Driving cycle development based on large-scale GPS data and implications on vehicle energy consumption

Ruoyun Ma[1], Xiaoyi He[1], Shaojun Zhang[2], Ye Wu[1,3*], Wei Shen[4], Weijian Han[5]

[1] School of Environment, and State Key Joint Laboratory of Environment Simulation and Pollution Control, Tsinghua University, Beijing 100084, China

[2] Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York 14853, USA

[3] State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, China

[4] Asia Pacific Research, Ford Motor Company, Unit 4901, Tower C, Beijing Yintai Center, No.2 Jianguomenwai Street, Beijing 100022, China

*Corresponding authors: Ye Wu (ywu@tsinghua.edu.cn)

Abstract

Researches have shown that inadequate representativeness of testing cycles could lead to significant discrepancy in vehicle energy and emission assessment between on-road performance. On-board measurement has been used in past researches to collect vehicle activity data but the amount of data is usually insufficient, and the statistical coverage of various road/traffic types could not be guaranteed. Using "big data" mining techniques, this study examines second-by-second GPS data of 459 private passenger cars in Beijing, covering over 17,000 sampling days to characterize vehicle speed profiles under various traffic condition and road types. We then applied the Markov Chain method to generate sub-cycles and corresponding weighting factors that have similar properties as real-world driving. The resulting cycles (i.e., Off-peak Cycle, Peak Cycle) are combination of sub-cycles representing different road types and traffic conditions, which depict fine-scale discrepancies of driving characteristics among different situations. An application of vehicle fuel consumption simulation shows that the developed typical driving cycle leads to up to 20% higher fuel consumption than regulation test cycles (i.e., NEDC, WLTC). This study proposes a practical method to construct driving cycle from massive GPS data and highlights the importance for developing representative driving cycles of a certain city/region for legislation fuel consumption and emission test.

Keywords

Driving cycle; large-scale GPS data; Markov Chain process; Spatial and temporal classification; Vehicle energy consumption

1. Introduction

Driving cycles are widely used in dynamometer vehicle fuel consumption and emission testing. A driving cycle reproduces typical vehicle travel patterns in a certain city or region, usually presented as velocity-time profiles. Standard driving cycles such as the NEDC (New European Driving Cycle), FTP-75 (Federal Test Procedure-75) and JC08 (Japan Cycle 08) are used in the authoritative testing for vehicle exhaust emissions and fuel consumption in different countries. In China, the NEDC has been adopted for national regulation for vehicle exhaust testing, which however, is widely acknowledged of significant deviation from real-world results. Zhang *et al*. [Zhang *et al*., 2014] reported that fuel consumption normalized to standard NEDC underrates about 10% compared to on-road results under average realistic driving pattern. Several studies in the Europe also indicates that in-use fuel consumption was significantly higher than the type-approval fuel consumption by 10%~50% [Ntziachristos *et al*., 2014; Rubino *et al*., 2007]. As NEDC is a modal driving cycle, which is designed feasibly for vehicle bench tests, it neglects speed fluctuation and consequent additional acceleration. Meanwhile, there is a prevalent consensus that the driving pattern of a certain city or region is distinguished by its fleet composition, road network topography and drivers' behavior [Andre *et al.* 2006]. Therefore, it is essential to aggregate multi-aspect local traffic information to develop its own driving cycle reproducing realistic characteristics of a specific city as far as possible.

China is a developing country with the world's largest auto market since 2009, and Beijing is the city with the largest vehicle population (i.e., 5.5 million vehicles by June 2017) [BTMB, 2017] within China. According to Ministry of Environmental Protection of China [MEP, 2018], mobile sources emission has become the primary emission source and accounted for 45% of local $PM_{2.5}$ emissions in Beijing. Meanwhile, increasing concerns about fuel consumption has been stressed in automotive sector, which takes up around 90% and 45% of total gasoline and diesel consumption in China [Wu *et al*., 2017]. As the prerequisite of related legislation and policy research, reliable and accurate test procedures of vehicle fuel consumption and emissions is important, where representative driving cycles should be developed locally.

The traditional micro-trip approach for driving cycle development divides the trip into segments, among which eigen micro-trips are extracted and juxtaposed together. This method has been adopted by researches in Europe [Andre, 2004], India [Kamble *et al*., 2009], Hong Kong [Hung *et al.*, 2007], and several major cities in Mainland China [Wang *et al.*, 2008]. However, micro-trio approach is disputed for ignoring the essential transfer patterns between driving modes, as well as the considerable uncertainty on the representativeness of selected segments [Bishop *et al.*, 2012]. A new approach using Markov Chain process proposed by Lin and Niemeier [Lin and Niemeier, 2003] has gained increasingly attention [Bishop *et al.*, 2012; Gong *et al.*, 2011; Nyberg *et al.*, 2014]. Compared with traditional methods, the Markov Chain approach captures how speed and driving status evolve during a trip without deconstructing intrinsic original behavior patterns. Shi et al. [Shi *et al.*, 2016] verified the Markov property of driving cycle from both theoretical and experimental aspects, and proved that the state transition matrix is a more appropriate expression of substantive features compared to velocity-acceleration joint distribution probability.

It is the primary principle for developing driving cycles to acquire realistic and representative driving data. The car-chasing technique has been prevalent for data collection, which means a target car selected randomly is followed by a chase car in defined routes with speed data reserved second by second [Hung *et al.*, 2007; Wang *et al.*, 2008; Lin and Niemeier, 2002]. Inevitably, route selection and test time become crucial for data reliability, where researchers' subjective judgement cannot be excluded completely. The other method for data collection is on-board measurement. The application of recording instruments like GPS addresses test route limitation and onerous experimental tasks, and is able to observe real-world driving characteristics in an extended period. However, the critical drawback it has to face lies in the constraints of small data scale or the access to realistic data, as 10 electric scooters and 9 plug-in electric vehicles were employed in Bishop et al.'s [Bishop *et al.*, 2012] and Gong et al.'s [Gong *et al.*, 2011] work. It appears consequential to expand the dataset to cover diverse travel behaviors, especially with the trend of applying large-scale data for environmental assessment nowadays.

The uniqueness of this study compared with previous studies is that large scale GPS trajectory data collected from 459 actual personal-owned cars were used as original dataset, with nearly 17,000 sampling days and 3.3 million km travelled in total. This is the first time that large scale vehicle trajectory data collected from a megacity in East Asia has been used in developing driving cycles. It can greatly improve representativeness for real-world traveling by actual users, who may have different driving behaviors from employed drivers driving on predetermined routes. The scale of the data profiles also ensures diverse spatial and temporal range.

This study tackles three major tasks: data collection and processing, driving cycle construction, and vehicle fuel consumption simulation. Sub-cycles defined by road types and traffic conditions (implied by peak hour driving or non-peak hour driving) were merged to create typical driving cycles. Vehicle fuel consumption simulation was further performed for typical driving cycles and compared with standard cycles.

2. Methodology

This study applied an iterative Markov Chain process to develop typical driving cycles. Markov process has been widely applied to problems with stochastic process in many fields, like behavior analysis, financial forecasting and so on. As the theoretical basis of this data-driven model, the principle of the Markov Chain process is introduced briefly first, following by detailed procedures of developing driving cycles (Figure 1).

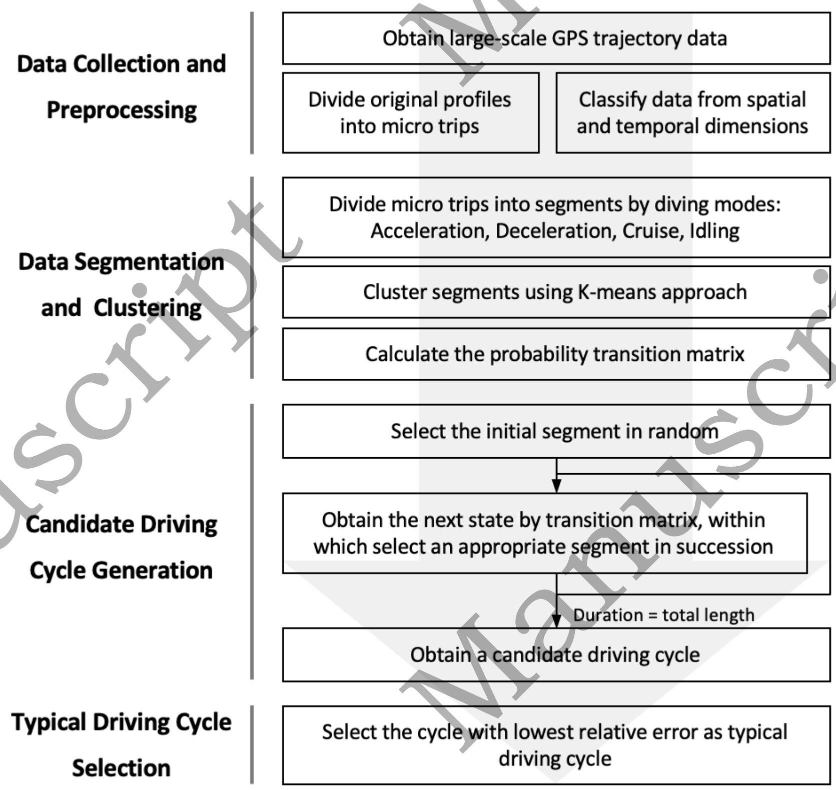


Figure 1. Flow Chart for Driving Cycle Development Process

Principal of the Markov Chain process

A Homogeneous Markov Chain describes a series of stochastic processes, where the status of the next moment, i.e., moment $n+1$, depends only on the status of the present moment n, instead of moments before moment n. Therefore, the probability of the status of moment $n+1$ could be described as

$$P\{X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n\} = P\{X_{n+1} = i_{n+1} | X_n = i_n\} \quad (2\text{-}1)$$

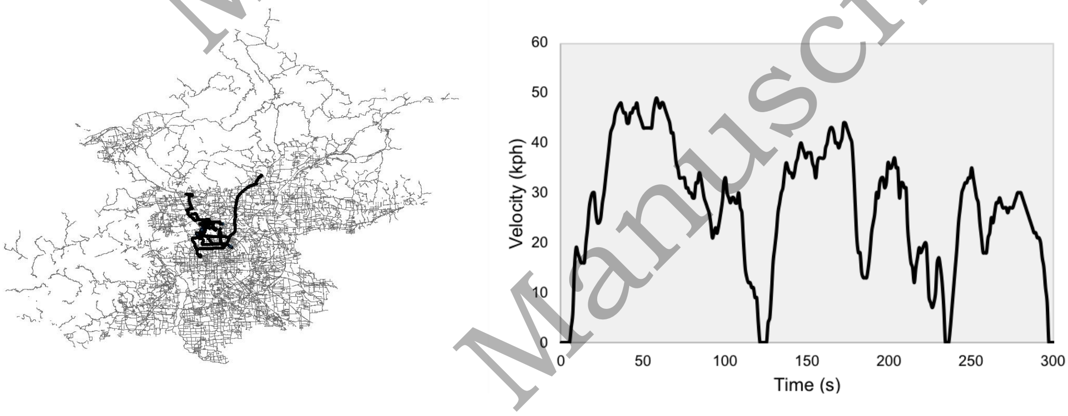
where $\{X_n, n = 0,1,2, \dots\}$ is a series of stochastic process. $P\{X_{n+1} = i_{n+1} | X_n = i_n\}$ is the probability of the next status happening of moment $n+1$, given the current status of moment n.

Second-by-second velocity of vehicles, denoted by $V(t_n)$, can be regarded as a random sequence with time advancing. As driving behavior next moment relies on the present instead of the past [Lin and Niemeier, 2003], the variation of vehicle driving states can be regarded as a Homogeneous Markov Chain.

Data Collection and Preprocessing

Our previous research obtained large-scale GPS trajectory data from 459 private gasoline cars in Beijing, China [He, et al, 2016]. Original data contains second-by-second time, location (i.e., latitude, longitude and altitude), speed, acceleration, and direction. A filtration program was developed to eliminate inaccurate data caused by unstable GPS signal. Then, a preprocessing step was conducted to divide original profiles into micro trips with stop periods no more than 5 minutes.

The spatial and temporal features of each data point are classified according to three road types (highway & freeway, arterial road and residential road) and two traffic conditions (peak hour travels from 7:00-9:00am and 17:00-19:00pm on weekdays, and off-peak hour travels for other time on weekdays and all weekends).



(a) (b)

Figure 2. Examples of (a) GPS trajectories and (b) speed profiles for a private car in Beijing

Data Segmentation and Clustering

One trip profile could last for from several minutes to several hours. They were divided into segments according to four types of driving modes: acceleration, deceleration, cruise, and idling (see Table 1 for definition). The derived segments usually last for several seconds (see Figure 3 for an example) and were used as the data pool for Markov chain modeling.

Table 1. Definition of four types of driving modes

Mode	Acceleration	Deceleration	Idling	Cruise
Acceleration (m/s²)	> 0.28	< -0.28	[-0.28, 0.28]	[-0.28, 0.28]
Velocity (km/h)	– –	– –	0	> 0

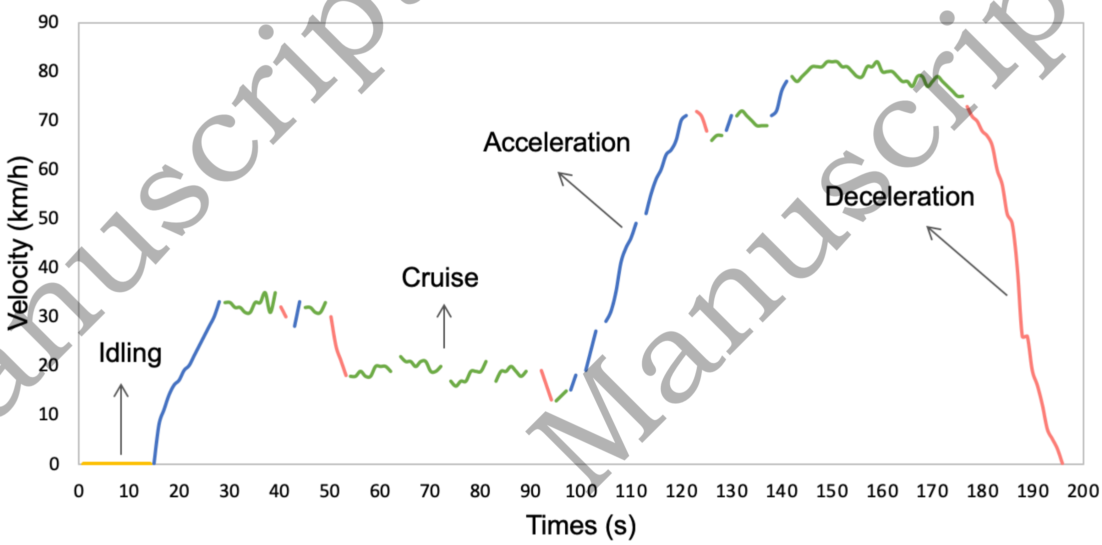


Figure 3. Example of data segments

To cooperate the original data profiles with the Markov model, a K-means clustering method, rather than divided by predetermined velocity thresholds, was adopted to gather trip segments clusters as the input of the Markov model. These segments are clustered into seven states, with average velocity ranging from 2~120 km/h at the interval of 10~20 km/h. Segments that correspond to one state share the most homogeneous mean speed and are most differentiated from segments corresponding to another state. Every segment corresponds to only one state, and the all states constitute the state space of the Markov Chain. By using the K-means approach, each cluster not only has enough data points with most similar characteristics, but also is most different compared with other clusters, which improves the efficiency of the Markov model.

After clustering the segments into seven states, the conditional probability of state i transferring to state j, $p_{ij} = P\{X_{n+1} = j|X_n = i\}\ (i, j \in S)$, can be calculated by counting the number of segments that transfer from state i to state j, denoted as N_{ij}.

$$p_{ij} = \frac{N_{ij}}{\sum_{j \in S} N_{ij}} \qquad (2\text{-}2)$$

There must be: (1) $p_{ij} \geq 0,\ i, j \in S$ (2) $\sum_{j \in S} p_{ij} = 1\ (\forall\, i \in S)$.

For the Homogeneous Markov Chain $\{X_n\}$, the transition matrix is constituted by transitional probabilities from every one state to another.

Candidate Driving Cycle Generation

Candidate driving cycles were generated in an iterative approach. One segment with the 0 initial velocity is selected in random, and its state serves as the initial state of the cycle. Then the state of the next segment is obtained according to corresponding transition probability, and the chain is succeeded by filtering segments from this state with appropriate initial velocity that can link up the last segment. By continuing this process to generate the next state and select the appropriate segment in succession until duration meets the total length, a candidate cycle was developed.

Typical Driving Cycle Selection

A series of candidate cycles had been developed, among which typical driving cycles that have the most similar characteristics with the original dataset were to be selected. The selection of candidate cycles depends on the comparison with two categories of sample characteristics parameters, numerical indexes including average velocity, running velocity (average velocity except idling), velocity_90 (the 90th percentile of velocity, the same below), average acceleration, average deceleration, and proportional indexes including ratio of four modes (acceleration, deceleration, idling and cruise) and ratio of each state. Typical driving cycle would be eventually selected with the lowest error within 5% (loosen the criteria to 10% for less important parameters).

3. Results

3.1 Traffic Characteristics Analysis

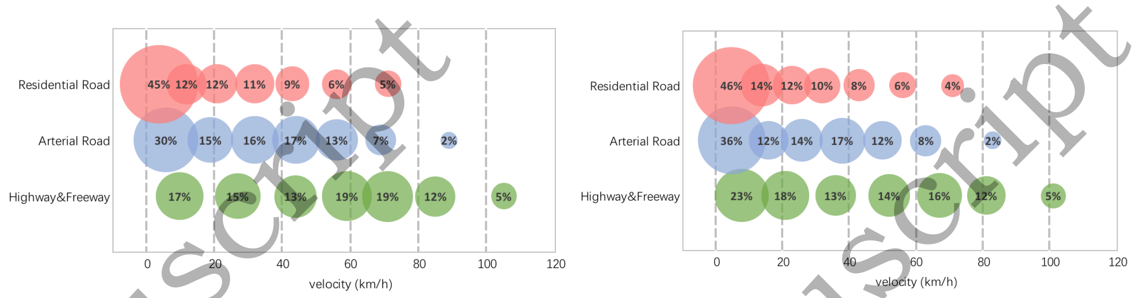
Velocity

Velocity is an important indicator of driving patterns and could impact vehicle emissions, fuel consumption, and traffic efficiency in urban scale. A k-means clustering approach was conducted to group driving segments with similar average velocity, which resulted in seven driving states for every driving situation. Vehicle velocity cluster centers for the seven states under different road types and traffic conditions are shown in Table 1. Generally speaking, the cluster centers of off-peak driving have higher velocity than that of peak hour driving in the same state. Exceptions were observed for residential road, where the velocity cluster centers for off-peak driving have similar or slightly lower velocities than peak hour driving. This could be explained that traffic condition has minor impact on residential road, compared with its impact on highway & freeway and arterial roads. We also observe that the discrepancies between off-peak hours and peak hours driving velocity cluster center are most significant for highway & freeway (2.0~7.9 km/h), followed by arterial road (0.9~6.9 km/h), and least significant for residential roads (-1.2~0.1 km/h).

Table 1. Velocity Cluster Centers based on a k-means Approach (km/h)

Driving States	Highway & Freeway		Arterial Road		Residential Road	
	Off-peak	Peak	Off-peak	Peak	Off-peak	Peak
1	8.2	6.2	4.1	3.2	2.3	2.8
2	24.5	19.0	17.3	13.5	10.5	11.7
3	41.7	33.8	30.1	24.2	19.5	20.7
4	56.6	50.3	42.1	35.5	29.8	30.3
5	69.5	64.7	53.9	47.6	41.0	41.1
6	83.1	79.3	67.2	61.2	53.7	53.6
7	103.0	98.7	86.9	80.0	68.7	68.6

Figure 2 further presents the velocity cluster distribution and proportions. Over 40% segments gather in clusters with average velocity over 60 km/h on highway & freeway during off-peak hours; on the other hand, over 60% segments gather in clusters with average velocity less than 20 km/h on residential road during peak hours. These differences coincide with the speed limits of usually 50-120 km/h for highway & freeway and less than 60 km/h for other roads, met by most sample segments. Remaining exceptions could be owing to irregular driving behavior erroneous road recognitions during space identification, which requires the promotion for higher precision of map matching in further researches.



(a) Off-peak hours (b) Peak hours

Figure 2. Velocity Centers and Proportions of Different Driving States. Center of circles indicate cluster velocity center and the areas of circles indicate proportions of each states.

The transition patterns among different driving states are given by one-step state transition probability matrix, as shown in Figure 3. Results show that the next moment segments are most likely to remain in the current state (probability > 0.5), especially for state 1 (lowest velocity) and state 7 (highest velocity). Driving states could transit to adjacent states, but are less likely to transit to states that has significant different average velocity, for example, from state 1 to state 5-7. Transition among different states is less frequent on highway & freeway driving than on arterial road and residential road, implying more stable driving patterns on highway & freeways.

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.837	0.147	0.011	0.003	0.001	0.000	0.000
State 2	0.146	0.719	0.124	0.009	0.002	0.000	0.000
State 3	0.015	0.125	0.706	0.146	0.007	0.001	0.000
State 4	0.003	0.012	0.101	0.763	0.119	0.002	0.000
State 5	0.001	0.003	0.008	0.126	0.777	0.084	0.000
State 6	0.001	0.001	0.003	0.006	0.146	0.796	0.048
State 7	0.000	0.001	0.001	0.001	0.003	0.129	0.865

(a) Off-peak Highway & Freeway

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.734	0.213	0.036	0.011	0.004	0.002	0.000
State 2	0.192	0.545	0.223	0.031	0.006	0.003	0.001
State 3	0.044	0.170	0.580	0.180	0.022	0.004	0.001
State 4	0.011	0.045	0.158	0.638	0.136	0.011	0.001
State 5	0.005	0.012	0.046	0.154	0.680	0.100	0.003
State 6	0.003	0.007	0.021	0.029	0.149	0.745	0.047
State 7	0.003	0.006	0.009	0.020	0.025	0.170	0.767

(b) Off-peak Arterial Road

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.787	0.167	0.025	0.011	0.006	0.003	0.001
State 2	0.223	0.543	0.202	0.022	0.005	0.003	0.001
State 3	0.041	0.203	0.546	0.182	0.022	0.005	0.002
State 4	0.019	0.034	0.193	0.582	0.154	0.015	0.003
State 5	0.009	0.010	0.045	0.167	0.636	0.123	0.010
State 6	0.004	0.006	0.014	0.036	0.148	0.699	0.093
State 7	0.003	0.004	0.007	0.015	0.028	0.147	0.795

(c) Off-peak Residential Road

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.821	0.164	0.010	0.003	0.001	0.000	0.000
State 2	0.161	0.704	0.125	0.008	0.002	0.000	0.000
State 3	0.019	0.168	0.686	0.120	0.007	0.001	0.000
State 4	0.005	0.016	0.113	0.736	0.127	0.003	0.000
State 5	0.002	0.004	0.011	0.111	0.781	0.091	0.001
State 6	0.001	0.002	0.003	0.007	0.135	0.806	0.048
State 7	0.000	0.001	0.001	0.001	0.004	0.127	0.866

(d) Peak Highway & Freeway

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.726	0.215	0.038	0.014	0.004	0.002	0.000
State 2	0.218	0.510	0.225	0.037	0.007	0.003	0.001
State 3	0.060	0.187	0.510	0.209	0.028	0.004	0.001
State 4	0.018	0.045	0.169	0.587	0.164	0.016	0.001
State 5	0.006	0.011	0.053	0.153	0.663	0.111	0.003
State 6	0.003	0.006	0.018	0.038	0.149	0.733	0.052
State 7	0.003	0.006	0.008	0.018	0.030	0.162	0.773

(e) Peak Arterial Road

probability	State 1	State 2	State 3	State 4	State 5	State 6	State 7
State 1	0.776	0.178	0.024	0.010	0.006	0.003	0.002
State 2	0.216	0.559	0.192	0.024	0.006	0.003	0.001
State 3	0.043	0.224	0.518	0.184	0.025	0.005	0.002
State 4	0.020	0.046	0.206	0.551	0.158	0.015	0.003
State 5	0.011	0.013	0.053	0.172	0.617	0.124	0.010
State 6	0.006	0.009	0.018	0.041	0.153	0.685	0.088
State 7	0.005	0.006	0.009	0.017	0.030	0.150	0.784

(f) Peak Residential Road

Figure 3. State Transition Probability Matrix (from the row state to the column state)

Acceleration and Deceleration

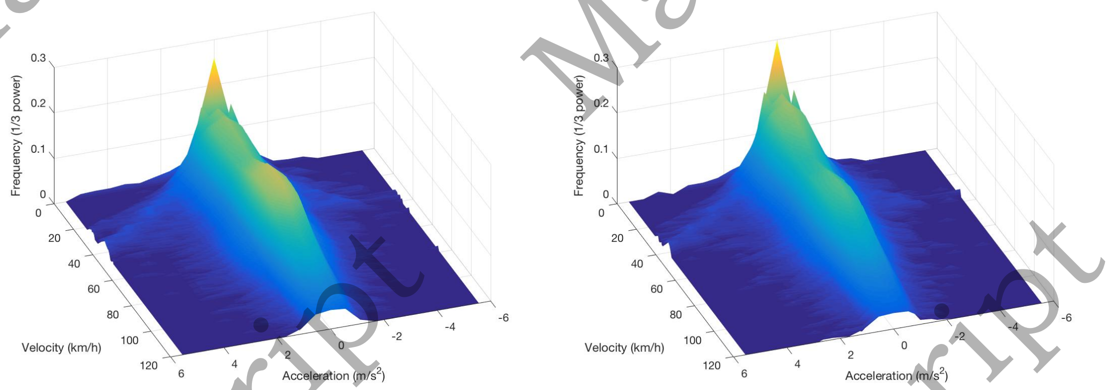
Acceleration and deceleration are significant indexes reflecting how fast driving velocity changes. Generalized acceleration refers to the rate of change of velocity, which could be then specified as acceleration phase and deceleration phase by whether greater or less than zero (with an threshold of ± $0.1 m/s^2$) .

Table 2 presents average acceleration and deceleration of each driving segments from acceleration and deceleration phases. There is a significant lift on absolute value from acceleration to deceleration for each situation, which is primarily because sharp braking is more likely to occur than sharp speeding up. Among three road types, absolute value of acceleration or deceleration are smallest for highway & freeway, followed by residential road and arterial road. Surprisingly, the travel time along does not have significant influence on the absolute value of acceleration or deceleration for a given road type, with the relative error less than 2%.

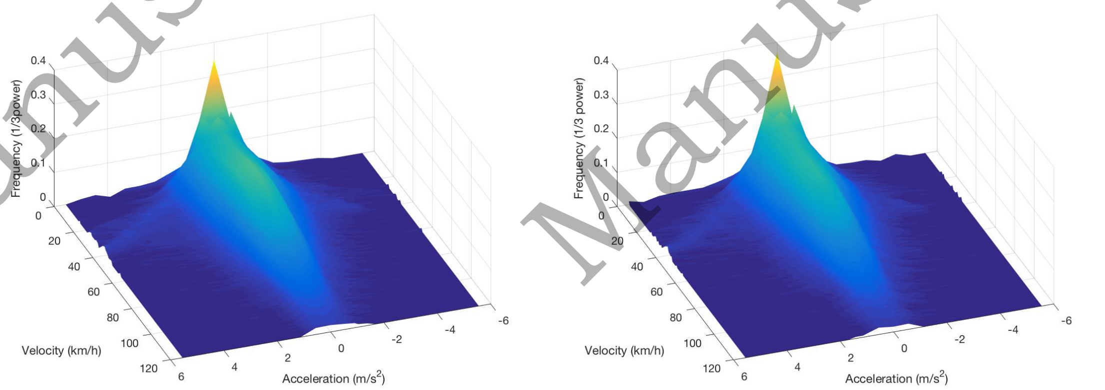
Table 2. Average Acceleration and Deceleration for Different Road Types and Traffic Conditions (m/s^2)

	Highway & Freeway		Arterial Road		Residential Road	
	Off-peak	peak	Off-peak	peak	Off-peak	peak
Acceleration	0.474	0.483	0.517	0.521	0.526	0.527
Deceleration	-0.494	-0.501	-0.539	-0.541	-0.537	-0.536

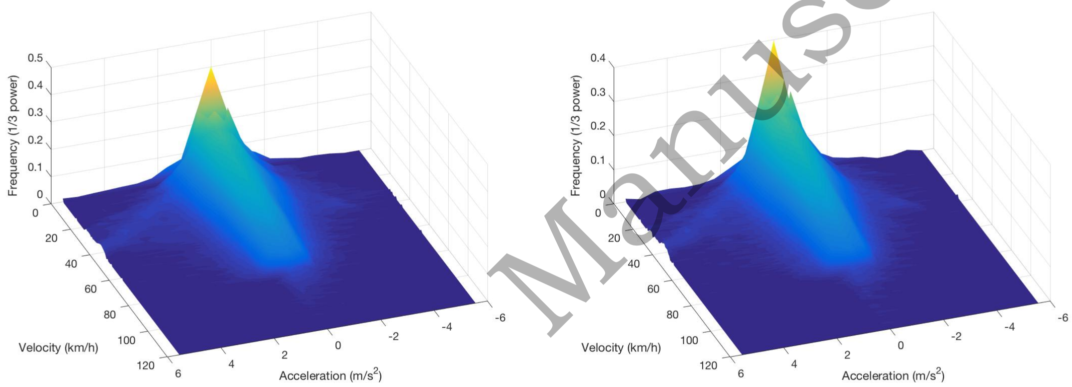
Figure 4 presents the joint distributions of velocity and acceleration of each case. Acceleration ranges from -2 m/s^2 to 2 m/s^2 mostly, while its distribution varies when velocity changes. The most aggressive acceleration and deceleration (e.g., <-4 m/s^2 or >4 m/s^2) appears when velocity ranges from 20 km/h to 40 km/h. Comparing acceleration-velocity distribution among different cases, we also observe the frequency of high-speed segments decreases from highway & freeway, arterial road, to residential road, while the frequency of zero velocity, indicating idling period, increases, which is in consistent with the velocity cluster center distribution shown in Figure 2.



(a) Highway & Freeway, Off-peak Hour (b) Highway & Freeway, Peak Hour



(c) Arterial Road, Off-peak Hour (d) Arterial Road, Peak Hour



(e) Residential Road, Off-peak Hour (f) Residential Road, Peak Hour

Figure 4. Joint Distribution of Acceleration and Velocity for Different Road Types and Traffic Conditions. Note that z-axis indicates the frequency of corresponding segments at 1/3 power

Share of Driving Modes

Figure 5 presents the shares of four driving modes, i.e., acceleration, deceleration, idling and cruise. The shares of acceleration and deceleration range from 14.5% to 18.5% among different road types and traffic conditions. For almost every case except off-peak residential road, the share of deceleration is slightly lower than acceleration, indicating that speeding down is usually more rapid than speeding up, corresponding with Table 2. From highway & freeway, to arterial road and to residential road, the share of idling increases from 36.2%~37.1% to 5.8%~ 8%, while the share of cruise decreases from 57.6%~62.6% to 33.0%~34.3%. This could be explained by the speed limits and traffic lights of different road types, as drivers are more likely to pause at low-speed driving or traffic lights. Besides, traveling during peak hours also accounts for the increasing idling share compared with off-peak hour driving, as moderate driving is more easily to be interrupted by traffic congestion during peak hours.

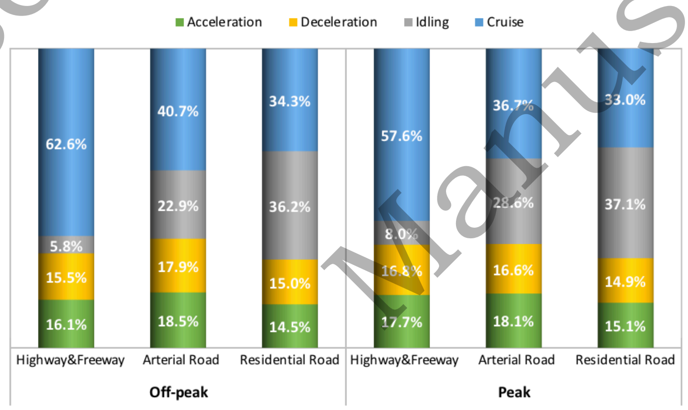
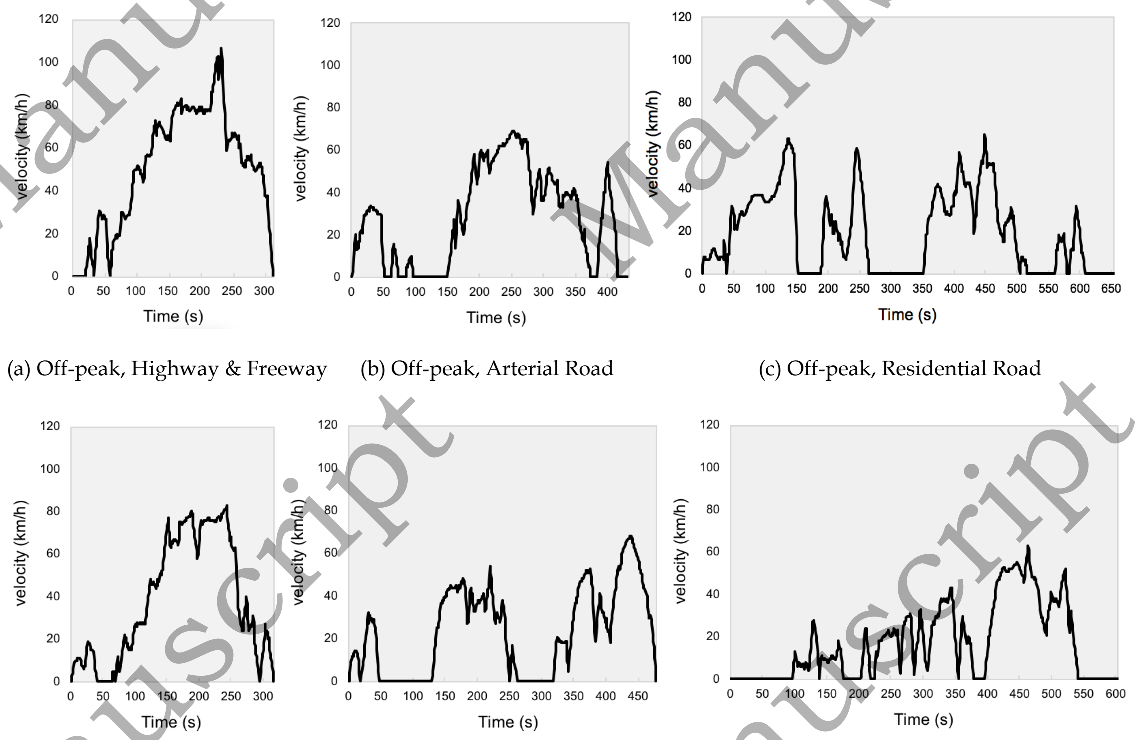


Figure 5. Ratio of Driving Modes for Different Road Types and Traffic Conditions

3.2 Typical driving cycle development

As differential characteristics are revealed while driving under different road types or traffic conditions, six sub-cycles has been developed respectively (Figure 6). They are derived from sample segments in corresponding category and complied with its transition probability matrix. The total length of a typical driving cycle for light-duty passenger vehicles in Beijing is defined by average length of all trips, which is 1398 seconds (i.e. 23 minutes and 18 seconds). As the duration of segments mostly (82.2%) last less than 5 seconds and there are no significant differences among categories, we set time proportion of highway & freeway, arterial road and residential road to be 22.3%, 30.9% and 46.8% respectively for off-peak hours, as well as 22.6%, 34.2% and 43.2% respectively for peak hours. The tolerance of 10 seconds for sub-cycles is adopted in this work.



(a) Off-peak, Highway & Freeway (b) Off-peak, Arterial Road (c) Off-peak, Residential Road

(d) Peak, Highway & Freeway (e) Peak, Arterial Road (f) Peak, Residential Road

Figure 6. Typical driving sub-cycle profiles for Light-duty vehicles in Beijing

Obvious differences emerge from six typical sub-cycles under different situations in Figure 6. Persistent high-speed driving could be achieved on highway and freeway during off-peak hours, which is more difficult for traffic peak hours. More interrupts would happen on residential rather than arterial roads, due to complex road condition and tightened speed limits. Idling time extends and overall velocity descends during rush hour. Parameters of both sample and sub-cycles and errors are presented as Table 3.

Table 3. Parameters and Error for 6 sub-cycles

		Highway&Freeway						Arterial road						Residential road					
		Off-peak			Peak			Off-peak			Peak			Off-peak			Peak		
		Sample	Cycle	Error	Sample	Cycle	Error	Sample	Cycle	Error	Sample	Cycle	Error	Sample	Cycle	Error	Sample	Cycle	Error
Velocity	Average velocity	49.1	50.5	2.9%	39.8	38.3	-3.9%	27.6	28.1	1.8%	23.8	23.4	-1.4%	18.7	18.9	1.0%	17.8	16.8	-5.9%
	Running velocity	52.6	54.4	3.3%	44.0	42.1	-4.2%	37.2	37.4	0.6%	33.7	33.7	0.0%	29.2	28.7	-1.7%	27.3	26.1	-4.6%
	Velocity_90	82.0	80.0	-2.4%	78.0	76.0	-2.6%	59.0	61.3	3.9%	55.0	51.0	-7.3%	50.0	47.0	-6.0%	48.0	47.0	-2.1%
Acceleration	Average acceleration	0.474	0.458	-3.5%	0.483	0.459	-5.0%	0.517	0.498	-3.6%	0.521	0.515	-1.1%	0.526	0.546	3.7%	0.527	0.535	1.5%
	Average deceleration	-0.494	-0.510	3.2%	-0.501	-0.530	5.7%	-0.539	-0.585	8.4%	-0.541	-0.590	9.1%	-0.537	-0.569	5.9%	-0.536	-0.580	8.2%
Ratio of modes	Deceleration	15.5%	15.8%	0.3%	16.8%	19.9%	3.1%	17.9%	20.6%	2.7%	16.6%	17.6%	1.0%	15.0%	17.6%	2.6%	14.9%	16.4%	1.5%
	Acceleration	16.1%	21.2%	5.2%	17.7%	21.8%	4.2%	18.5%	20.4%	1.8%	18.1%	19.9%	1.7%	14.5%	17.3%	2.8%	15.1%	16.7%	1.7%
	Idling	5.8%	6.1%	0.3%	8.0%	7.3%	-0.7%	22.9%	22.5%	-0.4%	28.6%	28.7%	0.1%	36.2%	32.9%	-3.4%	37.1%	34.6%	-2.4%
	Cruise	62.6%	56.9%	-5.7%	57.6%	50.9%	-6.6%	40.7%	36.6%	-4.1%	36.7%	33.9%	-2.8%	34.3%	32.3%	-2.0%	33.0%	32.3%	-0.7%
Ratio of states	State 1	17.0%	16.1%	-0.9%	22.6%	25.0%	2.4%	30.2%	31.9%	1.7%	36.0%	34.9%	-1.0%	45.2%	37.2%	-8.0%	46.3%	42.9%	-3.4%
	State 2	15.2%	15.4%	0.2%	18.1%	19.0%	0.9%	14.9%	11.8%	-3.1%	12.0%	9.8%	-2.2%	11.9%	12.2%	0.3%	14.4%	13.2%	-1.2%
	State 3	13.3%	7.4%	-5.9%	13.3%	12.7%	-0.6%	16.4%	17.6%	1.2%	13.8%	12.6%	-1.2%	11.6%	10.9%	-0.7%	11.9%	14.4%	2.5%
	State 4	18.6%	19.9%	1.3%	13.7%	7.6%	-6.1%	16.7%	15.7%	-1.0%	16.6%	19.5%	2.9%	11.2%	17.4%	6.2%	10.0%	10.3%	0.3%
	State 5	19.1%	16.7%	-2.3%	15.7%	13.9%	-1.7%	12.9%	12.3%	-0.6%	11.6%	17.2%	5.6%	8.7%	12.7%	4.0%	7.8%	9.6%	1.8%
	State 6	12.0%	20.6%	8.6%	11.8%	21.8%	10.0%	6.8%	10.6%	3.9%	7.6%	6.1%	-1.6%	6.4%	8.9%	2.5%	5.6%	9.4%	3.9%
	State 7	4.9%	3.9%	-1.0%	4.9%	0.0%	-4.9%	2.1%	0.0%	-2.1%	2.5%	0.0%	-2.5%	5.1%	0.8%	-4.3%	4.1%	0.2%	-3.9%

Apart from six sub-cycles depicting fine-scale patterns, two comprehensive driving cycle, Beijing Off-peak Driving cycle ("Off-peak Cycle" as follows) and Beijing Peak Driving cycle ("Peak Cycle" as follows), are constructed respectively by appending three sub-cycles on different road types. Thereinto, Off-peak Cycle reflects the overall features about how light-duty vehicles travel in Beijing generally, while Peak Cycle presents the particular focus on congested circumstances during peak hours. Besides, standard driving cycles including NEDC, WLTC, and CLTC-P [CATARC, 2018] are taken into account as well.

As presented in Table 4, congestion impairs the overall speed in Beijing by about 15% and slightly raise the ratio of idle, resulted from the comparison between Off-peak Cycle and Peak Cycle. CLTC shares the closest similarity with Off-peak Cycle in velocity and mode features, while NEDC overvalues the speed generally. The overestimation is expanded drastically by WLTC which underrates idling portion as well, showing the least representativeness of the typical driving pattern in Beijing. On the other hand, the absolute values of acceleration and deceleration in three standard cycles are more or less lower than those in our research, which requires further investigation.

Table 4. Key parameters of comprehensive driving cycles

index	Off-peak Cycle	Peak Cycle	CLTC	NEDC	WLTC
average velocity (km/h)	28.8	23.9	29.0	33.3	46.5
running velocity (km/h)	38.5	33.1	37.7	43.6	53.0
velocity_90 (km/h)	65.0	58.0	63.2	70.0	101.3
average acceleration (m/s^2)	0.506	0.507	0.415	0.480	0.477
average deceleration (m/s^2)	-0.558	-0.570	-0.464	-0.683	-0.514
ratio of deceleration	18.1%	17.6%	17.6%	15.1%	20.3%
ratio of acceleration	19.1%	19.0%	19.5%	19.0%	22.0%
ratio of idling	23.7%	26.4%	22.4%	23.5%	12.2%
ratio of cruise	39.1%	37.1%	40.4%	42.4%	45.6%

4. Vehicle Energy Consumption Simulation

4.1 The Operating Binning Method

The operating binning method was applied to relate the instantaneous vehicle fuel consumption with real-time driving conditions. The vehicle specific power (VSP), a proxy parameter that represents the instantaneous engine power demand, and vehicle velocity were selected as two indicators of vehicle operating conditions. VSP could be calculated based on real-world driving data using Equation (4-1) for light-duty vehicles [Zhang *et al.*, 2014].

$$VSP = v \cdot (1.1 \cdot a + 9.8 \cdot \sin(grade) + 0.132) + 3.02 \times 10^{-4} \cdot v^3 \tag{4-1}$$

where VSP (kW/t) represents vehicle specific power; v (m/s) is instantaneous driving velocity; a (m^2/s) is instantaneous driving acceleration; $grade$ is road inclination, generally regarded as 0 for Beijing in this study.

Driving conditions was divided into 28 operating mode bins defined by VSP and velocity, referring to [Zhang *et al.*, 2014], including a braking bin (Bin 0), an idling bin (Bin 1) and others representing cruise or acceleration. Distance-based fuel consumption (L/km) of a given cycle could be calculated using Equation (4-2),

$$FC = \frac{3600 \sum_i (\overline{FC_i} \cdot P_i)}{v \cdot D} \times 100 \quad (4\text{-}2)$$

where FC (L/100km) represents vehicle fuel consumption per hundred kilometers of a given driving cycle; $\overline{FC_i}$ (g/s) is average fuel consumption rate for a tested vehicle in operating mode bin i; P_i is the time percentage of operating mode bin i for a given driving cycle; v (km/h) is average velocity of driving sequence. D (g/L) is the density of fuel, set as 725 g/L (93# gasoline) in our research.

4.2 Vehicle Energy Consumption Simulation

A series of on-road experiments for vehicle energy consumption were conducted in Beijing using On-Board Diagnostic (OBD) devices [Lu et al., 2018]. Though OBD could record numerous parameters, we particularly extracted the instantaneous fuel consumption, location, velocity and acceleration for further analysis. These experiments involved 12 vehicles whose engine displacement varies from 1.5 L to 3.0 L, with one driver and one passenger on board. Average fuel consumption ($\overline{FC_i}$) for each bin could calculated from OBD profiles, and then used to simulate the distance-based fuel consumption (FC) for various driving cycles based on Equation (2-4).

Table 5 demonstrates the simulated fuel consumption for 12 tested vehicles under various driving cycles. The type-approval fuel consumption values, which were dynamometer testing results under national standard procedures (NEDC-type approval), are also given for comparison. For typical sub-cycles developed in this study, fuel consumption variations among road types and traffic conditions are significant. However, the discrepancies between various sub-cycles are different across vehicles, indicating the heterogeneous influence of driving cycles on specific vehicle models.

Table 5 Distance-based fuel consumption (L/100 km) simulation of 12 light-duty passenger vehicles under typical sub-cycles

Vehicle ID	NEDC-type approval	Off-peak sub-Cycles			Peak sub-Cycles		
		Highway & Freeway	Arterial Road	Residential Road	Highway & Freeway	Arterial Road	Residential Road
1	5.4	5.36	6.31	7.47	5.50	6.85	8.11
2	5.4	5.54	6.61	8.03	5.78	7.22	8.63
3	5.7	5.99	8.14	10.06	6.74	9.06	11.15
4	6.8	7.79	9.20	10.72	8.10	9.90	11.42
5	7.1	7.28	8.71	10.68	7.70	9.54	11.52
6	7.4	8.69	10.73	13.26	9.20	11.66	14.25
7	7.5	8.54	9.72	11.64	8.90	10.43	12.45
8	7.5	6.56	8.87	11.29	7.63	9.69	11.83
9	7.8	8.13	9.49	11.08	8.38	10.11	11.82
10	8.4	8.79	10.23	12.18	9.00	11.02	13.15
11	8.4	8.25	10.09	12.82	8.70	11.32	14.01
12	11	11.02	13.64	16.10	11.94	14.91	17.53
Average	7.37	7.66	9.31	11.28	8.13	10.14	12.16
STD[3]	1.49	1.55	1.84	2.19	1.64	2.01	2.40
Minimum	5.40	5.36	6.31	7.47	5.50	6.85	8.11
Maximum	11.00	11.02	13.64	16.10	11.94	14.91	17.53

The distance-based fuel consumption simulation results of two typical driving cycles developed in this study were demonstrated in Figure 7. Results show that NEDC-type approval fuel consumption median value is lower by 7.83% than simulated fuel consumption under NEDC-on road based on real-world performance (i.e., the NEDC-simulated). This could be explained that dynamometer testing did not capture the extra energy consumption caused by loading, use of accessories, etc., in real-world driving. Fleet average fuel consumption for off-peak driving cycle is 9.28 L/100 km, which is 15.5%, 17.0% and 7.8% higher than average simulated fuel consumption of NEDC, WLTC and CLTC, and that of peak-hour driving cycle is 9.86 L/100 km, which is 22.7%, 24.2% and 14.5% higher than that of NEDC, WLTC and CLTC.

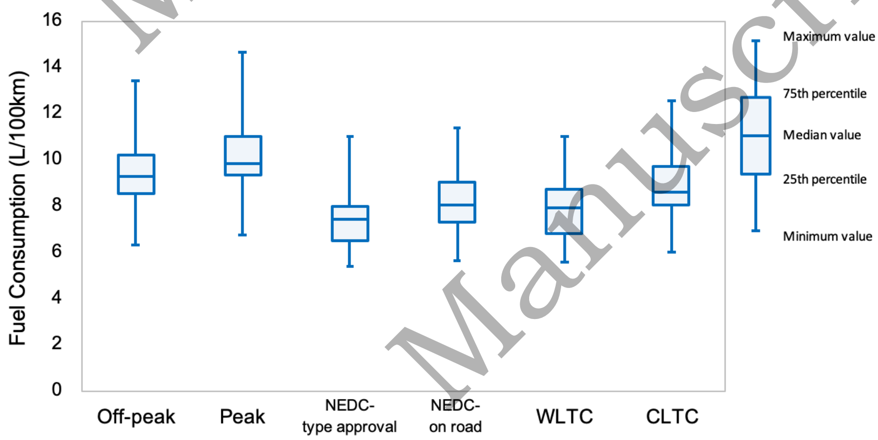


Figure 7 Distance-based fuel consumption of 12 light-duty passenger vehicles under two typical Beijing driving cycles and comparison with other cycles. The NEDC-type approval results are adopted from the official testing results released by China's Ministry of Industry and Information Technology (http://chaxun.miit.gov.cn/asopCmsSearch/)

We also evaluated the vehicle-level discrepancies between various driving cycles with NEDC, as shown in Table 6 and Table 7. For the 12 simulated vehicles, the fuel consumption of off-peak driving cycle is 9.8%~18.6% higher than NEDC-simulated and that of peak hour driving cycle is 19.3%~28.6% higher than NEDC-simulated. And the discrepancies enlarge to 16.2%~44.7% and 24.9~56.2% if compared with type approval (NEDC-type approval) fuel consumptions. On average, the WLTC cycle has slightly lower simulated fuel consumption than NEDC cycle by 3.6%, which is in consistent with previous studies. The WLTC has higher average velocity but is less aggressive than NEDC, thus leading to similar or even lower fuel consumption than the latter.

Table 6. Relative difference of fuel consumption simulation between comprehensive driving cycles and NEDC

	Off-peak – NEDC $_{\text{on road}}$	Peak – NEDC $_{\text{on road}}$	WLTC – NEDC $_{\text{on road}}$	CLTC – NEDC $_{\text{on road}}$
Average	15.0%	24.3%	-3.6%	8.7%
Range	[9.8%, 18.6%]	[19.3%, 28.6%]	[-12.2%, 0.4%]	[5.0%, 10.8%]

Table 7. Relative difference between fuel consumption simulation under comprehensive driving cycles and bench test results

	Off-peak – NEDC $_{\text{type approval}}$	Peak – NEDC $_{\text{type approval}}$	NEDC $_{\text{on road}}$ – NEDC $_{\text{type approval}}$	WLTC – NEDC $_{\text{type approval}}$	CLTC – NEDC $_{\text{type approval}}$
Average	26.2%	36.4%	9.7%	5.6%	19.2%
Range	[16.2%, 44.7%]	[24.9%, 56.2%]	[3.5%, 23.3%]	[-6.5%, 17.5%]	[11.1%, 36.6%]

5. Conclusion

In this research, typical driving cycles were developed using an iterative Markov Chain process for light-duty vehicles in Beijing. This is the first study that constructs driving cycles based on large-scale GPS data from 459 personal-owned passenger vehicles, covering around 3.3 million km. Methodology for data filtration, pre-processing and Markov Chain modeling has been established. The derived two typical driving cycles, i.e., off-peak driving cycle and peak hour driving cycle have average velocity of 28.8 km/h and 23.9 km/h respectively, lower than NEDC, WLTC and CLTC. Vehicle fuel consumption simulation was further conducted for various cycles. Results show that the fuel consumption of off-peak driving cycle is 9.8%~18.6% higher than NEDC-simulated and that of peak hour driving cycle is 19.3%~28.6% higher than NEDC-simulated. And the discrepancies enlarge to 16.2%~44.7% and 24.9~56.2% if compared with type-approval (NEDC-dynamometer) fuel consumptions. This research has explored a practical approach to develop robust driving cycles for a city, which can be employed reliably for exhaust emission assessment and also the field of new energy vehicles in the future.

Acknowledgement

This research was partially supported by Initiative Scientific Research Program, Tsinghua University.

Reference

André M. The ARTEMIS European driving cycles for measuring car pollutant emissions[J]. Science of the total Environment, 2004, 334: 73-84.

André M, Joumard R, Vidon R, et al. Real-world European driving cycles, for measuring pollutant emissions from high-and low-powered cars[J]. Atmospheric Environment, 2006, 40(31): 5944-5953.

Barth M, Younglove T, Wenzel T, et al. Analysis of modal emissions from diverse in-use vehicle fleet[J]. Transportation Research Record: Journal of the Transportation Research Board, 1997 (1587): 73-84.

Beijing Traffic Management Bureau (BTMB), 2017. National newly-increased vehicles reaches 9.38 million, with new drivers at 13.81 million in first half year 2017. Available at http://www.bjjtgl.gov.cn/jgj/95332/531059/index.html [in Chinese].

Bishop J D K, Axon C J, McCulloch M D. A robust, data-driven methodology for real-world driving cycle development[J]. Transportation Research Part D: Transport and Environment, 2012, 17(5): 389-397.

China Automotive Technology & Research Center (CATARC), 2018. China Automotive Driving Cycle Part I: Light-duty vehicles (for public comment). Available at http://www.catarc.org.cn/StandardRevision/detail/1147.html [in Chinese].

Gong Q, Midlam-Mohler S, Marano V, et al. An iterative markov chain approach for generating vehicle driving cycles[J]. SAE International Journal of Engines, 2011, 4(1): 1035-1045.

He X, Wu Y, Zhang S, et al. Individual trip chain distributions for passenger cars: Implications for market acceptance of battery electric vehicles and energy consumption by plug-in hybrid electric vehicles[J]. Applied Energy, 2016, 180: 650-660.

Hung W T, Tong H Y, Lee C P, et al. Development of a practical driving cycle construction methodology: A case study in Hong Kong[J]. Transportation Research Part D: Transport and Environment, 2007, 12(2): 115-128.

Kamble S H, Mathew T V, Sharma G K. Development of real-world driving cycle: Case study of Pune, India[J]. Transportation Research Part D: Transport and Environment, 2009, 14(2): 132-140.

Lin J, Niemeier D A. An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle[J]. Atmospheric Environment, 2002, 36(38): 5759-5770.

Lin J, Niemeier D A. Estimating regional air quality vehicle emission inventories: constructing robust driving cycles[J]. Transportation Science, 2003, 37(3): 330-346.

Lu S, Wu Y, Zhang S J, et al. 2018. Characterizing real world fuel consumption for light-duty passenger vehicles by using on-board diagnostic (OBD) system [J]. Acta Scientiae Circumstantiae, 38 (5): 1783-1790

Ministry of Environmental Protection (MEP) of China, 2015. Major city pollution sources identified. Available at http://www.chinadaily.com.cn/hqgj/jryw/2015-04-01/content_13477162.html [in Chinese].

Ntziachristos L, Mellios G, Tsokolis D, et al. In-use vs. type-approval fuel consumption of current passenger cars in Europe[J]. Energy Policy, 2014, 67: 403-411.

Nyberg P, Frisk E, Nielsen L. Generation of equivalent driving cycles using Markov chains and mean tractive force components[J]. IFAC Proceedings Volumes, 2014, 47(3): 8787-8792.

Rubino L, Bonnel P, Hummel R, et al. Mobile measurement of pollutant emissions and fuel consumption of road vehicles in real-world driving situations using portable emission measurement systems (PEMS)[J]. Final report. Eur. Commission, Ispra, Italy (61 pp. http://ec. europa. eu/clima/policies/transport/vehicles/docs/2007_pems_jrc_413194_en. pdf (accessed Sept. 09, 2013)), 2007.

Shi S, Lin N, Zhang Y, et al. Research on Markov property analysis of driving cycles and its application[J]. Transportation Research Part D: Transport and Environment, 2016, 47: 171-181.

Wang Q, Huo H, He K, et al. Characterization of vehicle driving patterns and development of driving cycles in Chinese cities[J]. Transportation research part D: transport and environment, 2008, 13(5): 289-297.

Wu Y, Zhang S, Hao J, et al. On-road vehicle emissions and their control in China: A review and outlook[J]. Science of The Total Environment, 2017, 574: 332-349.

Zhang S, Wu Y, Liu H, et al. Real-world fuel consumption and CO2 (carbon dioxide) emissions by driving conditions for light-duty passenger vehicles in China[J]. Energy, 2014, 69: 247-257.