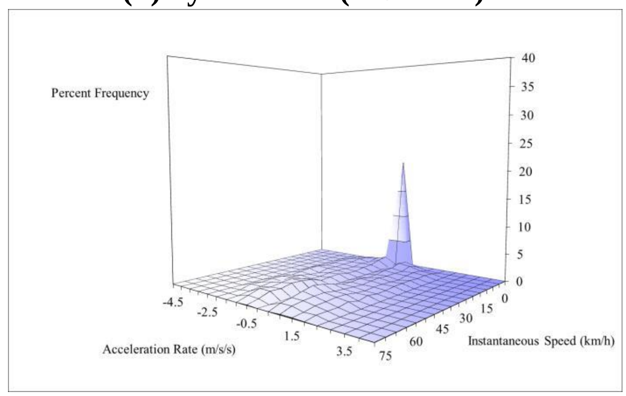
(a) All Data (Cluster 1) **(b) Synthesized (Cluster 1)**





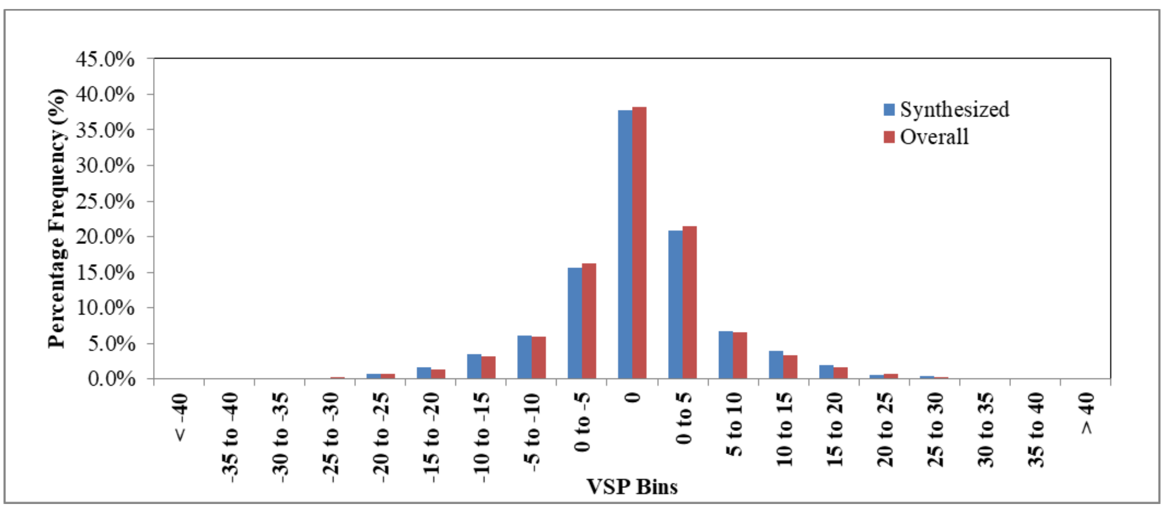
(a) All Data (Cluster 2) **(b) Synthesized (Cluster 2)**



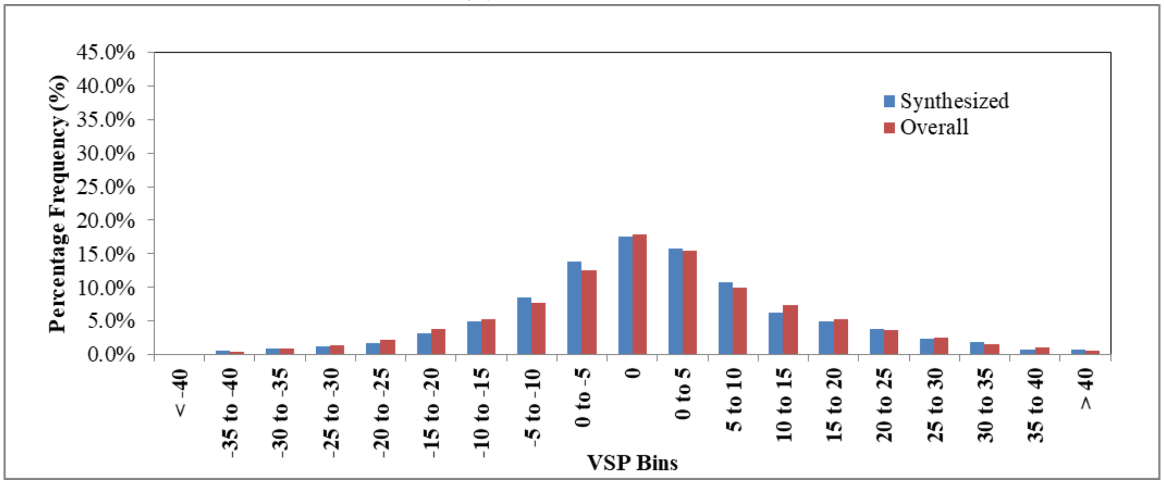


(a) All Data (Cluster 3) **(b) Synthesized (Cluster 3)**

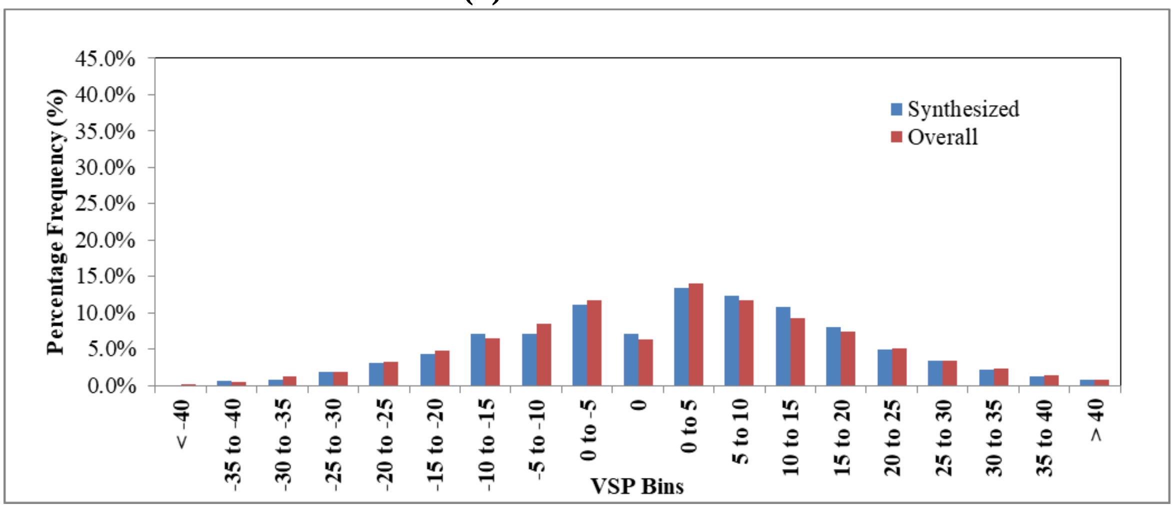
Figure 5 Comparisons of SAPDs for the Synthesized Cycles and the Whole Clusters



(a) For Cluster 1



(b) For Cluster 2



(c) For Cluster 3

Figure 6 Comparisons of VSP Distributions for the Synthesized Cycles and the Whole Clusters

Table Captions

	Top-Down Approach (Traditional)	Bottom-Up Approach (Proposed)
Stratification	□ Define based on purpose of driving cycle	□ No pre-defined stratification criteria
Test Route	□ Select based on stratification criteria	□ No pre-selected test route(s)
Data Collection	□ According to stratification criteria	□ Containing mixed characteristics
Analysis of Driving Characteristics	□ According to stratification criteria	□ Determine driving patterns by cluster analysis
Driving Cycle Development	□ According to stratification criteria	□ According to the identified clusters

Table 1 Characteristics of Top-Down and Bottom-Up Approaches for Driving Cycle Development

Operator	Number of Routes					No. of Buses	Daily Pax (in 1000)
	HKI	Kln	NT	Lantau/Airport	Cross-Harbour		
CTB	52	2	1	28	33	992	611
KMB	0	354		0	64	4065	2800
LW	0	0	0	36	0	279	125.5
NLB	0	0	1	26	0	156	96.5
NWFB	47	13		0	33	685	458
Total	99	369	2	88	132	6151	4091

Table 2 Information on the Franchised Bus Route Network in Hong Kong (as of 31 December 2019)

	HKI	Kln	Island and NT
HKI	6.3%	-	-
Kln	7.2%	21.6%	-
Island and NT	1.0%	18.3%	45.7%

Table 3 Distribution of Bus Journeys Collected

Abbr.	Name	Unit	Abbr.	Name	Unit
v_1	Average speed of the entire driving cycle	*km/h*	P_{idle}	Time proportions of idling modes	%
v_2	Average running speed	*km/h*	P_{acce}	Time proportions of acceleration modes	%
a	Average acceleration of all acceleration phases	m/s^2	P_{cruise}	Time proportions of cruising modes	%
d	Average deceleration of all deceleration phases	m/s^2	P_{dece}	Time proportions of deceleration modes	%
RMS	Root mean square acceleration	m/s^2	P_{creep}	Time proportions of creeping modes	%
PKE	Positive acceleration kinetic energy	m/s^2	M	Average number of acceleration/deceleration changes (and vice versa) within one micro-trip	*number of times*
c	Mean length of a micro-trip	*sec*			

Table 4 The list of 13 Assessment Parameters Adopted for Quantitative Analysis

Cluster	Weekends	Weekdays	Total
1	30.6%	69.4%	100.0%
2	50.0%	50.0%	100.0%
3	58.3%	41.7%	100.0%

Table 5 Clustered Statistics by Day of a Week

Cluster	AM Peak (07:00 – 10:00)	Inter-Peaks (10:00 – 17:00)	PM Peak (17:00 – 21:00)	After Peak (21:00 – 07:00)
1	85.7%	56.4%	100.0%	29.4%
2	0.0%	17.9%	0.0%	32.4%
3	14.3%	25.6%	0.0%	38.2%
Total	100.0%	100.0%	100.0%	100.0%

Table 6 Clustered Statistics by Time of a Day

Cluster	HKI	Inter-District	NT and Island	Kln
1	66.7%	11.1%	100.0%	97.1%
2	33.3%	35.6%	0.0%	2.9%
3	0.0%	53.3%	0.0%	0.0%
Total	100.0%	100.0%	100.0%	100.0%

Table 7 Clustered Statistics by Districts

	P_{idle} (%)	P_{acce} (%)	P_{cruise} (%)	P_{dece} (%)	P_{creep} (%)	*RMS* (m/s^2)	*PKE* (m/s^2)	*a* (m/s^2)	*d* (m/s^2)	v_1 (km/h)	v_2 (km/h)	*C* (*s*)	*M* (*time*)
Cluster 1	36.81	29.27	3.98	28.34	1.60	1.018	0.648	0.874	0.902	13.79	21.63	58.7	22.2
Cluster 2	17.57	37.93	6.41	37.33	0.75	0.968	0.518	0.827	0.838	38.57	46.59	161.5	76.6
Cluster 3	5.71	43.04	8.46	42.53	0.26	0.921	0.469	0.779	0.789	54.41	57.63	407.1	226.4

Table 8 Summary Statistics of the three Clusters

	P_{idle} (%)	P_{acce} (%)	P_{cruise} (%)	P_{dece} (%)	P_{creep} (%)	*RMS* (m/s^2)	*PKE* (m/s^2)	*a* (m/s^2)	*d* (m/s^2)	v_1 (km/h)	v_2 (km/h)	*C* (*s*)	*M* (*time*)	*AAPE*	*SSD*
Cluster 1															
Synthesized	36.68	29.32	3.86	28.53	1.58	0.984	0.619	0.851	0.873	13.80	21.79	59.79	23.1	38.35	5.50

Target	36.81	29.27	3.98	28.34	1.60	1.018	0.648	0.874	0.902	13.79	21.63	58.70	22.2	–	–
Cluster 2															
Synthesized	16.99	37.72	6.66	37.88	0.73	0.964	0.506	0.838	0.835	38.82	46.78	158.19	75.0	23.91	43.51
Target	17.57	37.93	6.41	37.33	0.75	0.968	0.518	0.827	0.838	38.57	46.59	161.50	76.6	–	–
Cluster 3															
Synthesized	6.89	43.00	9.25	40.59	0.24	0.848	0.427	0.708	0.750	56.82	61.03	482.75	268.00	120.51	29.49
Target	5.71	43.04	8.46	42.53	0.26	0.921	0.469	0.779	0.789	54.41	57.63	407.10	226.40	–	–

Table 9 Summary Statistics of Driving Cycles for the three Clusters

Location	v_1 (km/h)	a (m/s^2)	d (m/s^2)	P_{idle} (%)	P_{aece} (%)	P_{cruise} (%)	P_{dece} (%)	P_{creep} (%)	Source
Cluster 1 (Urban Congested)	13.8	0.851	0.873	36.7	29.3	3.9	28.5	1.6	This Study
Cluster 2 (Inter-district)	38.8	0.838	0.835	17.0	37.7	6.7	37.9	0.7	This Study
Cluster 3 (Express and Mid-Night)	56.8	0.708	0.750	6.9	43.0	9.3	40.6	0.2	This Study
Maharashtra Highway AM	37.7	0.280	0.370	5.2	26.0	46.6	21.5	-	Maurya and Bokare, 2012
Maharashtra Highway Off-Peak	44.1	0.330	0.310	4.3	27.0	53.5	15.5	-	Maurya and Bokare, 2012
Maharashtra Highway PM	31.0	0.280	0.600	9.6	38.2	38.5	18.5	-	Maurya and Bokare, 2012
Maharashtra Urban AM	15.1	0.500	0.470	35.9	28.3	-	27.4	-	Maurya and Bokare, 2012
Maharashtra Urban PM	18.6	0.600	1.500	39.9	29.2	-	29.3	-	Maurya and Bokare, 2012
Route 11	20.4	0.540	0.510	16.7	32.4	15.1	34.4	1.5	Kivekas et al., 2018a
Route 11	23.8	0.380	0.390	13.3	-	22.2	-	0.1	Kivekas, et al., 2018b
Route 24	17.3	0.680	0.730	15.9	-	9.3	-	0.1	Kivekas, et al., 2018b
Route 550	30.5	0.500	0.500	14.1	-	16.8	-	0.1	Kivekas, et al., 2018b
Route 03	18.2	0.650	0.680	18.5	-	10.6	-	0.3	Kivekas, et al., 2018b
Route 25	20.4	0.710	0.770	16.3	-	8.9	-	0.2	Kivekas, et al., 2018b

Kanchanaburi DC	-	-	-	-	52.1	1.8	46.2	-	Mongkonlerdmanee, 2019
Hanoi Bus DC	17.3	0.480	0.510	5.3	34.5	-	32.4	-	Nguyen, et al., 2016
HBDC 2018	16.8	0.500	0.520	7.6	34.2	14.1	32.7	11.4	Nguyen, et al., 2018
Chennai Bus DC	14.0	0.650	0.710	32.2	29.8	3.5	29.6	4.9	Nesamania, et al., 2011
Delhi AM DC	26.6	-	-	14.6	39.9	8.1	37.4	-	Kumar, et al., 2013
Delhi OffPeak DC	26.3	-	-	16.1	39.2	8.4	36.3	-	Kumar, et al., 2013
Delhi PM DC	27.8	-	-	14.0	39.7	9.7	36.6	-	Kumar, et al., 2013
XiBUS Arterial	18.4	0.509	0.504	10.8	40.3	16.8	32.3	-	Li, et al., 2016
XiBUS Composite	16.9	0.420	0.460	12.8	38.2	15.7	32.9	-	Li, et al., 2016
XiBUS Urban	15.8	0.404	0.580	18.3	35.5	22.0	24.3	-	Li, et al., 2016
XiBUS Highway	32.9	0.422	0.590	2.3	45.8	15.0	37.0	-	Li, et al., 2016
Xi'an Bus DC	18.2	-	-	23.6	27.8	25.3	23.3	-	Liu, et al., 2020
Shanghai HEB DC	23.0	0.710	0.830	34.0	33.0	5.0	28.0	-	Shen, et al., 2018
Fuzhou Bus DC	13.8	0.740	-	34.4	27.0	15.5	23.1	-	Peng, et al., 2019a
Mexico City Urban 1 DC	7.30	0.500	0.500	15.5	32.9	22.7	29.3	-	Quirama, et al., 2020
Mexico City Urban 2 DC	10.0	0.400	0.500	13.6	33.8	25.9	29.1	-	Quirama, et al., 2020

Table 10 Comparison of International Bus Driving Cycle Characteristics