



# Prediction of Electric Vehicle Energy Consumption in an Intelligent and Connected Environment

Qingchao LIU<sup>1</sup>, Fenxia GAO<sup>2</sup>, Jingya ZHAO<sup>3</sup>, Weiqi ZHOU<sup>4</sup>

Original Scientific Paper  
Submitted: 1 Feb. 2023  
Accepted: 19 May 2023

<sup>1</sup> lqc@ujs.edu.cn, Automotive Engineering Research Institute, Jiangsu University; School of Mechanical and Aerospace Engineering, Nanyang Technological University

<sup>2</sup> scisci@126.com, Automotive Engineering Research Institute, Jiangsu University; Jiangsu University Research Institute of Engineering Technology

<sup>3</sup> zhaojingya26@163.com, Automotive Engineering Research Institute, Jiangsu University; Jiangsu University Research Institute of Engineering Technology

<sup>4</sup> zwq@ujs.edu.cn, Automotive Engineering Research Institute, Jiangsu University; Jiangsu University Research Institute of Engineering Technology



This work is licensed under a Creative Commons Attribution 4.0 International License

Publisher:  
Faculty of Transport and Traffic Sciences,  
University of Zagreb

## ABSTRACT

Accurate energy consumption prediction is essential for improving the driving experience. In the urban road scenario, we discussed the influencing factors of energy consumption and divided the modes from various perspectives. The differences in energy consumption characteristics and distribution laws for electric vehicles using the IDM and CACC car-following models under different traffic flows are compared. An energy consumption prediction framework based on the LightGBM model is proposed. According to the study, driving range, acceleration, accelerating time, decelerating time and cruising time all significantly impact the overall energy consumption of electric vehicles. There are apparent differences in energy consumption characteristics and distribution laws under different traffic flows: average energy consumption is lower under low flow and increased under high flow. The CACC-electric vehicles consume more energy in low flow than IDM-electric vehicles. Under high flow, the opposite is true. The results show that the proposed framework has a high accuracy: the MAPE based on IDM datasets is 3.45% and the RMSE is 0.039 kWh; the MAPE based on CACC datasets is 5.57% and the RMSE is 0.042 kWh. The MAPE and RMSE are reduced by 33.7% and 50.6% (maximum extent) compared to the best comparison algorithm.

## KEYWORDS

electric vehicle; energy consumption prediction; green travel; signalised intersection.

## 1. BACKGROUND

Transportation electrification has been widely recognised as a necessary mean to improve energy efficiency and reduce emissions [1]. In the intelligent and connected environment, accurate and real-time energy consumption (EC) prediction for electric vehicles (EVs) is essential to improve the travel service experience, as well as provide support for optimising battery design [2], planning an energy-saving travel route [3] and improving charging infrastructure [4]. Considering that the EC characteristics and influencing factors of EVs are discrepant from fuel vehicles due to the apparent differences in the power system, and the urban road environment is relatively complicated, it is inevitable to analyse the influencing factors of EC and predict EC for EVs on urban roads. Most of the previous related studies considered the influencing factors such as weather, microscopic operating parameters, road conditions and vehicle and driver characteristics while ignoring the direct/indirect impact of vehicle interaction and signal control on EC under different traffic flows. In addition, the existing studies mainly focused on manual-driving vehicles. There is insufficient research of the EC law of connected automated vehicles (CAV, e.g. the car-following model for cooperative adaptive cruise control). These factors motivate us to analyse the EC law of non-CAV and CAV in combination with driving conditions and signal control under different traffic flows and predict the EC.

## 2. LITERATURE REVIEW

The battery capacity shortens the endurance mileage of EVs due to the technology's limitations. Hence, it is essential to focus on EC in real-time. EC is attributable to multiple influencing factors and changes unceasingly in a realistic driving environment. Therefore, the problem of EC prediction becomes relatively complicated. The influencing factors of EC roughly come from three categories based on the analysis of previous research: vehicle, environment and driver. Vehicle-related factors include vehicle speed, acceleration, vehicle-specific power and battery state of charge (SOC). Wu et al. [5] proposed a power estimation model based on longitudinal dynamics (LDM) to estimate EVs' instantaneous power and EC. Wang et al. [6] constructed an online EC prediction algorithm by combining LDM and other factors (e.g. driving behaviour and traffic states). The results showed that the mean absolute percentage error between the actual EC values and the online prediction results for each test was within 5%. Similar research has also been implemented. The EC prediction model for EVs developed by Fiori et al. [7] and Yang et al. [8] took acceleration into account. They found that the EVs with high EC usually possess high speed, significant acceleration, long space headway and short time headway. Environmental factors comprise ambient temperature road conditions. For example, Liu et al. [9] raised and further corrected an EC model combined with ambient temperature, vehicle auxiliary load analysis and other factors. Ahn et al. [10] quantified the impact of intersection control methods (e.g. roundabouts, traffic signals, two-way parking control) on EC according to the specific driving scenarios on the urban roads, and systematically studied the EC modes of EVs. Additionally, Wang et al. [11] developed an EC prediction algorithm for the future travel of EVs. The algorithm, composed of an offline and online algorithm, comprehensively considers diverse factors, including weather and traffic situations, to make future driving plans for electric vehicle (EV) drivers and make online adjustments. The EC can also be affected by driver-related factors (e.g. driving style). For instance, aiming at the problems of short driving range and unprecise remaining range prediction of EVs, Guo et al. [12] used a linear method-based drive cycle prediction model for real-world conditions prediction and subsequently proposed an EC prediction method integrating road information and driving style optimisation. Based on the analysis of influencing factors, most of the above studies have applied model-driven EC prediction methods, which rely on parameter assumptions and specified conditions, resulting in the difficulty of satisfying EVs' actual driving performance requirements. In addition, the traffic environment near signalised intersections is more complicated by contrast. Thus, the distribution law of EC needs to be analysed in combination with traffic signals and driving conditions.

Due to the rapid development of the Internet of Vehicles (IoV), vehicle operating data and traffic information can be obtained in real-time. Extensive practical driving data can be acquired based on various statistics and machine learning algorithms to predict the EC of EVs via the data-driven method. For example, Yi et al. [13] introduced a data-driven approach to establish a random EC model of EVs with a two-dimensional grid (average speed and ambient temperature), the accuracy of which depends on the number of available data samples collected. Hence, plenty of valuable data can ease the model precision in the IoV environment. Confronting the problem of range anxiety of drivers, De et al. [14] adopted the neural network to predict the unknown microscopic driving parameters of the roadway link before departure and then utilised multiple linear regression to predict EC with a mean absolute error of 12–14%. Modi et al. [15] implemented a method based on deep convolution neural network to predict the real-time EC of EVs, which can be carried out to strictly calculate the remaining ranges, thereby alleviating drivers' range anxiety. Moreover, Madhusudhanan and Na [16] innovatively proposed a cruise control system (CCS) to increase the endurance mileage for EVs. The CCS includes an EC model exploited based on actual data, reducing the EC of EVs driving on urban roads and highway travel by 36.6% and 15.4%, respectively. Combining the prediction of real-world driving situations, Zhang et al. [17] and Yao et al. [18] separately raised a machine learning-based EC prediction method, which achieved anticipatory effects compared with traditional ones. Furthermore, Fukushima et al. [19] developed a machine learning-based multi-vehicle mileage prediction system applicable for EVs driving in the highway environment. They applied data-driven models to predict the EC for EVs and recommended suitable charging sites for highway drivers. The above explained researches show that machine learning algorithms perform better under complicated real-world scenarios and can employ massive driving data to improve prediction accuracy. A light gradient boosting machine (LightGBM) [20] supports efficient parallel training with low memory consumption and high precision. It can rapidly process massive data and performs efficiently in time series prediction. In addition, CAVs can obtain more information than non-CAVs in the IoV environment, so its EC law is worth exploring.

Most existing studies consider vehicle characteristics and microscopic operating parameters on EC. However, on the one hand, the interaction between vehicles will affect driving behaviour, especially when the traffic flow is high. On the other hand, due to signal control, the EV is in different driving modes (such as accelerating, decelerating, cruising and idling) near the intersection. Moreover, the delay of EVs at signalised intersections is distinct under different traffic flows. Due to secondary queuing, the delay time is significantly prolonged under high traffic flow. These factors are closely related to EC. In addition, vehicles can realise cooperative driving in the IoV environment. Compared with non-CAVs, CAVs have distinct EC laws due to their communication advantages and different car following rules, which will also have an impact on EC prediction. The existing research rarely analyses and compares the EC law of non-CAVs and CAVs at a signalised intersection, especially considering the influence of vehicle interaction and signal control under different traffic flows. The data-driven method has been widely applied [21, 22]. In particular, LightGBM is usually utilised in traffic flow prediction [23], risk factors analysis on crashes [24], lane-changing risk prediction [25] and other fields in the intelligent transportation system, but it has not been used in EC prediction.

Therefore, this paper compared the EC and its distribution for EVs with the intelligent driving model (IDM-EVs) and the cooperative adaptive cruise control (CACC-EVs) system driving at signalised intersections under different traffic flows, taking into account vehicle interactions and signal control factors, and divided the multi-mode EC from perspectives of traffic flow and vehicles. An EC prediction framework based on the machine learning algorithm – LightGBM was developed afterward. The data used in this study and the proposed framework can be further applied to evaluate and optimise EVs' EC and endurance mileage, which is crucial for alleviating “range anxiety”.

The main contributions of this paper are: (1) The comparative experimental method was adopted to study the differences in EC distribution between IDM-EVs and CACC-EVs during different driving stages. The braking energy regeneration of EVs was considered in the process. (2) A multi-mode EC division method was introduced, and the EC in different modes was studied from perspectives of traffic flow and vehicles to explore the EC law in depth. (3) An EC prediction framework for EVs based on the LightGBM algorithm was proposed, comprehensively considering various vital influencing factors in practical applications to improve prediction performance. The remainder of this paper is organised as follows; influencing factors of EC are elaborated in Section 3. Section 4 introduces the EC prediction framework and compares the prediction results with the comparison algorithm. Finally, Section 5 summarises the main conclusions of this paper and discusses further prospects.

### 3. ELECTRIC VEHICLE ENERGY CONSUMPTION INFLUENCING FACTORS

Unstable EC will occur and sometimes fluctuate significantly during the driving process of EVs. As shown in *Figure 1*, the energy consumption rate (ECR) varies within a specific range during the driving process, mainly due to the influence of various factors on EC. In addition, energy regeneration (ER) cannot be neglected in the whole driving process. The battery system of an IDM-EV recovers an average of 9.87% of the energy per kilometre, and the battery system of a CACC-EV recovers an average of 12.50% of the energy per kilometre.

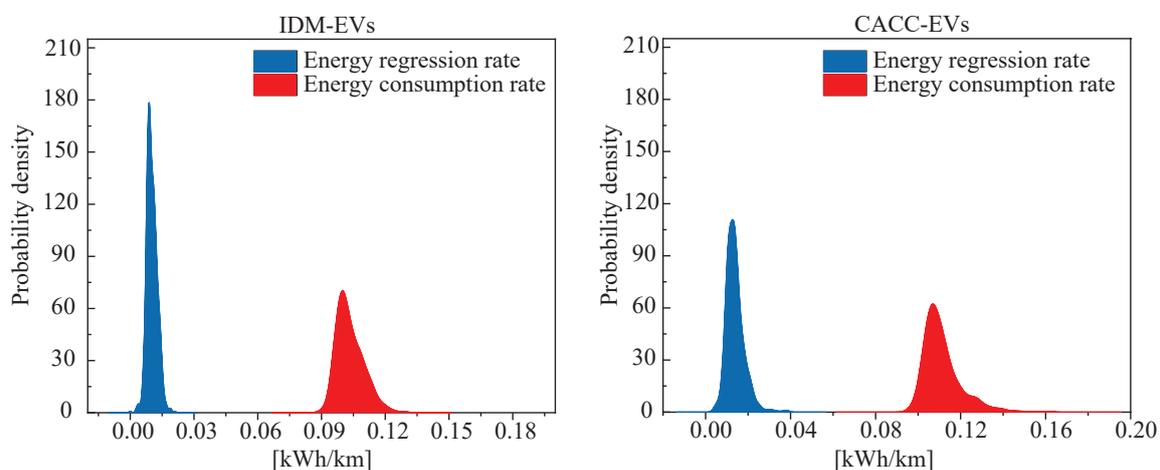


Figure 1 – The distribution of ECR and ERR of EVs

The influence of acceleration and speed on EC has attracted much attention in previous studies. As shown in Figure 2, acceleration shows an apparent positive correlation with ECR and energy regeneration rate (ERR), which can reach a high level at high acceleration/deceleration stages. The relationship will slightly differ due to the distinct car-following models. Figure 3 shows the relationship between cruising speed and ECR. ECR decreases with the increase of cruising speed before the cruising speed reaches 24 km/h and then increases gradually. Multiple variables simultaneously influence EC, but it is impossible to comprehensively explain their coupling effect on the EC of EVs by theoretical methods. This section analyses the EC of EVs driving in different areas and summarises the EC laws and the influencing factors by mining the EC modes.

### 3.1 Experimental settings

This paper validates the proposed method for predicting EC in an urban road environment using a simulation of urban traffic (SUMO). Based on C++ and Python, SUMO is a microscopic, continuous spatial discrete traffic simulation software that simulates various road traffic networks. In SUMO, users can set a variety of contents, including signal intersections, traffic flow, vehicle parameters and driving routes. The XML file generated at the end of the simulation records the entire road network information. The simulated road network is shown in Figure 4, including five signalised intersections and 16 road sections. Aiming to study the effectiveness of the proposed EC prediction method under different traffic demands, the traffic flows at each entrance are set to 600 vehs/h and 1,200 vehs/h, respectively. Each entrance is 2 km long and contains three motor vehicle lanes: straight, left-turn and straight-right mixed lanes. The total mileage of EVs can reach up to 12 km. Each intersection is controlled by a conventional four-phase timing signal with a cycle time of 132 s, of which the straight (and right-turn) phase is 35 s, and the left-turn phase is 25 s. The total simulation time is 3,600 s, and the vehicle’s free-flow speed is 50 km/h. In addition, the vehicle length is 5 m. According to Fiori et al. [26] and considering the energy loss generated by EVs during driving, the recuperation efficiency is set to 0.8. Other detailed specifications are shown in Table 1.

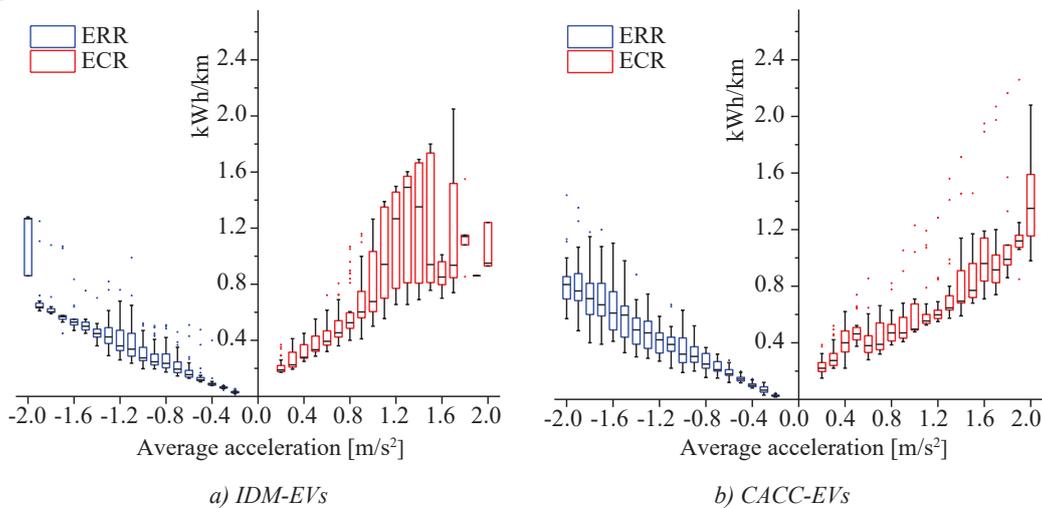


Figure 2 – Relationship of average acceleration with ECR and ERR

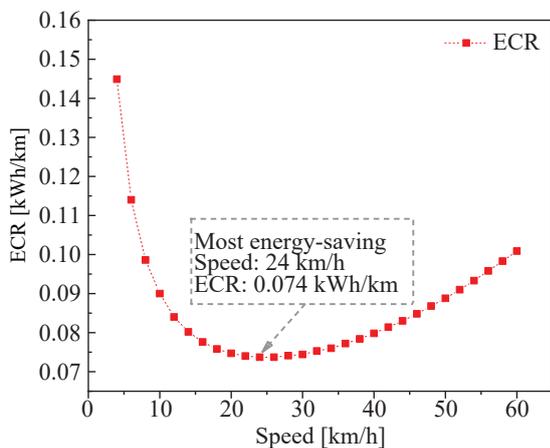


Figure 3 – Relationship of cruising speed with ECR

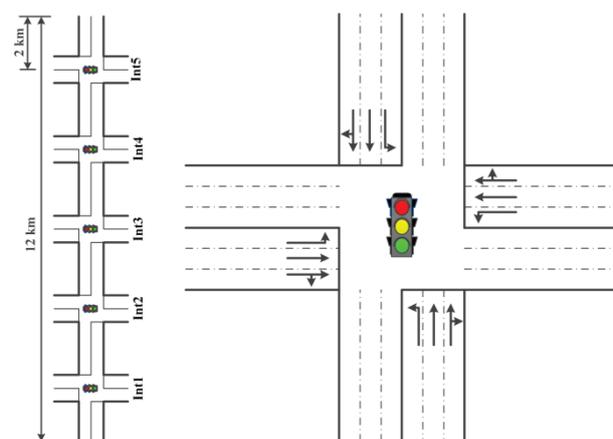


Figure 4 – Road layout

Table 1 – The specifications of the studied EVs

Parameters	Value
Maximum acceleration	3 m/s <sup>2</sup>
Maximum deceleration	3 m/s <sup>2</sup>
Maximum velocity	70 km/h
Car-following model	IDM; CACC
Lane-changing model	LC2013
tau	1.4 (When the car-following model is IDM); 0.6 (When the car-following model is CACC)
Battery capacity	20 kWh
Maximum power	85 kW
Vehicle mass	1,500 kg
Recuperation efficiency	0.8

### 3.2 Energy consumption analysis by areas

Due to EVs near the intersection being controlled by signals and producing more complex driving behaviours, which results in differences in the distribution of EC in different areas (e.g. between the entrance road and the inside of the intersection), it is necessary to analyse the EC by areas. The distribution of EC near the signalised intersection is directly related to the driving state of EVs, except for the traffic flow. EVs on the entrance/exit lane beyond a certain distance from the intersection are less affected by the signal light and usually maintain the cruising state so the EC tends to be stable. Due to signal control, EVs start braking within a certain distance from the intersection on the entrance road, resulting in high energy recovery. The EC of EVs lining up on the entrance road is low and stable. EVs restart at the end of the red light and accelerate inside the intersection and near the exit road, resulting in high EC. To further clarify the driving states of EVs for data statistics, a typical travel segment is taken as an example, as shown in Figure 5. According to the variation range of speed and acceleration, each travel segment is divided into five driving states: starting, accelerating, cruising, decelerating and idling. The division criteria are determined based on statistics of all driving data, as shown in Table 2. It is noted that the starting state is the accelerating process from the vehicle’s stationary to the first cruising or decelerating state [17].

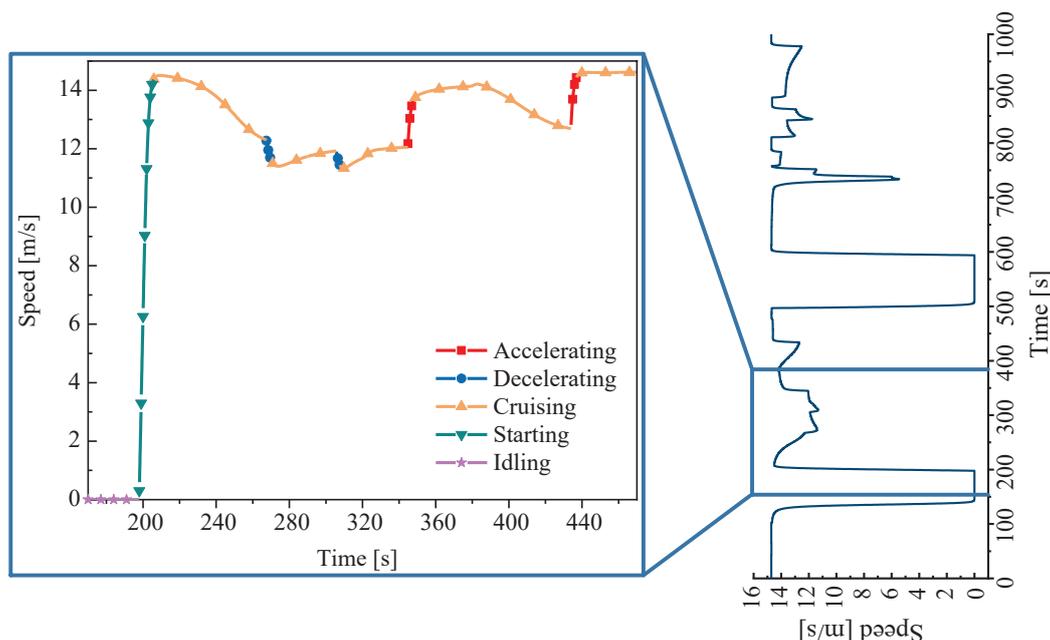


Figure 5 – Example of driving states

Table 2 – Division principle of driving states

Driving states	Division principle	
	Speed ( $v$ , m/s)	Acceleration ( $a$ , m/s <sup>2</sup> )
Accelerating	$v > 1.11$	$a > 0.15$
Decelerating	$v > 1.11$	$a < -0.15$
Cruising	$v > 1.11$	$-0.15 < a < 0.15$
Idling	$v \geq 1.11$	--
Starting	$v \geq 0$	$a > 0.15$

### 3.3 Energy consumption mode mining

#### Energy consumption mode division

For exploring the EC laws of EVs in each driving state, this subsection discusses two perspectives of traffic flow and vehicles, divides the EC modes according to driving conditions and summarises the influencing factors of EC. From the traffic flow perspective, Figure 6 shows the method of dividing EC modes. Based on the trajectory division, in Figure 6a one of the typical trajectory lines is taken as a specific illustration.  $S_1$  is the cruising stage (green segment), in which EC is positively correlated with vehicle speed, driving time and distance, etc. and its EC mode is noted as  $P_1$ .  $S_2$  is the stage (red segment) when the vehicles on the entrance road start to brake and queue due to the red light. EVs begin to recover energy when the deceleration reaches a specific value, the peak value of which has a great relationship with the magnitude of deceleration. In addition, the behaviours such as stop-and-go and changing lanes frequently (related to the law of car-following) also increase the EC. The EC mode at this stage is recorded as  $P_2$ .  $S_3$  is the queuing stage (blue segment). The EVs produce a low and stable EC to ensure the operation, and their EC mode is recorded as  $P_3$ . The total EC during this stage is related to the queueing delay.  $S_4$  is the stage (brown segment) when the EVs start and accelerate to a constant speed after the green light. The fleet begins dissipating and EC increases in the stage. The EC mode is recorded as  $P_4$ . Figure 6b shows the EC distribution states at  $P_1, P_2, P_3$  and  $P_4$ .

To further analyse EC for each mode, an EV controlled to drive 1 km (including a signalised intersection) was selected as the research object under low and high traffic flow, as shown in Figure 7 (the negative value indicates energy recovery). The total EC of CACC-EVs is 13.2% higher than IDM-EVs under low traffic flow and 22.7% lower than IDM-EVs under high traffic flow. After analysis, it is difficult to form a fleet among the CACC-EVs under low traffic flow, and the car-following advantage is lost when losing the “guidance” of the leader. However, CACC-EVs can readily form a fleet under high traffic flow, which can significantly optimise driving trajectories to reduce the EC. Moreover, the results show that the higher the traffic flow in a particular range is, the more energy-saving the CACC-EVs can be.

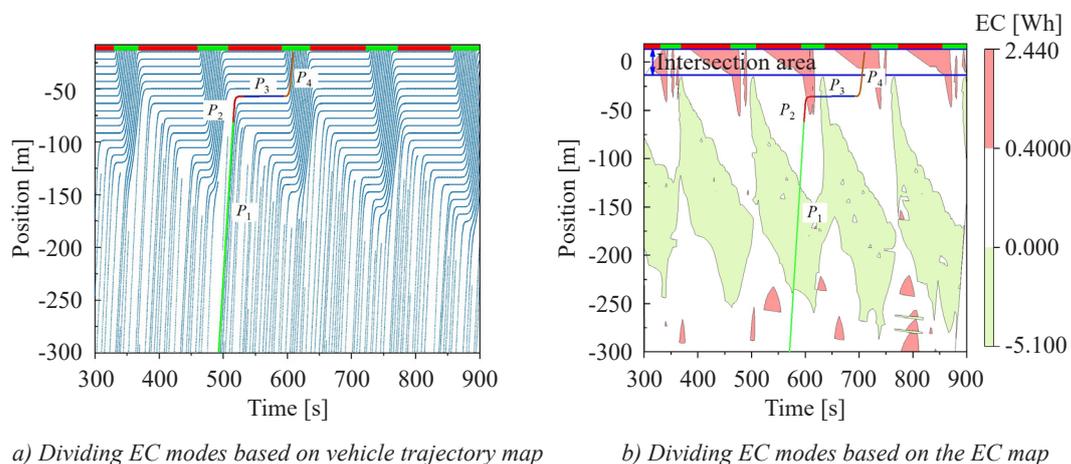


Figure 6 – Multi-mode division of EC

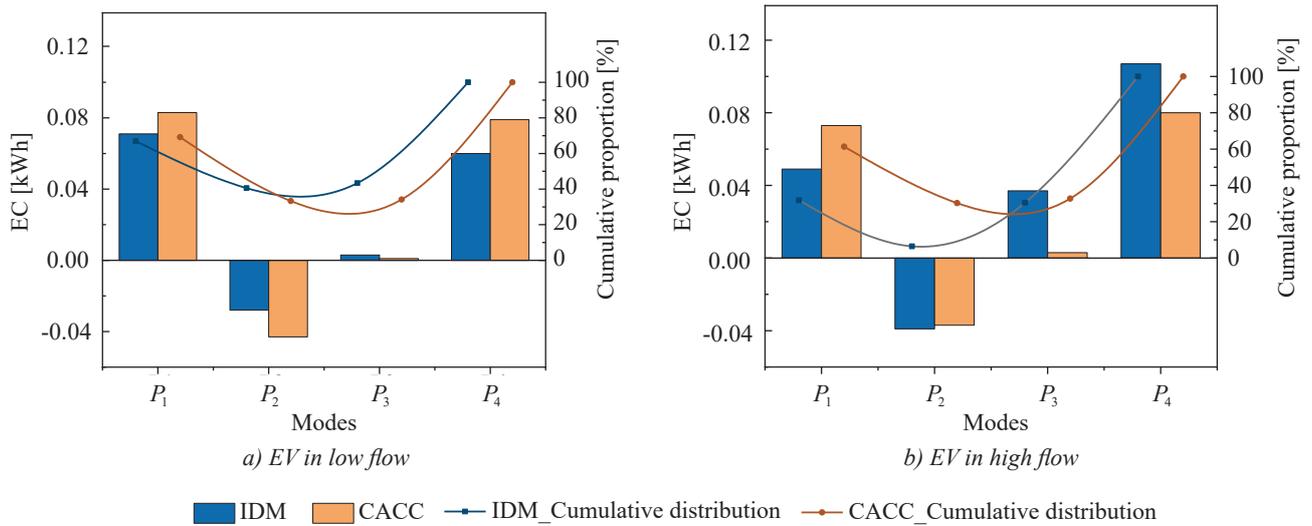


Figure 7 – EC in different modes

From the perspective of vehicles, the driving state of EVs during travel continues to change according to the traffic environment, and it is necessary to further analyse its EC law from the microscopic level. At signalised intersections, EVs generally experience several processes of cruising–decelerating–queuing–starting and accelerating–cruising when they are queuing due to the red light, which is denoted as  $s_1$  (cruising),  $s_2$  (braking),  $s_3$  (queuing),  $s_4$  (accelerating). As shown in Figure 8, both the  $s_1$  and  $s_3$  stages produce stable EC. The EC in the  $s_1$  stage is mainly related to the driving speed and distance (or time). In contrast, the EC in the  $s_3$  stage is related

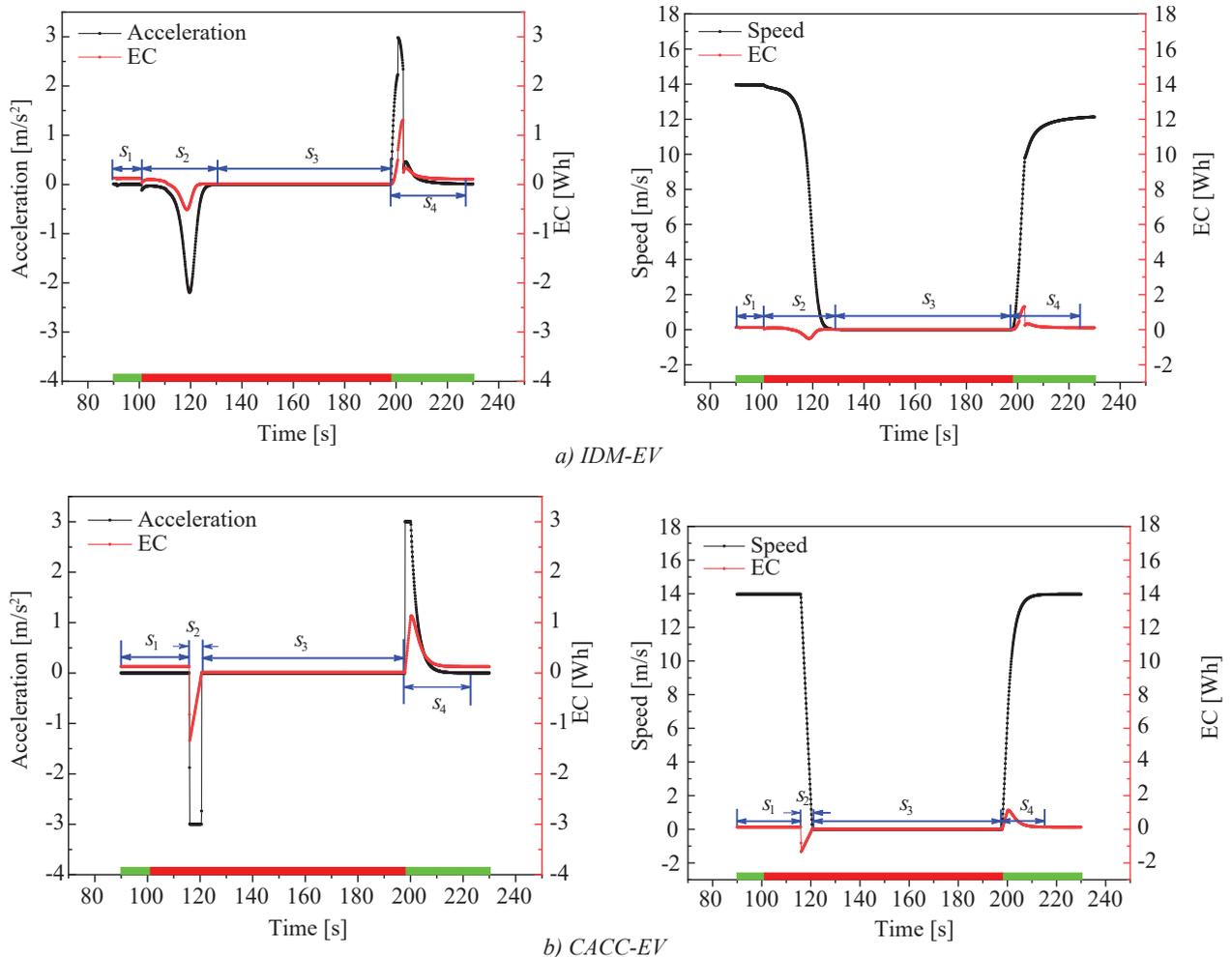


Figure 8 – EC of a single vehicle

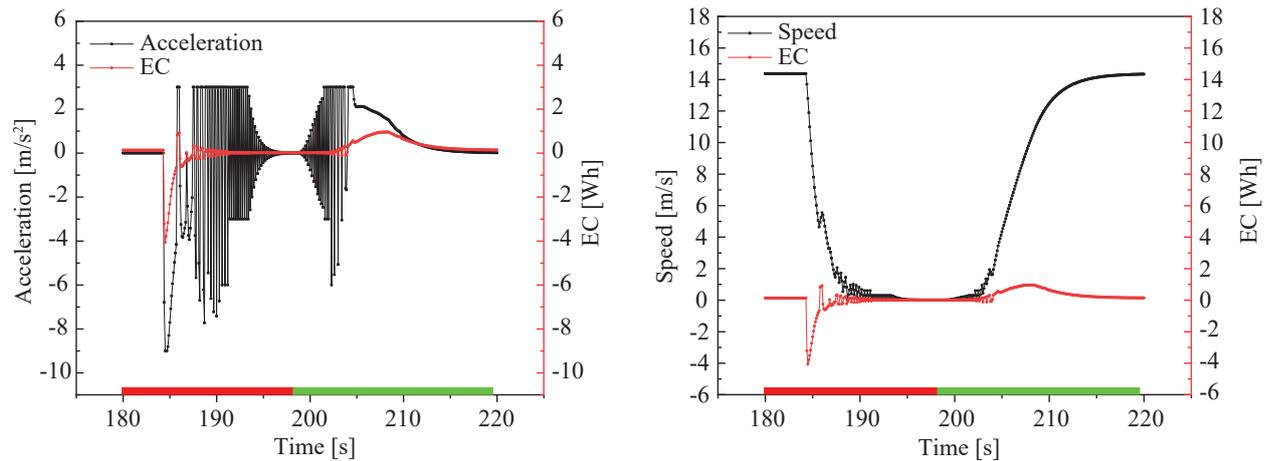


Figure 9 – EC of a following CACC-EV in a fleet

primarily to queuing delay. The EV reclaims part of the energy during the  $s_2$  stage and consumes plenty of energy in the  $s_4$  stage. The period of the two stages for the CACC-EV is usually shorter than the IDM-EV. The correlation curves of the two EVs in Figure 8 show noticeable differences mainly caused by car-following laws: IDM-EVs prefer to gradually complete driving tasks, while CACC-EVs pay more attention to efficiency. For example, the acceleration of the former gradually increases when the vehicle starts, while the latter tends to take maximum acceleration.

The study found that when CACC-EVs are queuing, the EC law of the leader is different from the following vehicles, as shown in Figures 8 and 9. The CACC-EV in Figure 8 is the fleet leader, while the identical EV in Figure 9 is one of its followers. During the  $s_2$  stage, the leader gradually decelerates and mainly recovers energy. Nevertheless, the continuous changes in the acceleration of the follower cause the speed to increase and decrease, and the EC fluctuates consequently. However, the overall recovered energy is more than the consumed amounts. A similar phenomenon also exists in the  $s_4$  stage. It should be noted that if some vehicles do not line up at the intersection, there is no  $s_3$  stage. In addition, the vehicle possesses lateral and longitudinal acceleration during left-turning, and its EC is slightly different from the straight-going one. The main difference is reflected inside the intersection, corresponding to the  $s_4$  phase, as shown in Figure 10. There are significant fluctuations in the EC of the IDM-EV during left-turning, especially for the following vehicle, which is mainly related to the changes in its acceleration. However, the EC of the straight-going vehicle (Figure 8) shows a process of increasing and then decreasing to flatness in this stage. The CACC-EV leader and the following showed visible fluctuations in EC during left-turning, and the maximum EC was higher than the corresponding straight-going vehicle (Figures 8 and 9).

#### Multi-mode energy consumption laws

During the  $S_1$  stage (the EC mode is  $P_1$ ), the vehicle is cruising. Its EC is mainly related to vehicle speed and distance (or driving time). During the  $S_2$  stage (the EC mode is  $P_2$ ), the EV starts to brake and queue, and the recovered energy is closely correlated with the deceleration. During the  $S_3$  stage (the EC mode is  $P_3$ ), the vehicle is idling, the EC of which is mainly related to queuing delay. Moreover, during the  $S_4$  stage (the EC mode is  $P_4$ ), the vehicle starts and further accelerates, the EC of which is relatively correlative with the acceleration.

There are differences in the EC laws for each mode of IDM-EVs and CACC-EVs. Compared to IDM-EVs, this study found that CACC-EVs consume more energy in the  $S_1$  and  $S_4$  stage, more energy is recovered in the  $S_2$  stage, and less EC in the  $S_3$  stage under the same traffic flow. In addition, when driving the same mileage near the signalised intersection, compared to the IDM-EV, the total EC of the CACC-EV is higher at low traffic flow and lower at high traffic flow, which means that the CACC-EV saves more energy and possesses higher driving efficiency under high traffic flow.

Based on the EC modes analysis, the influencing factors of EC can be summarised into two aspects; one is related to vehicle travel, including car-following rules and driving behaviours; the other is related to signal control, including queuing caused by red lights. Specifically, the analysis from a temporal perspective includes factors such as travel time, delay, accelerating and decelerating time, cruising time, and idling time. The analysis from a spatial perspective includes factors such as driving distance, acceleration and deceleration, the number of lane changes and stops, and car-following behaviours between EVs.

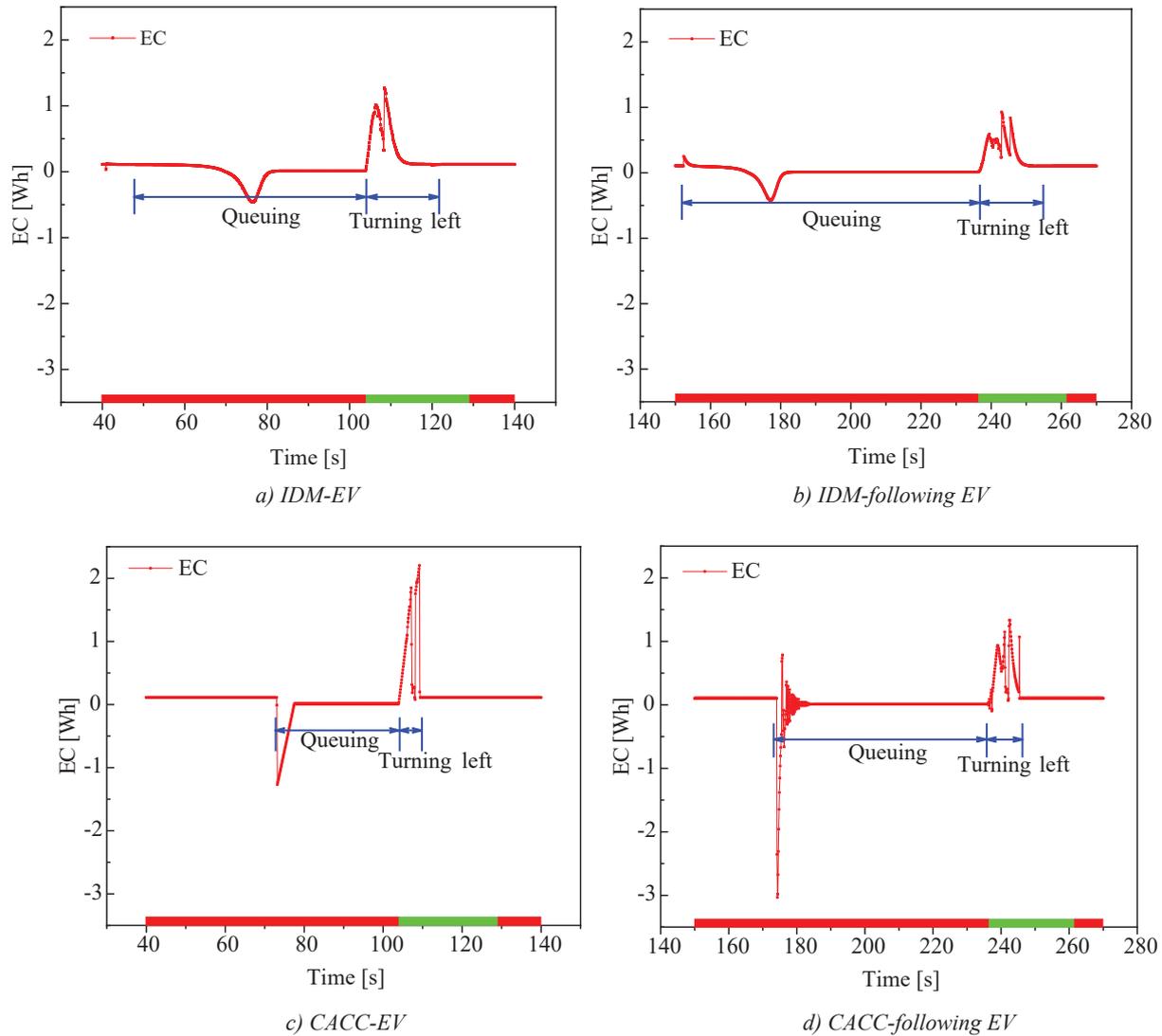


Figure 10 – EC of a left-turn vehicle

#### 4. ENERGY CONSUMPTION PREDICTION CONSIDERING THE INFLUENCE OF MULTIPLE FACTORS

Section 3 analysed the EC law of EVs from the perspective of traffic flow and vehicles and summarised the influencing factors of EC. It is challenging to quantify the effects of some influencing factors on EC since they couple in nonlinear ways. Machine learning algorithms can efficiently model the nonlinear relationship between multiple influencing factors and EC. This subsection proposes a LightGBM-based EC prediction framework that integrates the extracted features related to vehicle travel and signal control.

##### 4.1 Datasets

The dataset is divided into IDM and CACC datasets according to vehicle type, each containing driving segments of 1,000 EVs in different traffic flows. Each driving segment is in the range of 4–12 km in length and contains  $S_1$ – $S_4$  driving stages. The training and test sets are split according to 4:1, and the constructed datasets is shown in Tables 3 and 4. Table 3 relates to vehicle travel and includes statistical information for each segment, such as driving time, driving range, acceleration and deceleration and average travel speed. The classification standard of the accelerating, decelerating, cruising, idling and starting of the EV in Table 3 is recorded in Table 2. Table 4 relates to the signal control and includes the signal states corresponding to the arrival of vehicles at the intersection in each segment. The signal phases are divided according to 1:2:1. For example, the green light for straight-going and right-turn is 35 s, and its early, mid and late stages are 8.75 s, 17.5 s and 8.75 s, respectively.

Table 3 – Vehicle-related variables

Input/output	Parameters	Unit	Training			...	Test		
			veh <sub>1</sub>	veh <sub>2</sub>	veh <sub>i</sub>		veh <sub>1</sub>	veh <sub>2</sub>	veh <sub>j</sub>
Input	Driving time	s	1,031.3	1,155.1	395.6	...	301.9	505.0	714.0
	Travel delay	s	216.6	316.1	88.7	...	7.6	74.4	143.9
	Waiting time	s	144.6	263.8	73.0	...	0.0	38.2	91.6
	Driving range	km	12.0	10.0	4.0	...	4.0	6.0	8.0
	95 % quantile of acceleration	m/s <sup>2</sup>	2.99	2.98	2.94	...	1.06	2.80	2.35
	5 % quantile of deceleration	m/s <sup>2</sup>	-2.46	-2.00	-2.27	...	-1.17	-2.18	-2.19
	Average travel speed	m/s	11.63	8.64	10.09	...	13.23	11.86	11.19
	Waiting count	times	2	5	1	...	0	3	3
	Lane-change count	times	17	13	7	...	5	10	13
	Accelerating time	s	41.3	43.0	19.6	...	10.0	34.2	51.6
	Decelerating time	s	39.7	46.0	17.7	...	10.6	25.9	39.4
	Cruising time	s	799.2	787.3	282.0	...	281.3	397.7	517.6
	Idling time	s	150.5	275.5	76.0	...	0.0	44.1	100.1
	Starting time	s	0.6	3.3	0.3	...	0	3.1	5.3
	Average time headway	s	13.8	506.0	14.4	...	11.3	20.0	572.3
	Minimum time headway	s	1.77	1.15	1.2	...	1.83	1.81	1.93
	Average space headway	m	170.0	124.5	186.3	...	151.8	169.0	221.3
Minimum space headway	m	22.5	2.0	14.8	...	13.3	2.0	2.0	
Output	EC <sub>b</sub>	kWh	1.14	0.94	0.38	...	0.36	0.62	0.82

Notes: (1) Travel delay: refers to the total time delay during the travel per EV, including queuing delay, braking delay, etc.

(2) Accelerating time, Decelerating time, Cruising time, Idling time, Starting time: time statistics for different driving states during the travel per EV.

The variables in Table 4 require additional clarification. The signal states corresponding to the arrival at the intersection have a different impact on the driving behaviours of the vehicle, which in turn affects the EC. The corresponding state of the vehicle arriving at the intersection is shown in Figure 11. The serial number ①~⑦ indicates seven EVs, corresponding to different signal states when they arrive at the intersection. EVs ①~③ line up due to red lights, but the queuing time differs. EV ① lines up for longer, while EV ③ queues for a shorter time. Moreover, EVs often queue twice under high traffic flow, namely prolonged queuing time, resulting in increased EC due to time delays. EVs ④~⑥ arrive at the intersection during the green light, and they will adopt different accelerations and decelerations to pass the intersection before the end of the green light. When the EV ⑦ arrives at the intersection during the yellow light, it usually slows down. It passes through the intersection after confirming safety.

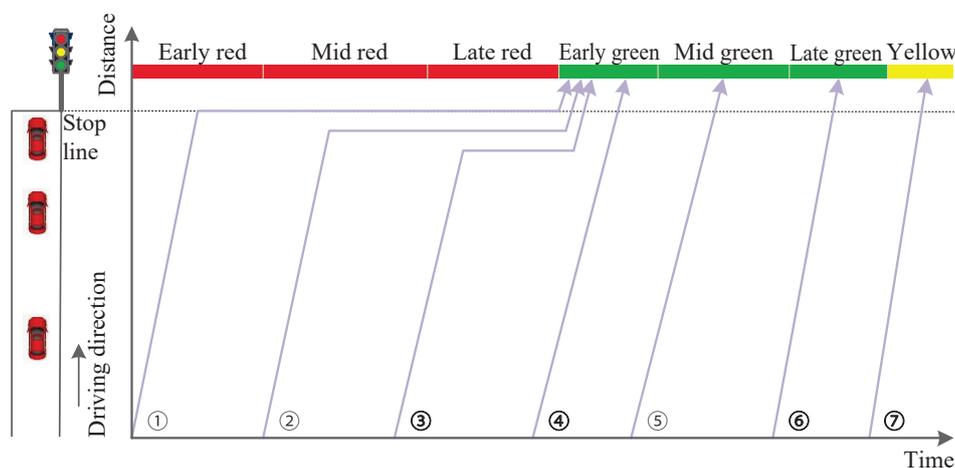


Figure 11 – Schematic diagram of the corresponding signal phase when the vehicle arrives at the intersection

Table 4 – Signal control-related variables

Input	Parameters	Intersection number	Unit	Training				Test			
				veh <sub>1</sub>	veh <sub>2</sub>	veh <sub>3</sub>	veh <sub>l</sub>	veh <sub>1</sub>	veh <sub>2</sub>	veh <sub>3</sub>	veh <sub>j</sub>
Input	Yellow light stage	①	-	0	0	0	...	0	0	0	...
		②	-	0	0	0	...	0	0	0	...
		③	-	0	0	0	...	0	0	0	...
		④	-	0	0	0	...	0	0	0	...
		⑤	-	0	0	0	...	0	0	0	...
	Early green stage	①	-	0	1	0	...	0	0	0	...
		②	-	0	0	0	...	0	0	0	...
		③	-	0	0	0	...	0	0	1	...
		④	-	0	0	0	...	0	0	0	...
		⑤	-	0	0	0	...	1	0	1	...
	Mid green stage	①	-	0	0	0	...	0	0	0	...
		②	-	0	0	0	...	0	0	0	...
		③	-	0	0	0	...	0	0	0	...
		④	-	1	0	0	...	0	0	1	...
		⑤	-	0	0	0	...	0	1	0	...
	Late green stage	①	-	0	0	0	...	0	0	0	...
		②	-	1	0	0	...	0	0	0	...
		③	-	0	0	0	...	0	0	0	...
		④	-	0	0	0	...	0	0	0	...
		⑤	-	1	0	0	...	0	0	0	...
	Early red stage	①	-	0	0	1	...	0	0	0	...
		②	-	0	1	0	...	0	0	0	...
		③	-	1	1	0	...	0	0	0	...
		④	-	0	1	0	...	0	0	0	...
		⑤	-	0	0	0	...	0	0	0	...
Mid red stage	①	-	1	0	0	...	0	0	0	...	
	②	-	0	0	0	...	0	0	0	...	
	③	-	0	0	0	...	0	0	0	...	
	④	-	0	0	0	...	0	0	0	...	
	⑤	-	0	0	0	...	0	0	0	...	
Late red stage	①	-	0	0	0	...	0	0	0	...	
	②	-	0	0	0	...	0	0	0	...	
	③	-	0	0	0	...	0	0	0	...	
	④	-	0	0	0	...	0	1	0	...	
	⑤	-	0	0	0	...	0	0	0	...	
Output	EC <sub>b</sub>		kWh	1.14	0.94	0.38	...	0.36	0.62	0.82	...

Note: The numbers ①~⑤ in the “Position” column in the table are the numbers of the five signalised intersections, indicating that the EVs arrive at the intersection’s stop line corresponding to the number at the corresponding time stage. The value 1 indicates that the vehicle comes at the designated intersection in the corresponding signal phase. Otherwise, it is marked as a value of 0.

## 4.2 Conventional EC prediction method

The conventional prediction method for EC utilises historical driving data to estimate the EC of future travels. This calculation process is relatively simple. The average ECR of each historical trip can be regarded as approximately constant within a specific range. First, calculate the average ECR of the historical driving range and then predict the EC of future travels, as shown in Equation 1

$$EC^{pred}(x) = ECR^{pred}(x) \cdot M_{future}(x) \approx \frac{EC^{past}(x)}{M_{past}(x)} \cdot M_{future}(x) \quad (1)$$

where  $EC^{pred}$  represents the predicted EC of the EV within a certain driving distance,  $M_{future}$ ;  $ECR$  is the energy consumption rate (kWh/km) and  $ECR^{pred}$  is the predicted value of  $ECR$ ;  $M_{future}$ ,  $M_{past}$  refers to the future and past driving distance of the EV, respectively;  $EC^{past}$  represents the EC value of the EV within a certain driving distance  $M_{past}$  in the past.

In this study, the above mentioned conventional method is used to predict the EC for EVs and compare it with the actual value, as shown in Table 5. However, many factors are affecting the EC. The conventional method utilises the historical ECR to predict the future EC. It only considers the driving distance, thus ignoring other factors that may have a non-negligible impact on the EC, resulting in inconsistent prediction results.

Table 5 – The examples of EC prediction results using the conventional method

$ECR^{past}$ [kWh/km]	Test range [km]	$EC^{pred}$	$EC^{real}$	Relative error [%]
0.0996	9.9740	0.9931	0.9921	0.10
0.0922	5.9895	0.5525	0.6038	8.49
0.0888	7.9895	0.7096	0.8885	20.14
0.0945	9.9895	0.9441	1.0517	10.23
0.1054	5.9767	0.6298	0.6227	1.14
0.1028	3.9949	0.4107	0.3971	3.42
0.0946	11.9740	1.1322	1.1991	5.58
0.1077	3.9922	0.4300	0.4013	7.14
0.0830	11.9767	0.9944	1.1628	14.48
0.0889	5.9586	0.5296	0.6051	12.47

## 4.3 The EC prediction framework based on LightGBM model

The eXtreme Gradient Boosting (XGBoost) model is a large-scale parallel-boost tool with high prediction accuracy and flexibility. However, it has shortcomings. Although the utilisation of pre-sorting and approximation algorithms can reduce the amount of calculation to find the best-split point, the dataset must still be traversed during the node-splitting process. The space complexity of the pre-sorting process is unduly high, and not only needs to store the feature value but needs to store the index of the statistical gradient value of the sample corresponding to the feature, which is equivalent to consuming twice the memory [27]. The LightGBM is a light gradient boosting machine, which was improved based on the shortcomings of XGBoost. Compared with XGBoost, on the one hand, LightGBM occupies less memory, reduces spatial complexity and the number of features during training, and dramatically diminishes memory consumption. On the other hand, LightGBM is faster, which significantly decreases the temporal complexity, reduces unnecessary calculations and optimises the cache.

The EC prediction framework based on the LightGBM model is proposed in this paper, as shown in Figure 12. The framework mainly includes five parts: ① data collection; ② data preprocessing; ③ influencing factors analysis of EC; ④ feature extraction; ⑤ training and optimisation for the prediction model. This prediction framework is mainly composed of the model training and prediction phases.

*Data collection:* simulation in SUMO to get the raw data and use of Python to eliminate data noise.

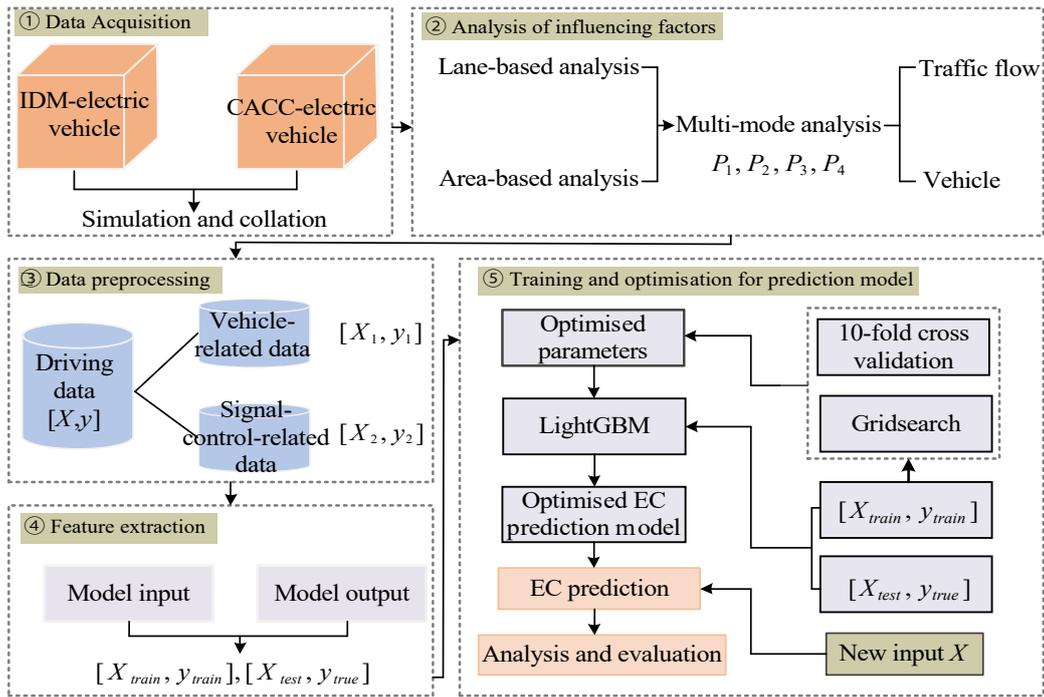


Figure 12 – The EC prediction framework

*Data preprocessing:* according to the characteristics of the urban road environment, the data were divided into vehicle travel-related data and signal control-related data.

*Influencing factors analysis of EC:* in previous related studies, the speed and acceleration of EVs have attracted much attention, but in fact, the EC of EVs is affected concurrently by many influencing factors. This framework first analysed the EC laws of EVs with different trajectories according to lanes and areas, then summarised four EC modes from perspectives of traffic flow and vehicles, and induced the influencing factors of EC.

*Feature extraction:* this framework considered travel-related factors (Table 3), including travel time, distance, delay, speed and acceleration. In addition, the influence of signal control is also taken into account (Table 4), including the seven phases, early green stage, mid-green stage, late green stage, early red stage, mid-red stage, late red stage and yellow stage.

*Training and optimisation for the prediction model:* LightGBM implements machine learning algorithms under the tree boosting framework, efficiently constructing boosting trees and realising parallel operation. It is so efficient that it can solve the problem of EC prediction under various working conditions. In the meantime, compared to XGBoost, LightGBM proposes two improvements: gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB) [28].

*GOSS:* sort according to the absolute value of the gradient of all samples, select the first  $a \cdot 100\%$  samples (to form set  $A$ ), and randomly select  $b \cdot 100\%$  samples (to form set  $B$ , its size is  $b \cdot |A^c|$ ) from the remaining  $(1 - a) \cdot 100\%$  samples (called a set  $A^c$ ) to generate a small gradient sample set. When calculating the information gain for small gradient samples, multiply it by a constant  $\frac{1-a}{b}$ . Therefore, LightGBM only pays more attention to the under-trained samples and does not excessively change the original data distribution.

GOSS divides the data on the set  $A \cup B$  by estimating the variance gain  $\tilde{V}_j(d)$ :

$$\tilde{V}_j(d) = \frac{1}{n} \left( \frac{\left( \sum_{x_i \in A_i} g_i + \frac{1-a}{b} \sum_{x_i \in B_i} g_i \right)^2}{n_i^j(d)} + \frac{\left( \sum_{x_i \in A_r} g_i + \frac{1-a}{b} \sum_{x_i \in B_r} g_i \right)^2}{n_r^j(d)} \right) \quad (2)$$

where  $A_l = \{x_i \in A : x_{ij} \leq d\}$ ,  $A_r = \{x_i \in A : x_{ij} > d\}$ ,  $B_l = \{x_i \in B : x_{ij} \leq d\}$ ,  $B_r = \{x_i \in B : x_{ij} > d\}$ ,  $\frac{1-a}{b}$  is used to normalise the sum of gradients of dataset  $B$ .

Let  $\varepsilon(d) = |\tilde{V}_j(d) - V_j(d)|$  be the approximate error of GOSS  $\bar{g}_l^j(d) = \frac{\sum_{x_i \in (A \cup A^c)_r} |g_i|}{n_l^j(d)}$ , then:

$$\varepsilon(d) \leq C_{a,b}^2 \ln 1/\delta \cdot \max\left\{\frac{1}{n_l^j(d)}, \frac{1}{n_r^j(d)}\right\} + 2DC_{a,b} \sqrt{\frac{\ln 1/\delta}{n}} \tag{3}$$

where  $C_{a,b} = \frac{1-a}{\sqrt{b}} \max_{x_i \in A^c} |g_i|$ , and  $D = \max(\bar{g}_l^j(d), \bar{g}_r^j(d))$ .

Therefore, the asymptotic approximation probability of GOSS is  $O(\frac{1}{n_l^j(d)} + \frac{1}{n_r^j(d)} + \frac{1}{\sqrt{n}})$  if the split is unbalanced (such as  $n_l^j(d) \geq O(\sqrt{n})$  or  $n_r^j(d) \geq O(\sqrt{n})$ ), the approximation probability depends on the second term of Equation 2 (if  $n \rightarrow \infty$  the second term tends to 0), which means that when there are vast samples, the approximation probability is entirely accurate.

*EFB*: in actual application, although there are great features, many are sparse, and EFB provides an almost lossless method to reduce the possibility of the number of practical features. It bundles mutually exclusive features into one, designs a suitable algorithm for the optimisation bundling problem as a graph colouring problem, and utilises a greedy algorithm to approximate the solution. For feature bundling, the complexity can be reduced from  $O(\text{data} \cdot \text{feature})$  to  $O(\text{data} \cdot \text{bundle})$  (*bundle* is much smaller than *feature*), which is the reason for its acceleration.

Combined with the above advantages, the framework based on LightGBM has preferable applicability to this study’s EC prediction. In addition, a model optimisation framework combining KFold and GridSearch was used in the model training and optimisation process. The model performance was evaluated comprehensively by KFold with sufficient training samples. Hyperparameters were empirically established, and GridSearch was applied to stepwise test different combinations of hyperparameters to improve model accuracy.

#### 4.4 Model evaluation and comparison

This paper utilises root mean square error (RMSE) and mean absolute percentage error (MAPE) to evaluate the prediction performance of the framework. RMSE and MAPE are respectively defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \tag{4}$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{5}$$

where  $\{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$  is the predicted value and  $\{y_1, y_2, \dots, y_n\}$  is the actual value. The smaller the value of RMSE and MAPE is, the more accurate the prediction results can be.

The framework is trained using the training set and evaluated by the 10-fold cross-validation training strategy. As shown in Figure 13, the model performed differently on each training and verification sample. The average indexes of 10-fold cross-validation determined the final model prediction performance. MAPE was applied in model training and optimisation to evaluate the model’s accuracy. Based on 10-fold cross-validation and evaluation indicators, the critical hyperparameters in LightGBM, including learning rate, maximum tree depth, sub-sample training examples and regularisation terms, were optimised to improve model performance and reduce prediction errors. As shown in Figure 14, taking the learning rate optimisation in the training process as an example, a grid with a learning rate of 0.01 to 0.1 was set, and ten cross-validations were used for testing. For the IDM-EVs, the results show that the error is the smallest when the learning rate is 0.04, while for the CACC-EVs, the error is the smallest when the learning rate is 0.03.

The trained model is employed to predict EC based on the test set. As illustrated in Figure 15, the  $R^2$  values of the EC prediction results are 0.96 and 0.95, respectively, which indicates that more than 95% of the variation

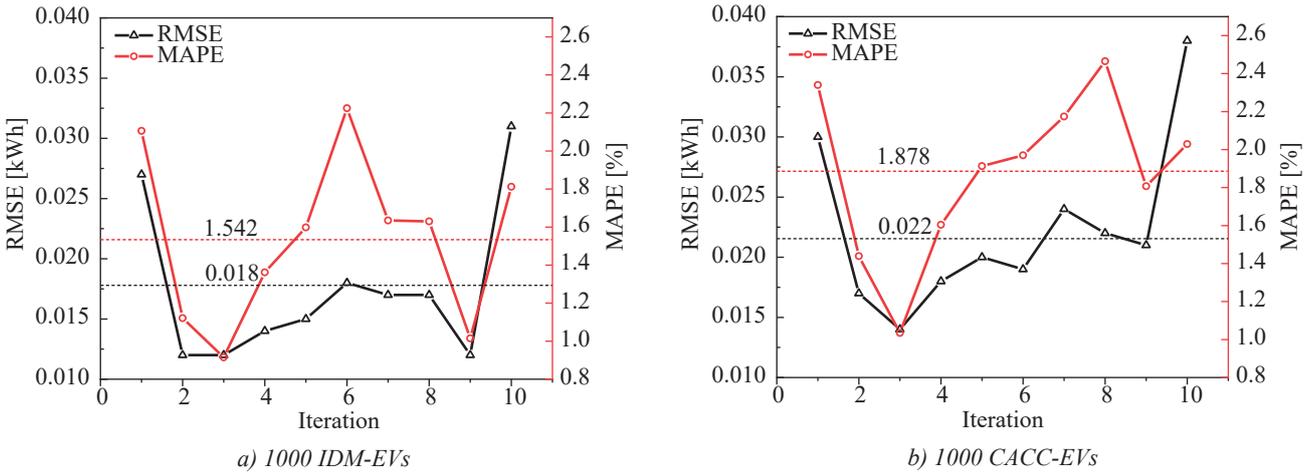


Figure 13 – 10-fold cross-validation

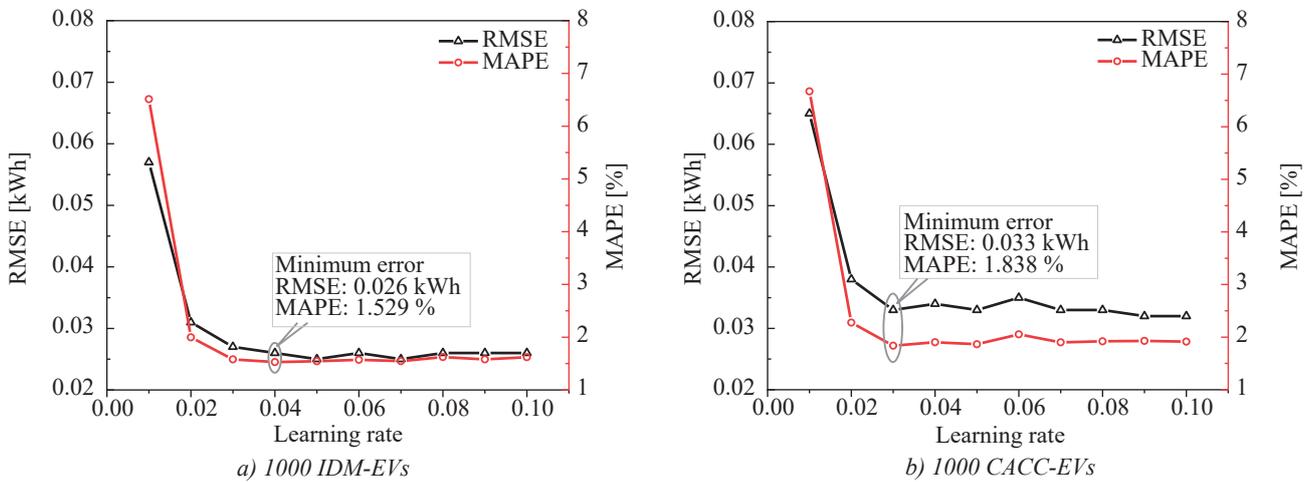


Figure 14 – Grid search

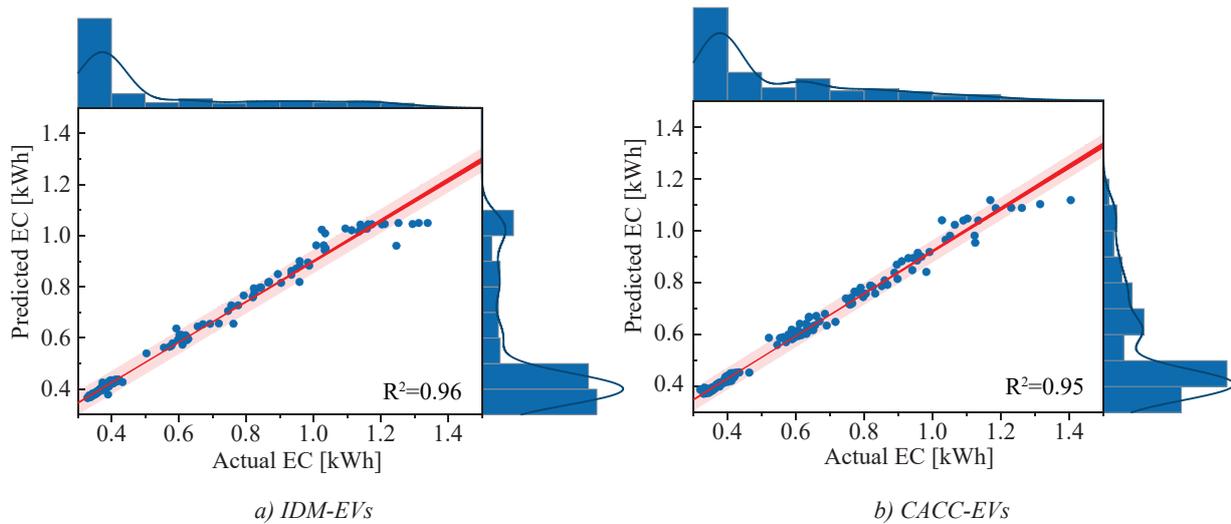


Figure 15 – The prediction results based on the proposed framework

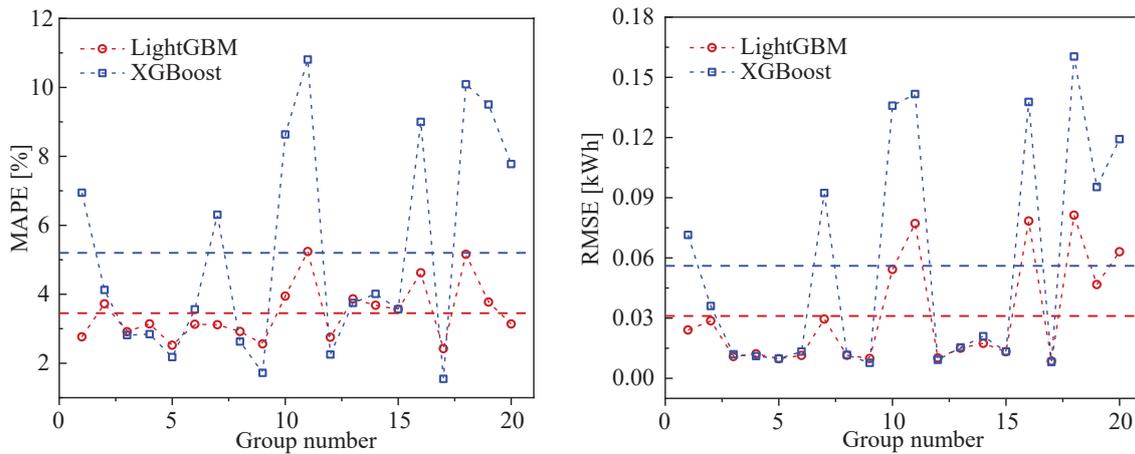
characteristics of EC can be explained using the variables in *Tables 3 and 4*. To verify the prediction performance of the LightGBM model, in addition to the conventional method, this study adopts two other comparison algorithms, including the XGBoost model and the gradient boosted regression tree (GBRT) model. They are gradient boosting algorithms that are highly accurate at handling the complex nonlinear relationship between

feature inputs and EC. The results of the two evaluation indicators are shown in Table 6. For the IDM dataset, the average prediction errors of the LightGBM model are 3.45% (MAPE) and 0.039 kWh (RMSE), which are 33.7% and 50.6% lower than the best comparison algorithm (XGBoost model), respectively. For the CACC dataset, the average prediction errors of the LightGBM model are 5.57% (MAPE) and 0.042 kWh (RMSE), which are 16.1% and 34.4% lower than the best comparison algorithm (XGBoost model), respectively. The conventional method shows the lowest prediction performance on both datasets. Figure 16 shows that the proposed framework exhibits better predictions in most cases of the test set. Based on the analysis of the above results, the proposed framework shows the best performance in predicting the EC of IDM-EVs and CACC-EVs.

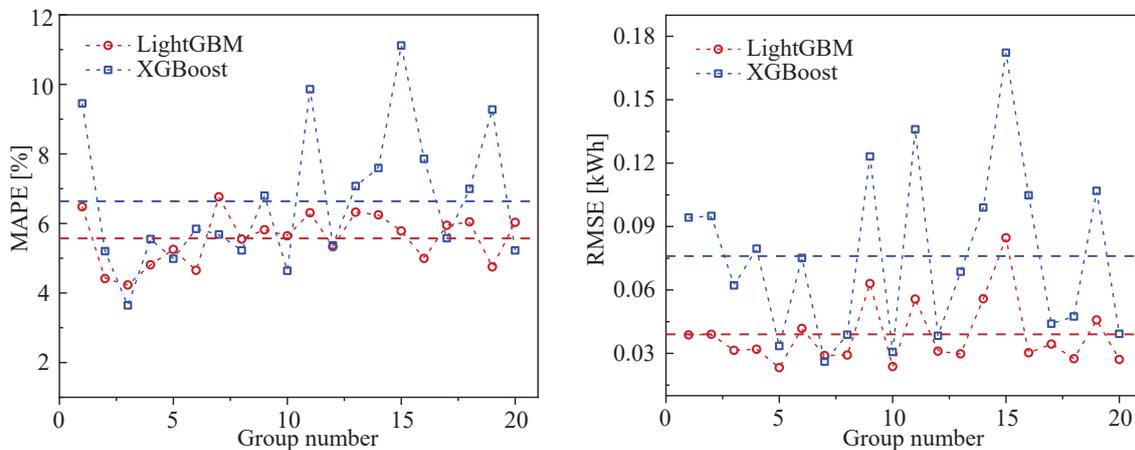
Table 6 – Prediction performance of LightGBM model and comparison algorithms

Model	MAPE [%]		RMSE [kWh]	
	IDM-EVs	CACC-EVs	IDM-EVs	CACC-EVs
Conventional method	8.62	9.54	0.085	0.082
GBRT	6.98	7.45	0.074	0.081
XGBoost	5.20	6.64	0.079	0.064
LightGBM	3.45	5.57	0.039	0.042

To understand the impact of the extracted features on the EC prediction and to provide the necessary interpretability for the prediction results of the proposed framework, this study obtained the feature importance ranking of the LightGBM model on the test set, as shown in Figure 17. The abscissa represents 53 features,



a) IDM-EVs



b) CACC-EVs

Figure 16 – The EC prediction results of 20 group EVs

sorted according to the features shown in *Tables 3 and 4*. The ordinate represents the feature importance value. The top five essential features are marked in the figure. The 19th to 53rd features on the abscissa represent the signal phases corresponding to vehicle arrival at each intersection (as shown in *Table 4*). Compared with the first 18 features, the latter 35 features are of lower importance, but this shows that signal control has a certain degree of influence on the EC of EVs. Signal control factors may not be neglected in future related microscopic studies. It can be seen that there are certain differences in the EC feature importance between IDM-EVs and CACC-EVs. The driving range, acceleration, accelerating time, decelerating time and cruising time marked in the figure correspond to larger feature importance values, which are significant factors affecting the EC of EVs.

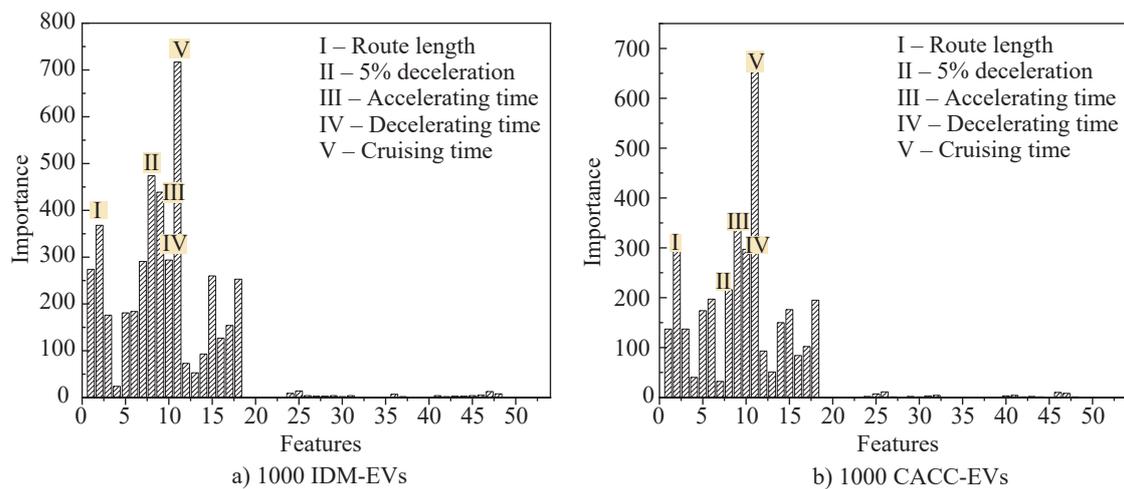


Figure 17 – Feature importance maps for EC prediction

## 5. CONCLUSION

This paper analyses EC by driving areas and defines accelerating, decelerating, cruising, idling and starting states for EVs. The four EC modes ( $P_1, P_2, P_3, P_4$ ) are classified according to the driving states from the perspectives of traffic flow and vehicles, and the EC laws and distribution differences between IDM-EVs and CACC-EVs under different traffic flows are compared and analysed. Based on the above analysis, the influencing factors of EC are summarised and extracted, which are mainly related to vehicle travel and signal control. Finally, the EC prediction framework based on the LightGBM algorithm is proposed, and its prediction results are validated using other advanced comparison algorithms. The results show that the proposed framework exhibits the best performance in EC prediction for IDM-EVs and CACC-EVs. For the IDM dataset, the average prediction error of the proposed framework is 3.45% (MAPE) and 0.039 kWh (RMSE), which is 33.7% and 50.6% lower in relation to the best comparison algorithm (XGBoost model). For the CACC dataset, the average prediction error of the proposed framework is 5.57% (MAPE) and 0.042 kWh (RMSE), which is 16.1% and 34.4% lower in relation to the XGBoost model. In addition, this paper brings about the following conclusions:

- 1) Due to the relatively complex urban road environment, many factors affect the EC of EVs. This paper extracted vehicle travel-related and signalled control-related factors. The final results show that driving distance, acceleration, decelerating and cruising are critical. They are significant influencing factors in the overall EC during travel.
- 2) There is an apparent discrepancy in the EC and its EV distribution under high and low traffic flows. Under low traffic flow, the EV gains higher freedom, and the interplay between vehicles is inconspicuous so that the average EC of EVs is low; under high traffic flow, the driving freedom of EV is reduced, and the interplay between EVs is excellent, especially the queuing time is prolonged at signalised intersections, resulting in increased idling EC. Previous studies [5, 6, 8] did not specifically consider the factor of flow, which may weaken the generalisation ability of the energy consumption prediction algorithm.
- 3) This paper compared and analysed the EC of IDM-EVs and CACC-EVs. Compared to IDM-EVs, the results show that CACC-EVs consume more energy at lower flow, while the opposite is true at higher flow. It is difficult to form a fleet among the CACC-EVs under low traffic flow, and the car-following advantage is lost when losing the “guidance” of the leader. However, CACC-EVs can readily form a fleet under

high traffic flow, which can significantly optimise driving trajectories to reduce the EC. Previous research [13–15] predicted the EC of manual driving vehicles. However, intelligent connected EVs possess apparent advantages in the IoV environment, so its EC law is worth exploring.

EC has become one of the essential concerns of EV drivers. Accurate and real-time prediction of EC can significantly improve the range anxiety for drivers. Because of the above mentioned situation, the prediction framework of electric vehicle energy consumption proposed in this study possesses high accuracy and good real-time performance, suitable for various traffic scenarios.

This study also has limitations, mainly reflected in ① the analysis is only based on signalized intersections; ② the degradation of EV battery is not considered; ③ only urban road scenarios are considered. Therefore, to improve the prediction performance of the proposed framework, the above factors be considered in future studies, such as: analysing the trajectories and EC laws at the actuated signalised intersection and comparing it with the case of fixed-time signalised control; take into account the impact of battery status on EC; applying the proposed prediction framework to a variety of scenarios, for instance, highway sections, up and down ramps. In addition, with the rapid development of the IoV, the framework can be utilised to optimise travel paths for intelligent connected EVs and further applied in the operation of urban road charging infrastructure.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (52372413, U20A20333, U20A20331, 52225212, 52072160, 51905223); Key Research and Development Program of Jiangsu Province (BE2020083-3, BE2019010-2, BE2021011-3); Natural Science Foundation of Jiangsu Province (BK20180100); Six Talent Peaks Project of Jiangsu Province (2018-TD-GDZB-022); Research project of Jiangsu Provincial Public Security Department (2021KO002Z); Transportation Science and Technology Project of Jiangsu Province (2021G05, 2022Y03); Postgraduate Research & Practice Innovation Program of Jiangsu Province (S-JCX23\_2049) and the Young Talent Cultivation Project of Jiangsu University.

## REFERENCES

- [1] Tu W, et al. Acceptability, energy consumption, and costs of electric vehicle for ride-hailing drivers in Beijing. *Applied Energy*. 2019;250:147-160. DOI: 10.1016/j.apenergy.2019.04.157.
- [2] Smith R, Shahidinejad S, Blair D, Bibeau EL. Characterization of urban commuter driving profiles to optimize battery size in light-duty plug-in electric vehicles. *Transportation Research Part D: Transport and Environment*. 2011;16(3):218-224. DOI: 10.1016/j.trd.2010.09.001.
- [3] Basso R, et al. Energy consumption estimation integrated into the electric vehicle routing problem. *Transportation Research Part D: Transport and Environment*. 2019;69:141-167. DOI: 10.1016/j.trd.2019.01.006.
- [4] Speidel S, Bräunl T. Driving and charging patterns of electric vehicles for energy usage. *Renewable and Sustainable Energy Reviews*. 2014;40:97-110. DOI: 10.1016/j.rser.2014.07.177.
- [5] Wu X, Freese D, Cabrera A, Kitch WA. Electric vehicles' energy consumption measurement and estimation. *Transportation Research Part D: Transport and Environment*. 2015;34:52-67. DOI: 10.1016/j.trd.2014.10.007.
- [6] Wang J, Besselink I, Nijmeijer H. Online prediction of battery electric vehicle energy consumption. *World Electric Vehicle Journal*. 2016;8(1):213-224. DOI: 10.3390/wevj8010213.
- [7] Fiori C, Ahn K, Rakha HA. Power-based electric vehicle energy consumption model: Model development and validation. *Applied Energy*. 2016;168:257-268. DOI: 10.1016/j.apenergy.2016.01.097.
- [8] Xing Y, Lv C, Cao D, Lu C. Energy oriented driving behavior analysis and personalized prediction of vehicle states with joint time series modeling. *Applied Energy*. 2020;261:114471. DOI: 10.1016/j.apenergy.2019.114471.
- [9] Liu K, Wang J, Yamamoto T, Morikawa T. Exploring the interactive effects of ambient temperature and vehicle auxiliary loads on electric vehicle energy consumption. *Applied Energy*. 2018;227:324-331. DOI: 10.1016/j.apenergy.2017.08.074.
- [10] Ahn K, Park S, Rakha HA. Impact of intersection control on battery electric vehicle energy consumption. *Energies*. 2020;13(12):3190. DOI: 10.3390/en13123190.
- [11] Wang J, Besselink I, Nijmeijer H. Battery electric vehicle energy consumption prediction for a trip based on route information. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*. 2018;232(11):1528-1542. DOI: 10.1177/0954407017729938.
- [12] Guo J, Jiang Y, Yu Y, Liu W. A novel energy consumption prediction model with combination of road information and driving style of BEVs. *Sustainable Energy Technologies and Assessments*. 2020;42:100826. DOI: 10.1016/j.seta.2020.100826.
- [13] Yi Z, Smart J, Shirk M. Energy impact evaluation for eco-routing and charging of autonomous electric vehicle fleet: Ambient temperature consideration. *Transportation Research Part C: Emerging Technologies*. 2018;89:344-363. DOI: 10.1016/j.trc.2018.02.018.

- [14] De Cauwer C, et al. A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions. *Energies*. 2017;10(5):608. DOI: 10.3390/en10050608.
- [15] Modi S, Bhattacharya J, Basak P. Estimation of energy consumption of electric vehicles using deep convolutional neural network to reduce driver's range anxiety. *ISA Transactions*. 2020;98:454-470. DOI: 10.1016/j.isatra.2019.08.055.
- [16] Madhusudhanan AK, Na X. Effect of a traffic speed based cruise control on an electric vehicle's performance and an energy consumption model of an electric vehicle. *IEEE/CAA Journal of Automatica Sinica*. 2020;7(2):386-394. DOI: 10.1109/JAS.2020.1003030.
- [17] Zhang J, Wang Z, Liu P, Zhang Z. Energy consumption analysis and prediction of electric vehicles based on real-world driving data. *Applied Energy*. 2020;275:115408. DOI: 10.1016/j.apenergy.2020.115408.
- [18] Yao J, Moawad A. Vehicle energy consumption estimation using large scale simulations and machine learning methods. *Transportation Research Part C: Emerging Technologies*. 2019;101:276-296. DOI: 10.1016/j.trc.2019.02.012.
- [19] Fukushima A, et al. Prediction of energy consumption for new electric vehicle models by machine learning. *IET Intelligent Transport Systems*. 2018;12(9):1174-1180. DOI: 10.1049/iet-its.2018.5169.
- [20] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17 Aug. 2016, San Francisco, CA*. 2016. p. 785-794. DOI: 10.1145/2939672.2939785.
- [21] Chang Y, Yang W, Zhao D. Energy efficiency and emission testing for connected and automated vehicles using real-world driving data. *2018 21st International Conference on Intelligent Transportation Systems (ITSC), 4-7 Nov. 2018, Maui, HI, USA*. 2018. p. 2058-2063. DOI: 10.1109/ITSC.2018.8569806.
- [22] Gawron JH, et al. Life cycle assessment of connected and automated vehicles: Sensing and computing subsystem and vehicle level effects. *Environmental Science & Technology*. 2018;52(5):3249-3256. DOI: 10.1021/acs.est.7b04576.
- [23] Mei Z, Xiang F, Zhen-hui L. Short-term traffic flow prediction based on combination model of XGBoost-lightGBM. *2018 International Conference on Sensor Networks and Signal Processing (SNSP), 28-31 Oct. 2018, Xi'an, China*. 2018. p. 322-327. DOI: 10.1109/SNSP.2018.00069.
- [24] Wen X, Xie Y, Wu L, Jiang L. Quantifying and comparing the effects of key risk factors on various types of roadway segment crashes with LightGBM and SHAP. *Accident Analysis & Prevention*. 2021;159:106261. DOI: 10.1016/j.aap.2021.106261.
- [25] Chen T, Shi X, Wong YD, Yu X. Predicting lane-changing risk level based on vehicles' space-series features: A pre-emptive learning approach. *Transportation Research Part C: Emerging Technologies*. 2020;116:102646. DOI: 10.1016/j.trc.2020.102646.
- [26] Fiori C, Ahn K, Rakha H A. Power-based electric vehicle energy consumption model: Model development and validation. *Applied Energy*. 2016;168:257-268. DOI: 10.1016/j.apenergy.2016.01.097.
- [27] Liu H, et al. LightGBM-based prediction of remaining useful life for electric vehicle battery under driving conditions. *2020 IEEE Sustainable Power and Energy Conference (iSPEC), 23-25 Nov. 2020, Chengdu, China*. 2020. p. 2577-2582. DOI: 10.1109/iSPEC50848.2020.9351029.
- [28] Ke G, et al. LightGBM: A highly efficient gradient boosting decision tree. *31st Annual Conference on Neural Information Processing Systems (NIPS), 4-9 Dec. 2017, Long Beach, CA*. 2017. p. 30. <https://webvpn.ujs.edu.cn/https/77726476706e69737468656265737421f4fb0f9d243d265f6c0f/doi/pdf/10.5555/3294996.3295074/3294996.3295074.pdf> [Accessed 31st May 2023].

刘擎超, 高芬霞, 赵晶娅, 周卫琪  
智能网联环境下电动汽车能耗预测研究

#### 摘要

准确的能耗预测对于改善驾驶体验至关重要。在城市道路场景中, 我们讨论了能源消耗的影响因素, 并从各个角度对模式进行了划分。比较了使用IDM和CACC跟车模型的电动汽车在不同交通流量下的能耗特征和分布规律的差异。提出了一种基于LightGBM模型的能耗预测框架。研究表明, 续航里程、加速、加速时间、减速时间和巡航时间都会显著影响电动汽车的整体能耗。不同交通流量下的能耗特征和分布规律存在明显差异: 低流量下平均能耗较低, 高流量下平均功耗较高。CACC电动汽车在低流量下比IDM电动汽车消耗更多的能量。在高流量下, 情况正好相反。结果表明, 该框架具有较高的精度: 基于IDM数据集的MAPE为3.45%, RMSE为0.039kWh; 基于CACC数据集的MAPE为5.57%, RMSE为0.042kWh。与最佳比较算法相比, MAPE和RMSE分别降低了33.7%和50.6% (最大程度)。

#### 关键词:

车辆安全系统; 主动安全系统; 智能座舱; 疲劳检测; 成像光电容积描记法