Int j simul model 21 (2022) 3, 489-500
Original scientific paper

A SIMULATION STUDY ON THE IDENTIFICATION OF ECO-DRIVING BEHAVIOUR

Yan, L. X.[*,#]; Jia, L.[*]; Guo, J. H.[*] & Lu, S.[**]
[*] School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China
[**] Institute of Intelligence Science and Engineering, Shenzhen Polytechnic, Shenzhen 518000, China
E-Mail: yanlixinits@163.com ([#] Corresponding author)

Abstract
Eco-driving is considered as one of the effective ways to reduce energy consumption. The objective of this study is to establish an eco-driving behaviour identification model and analyse the eco-driving behaviour characteristics. First, this study built a driving simulation platform and conduct an experiment for research. Then, based on significant driving behaviours related to fuel consumption, Piecewise Linear Representation (PLR) method was used to fit multivariate time series consisting of Significant driving behaviour variables. At last, the features of each time series segment were extracted as the input of a Random Forest (RF) model to recognize the eco-driving behaviour. The results indicated that the depth of the acceleration pedal, the depth of the clutch pedal, the depth of the brake pedal, steering wheel angle, and gear position had significant effects on fuel consumption. The eco-driving behaviour identification model demonstrated a high predictive power with a prediction accuracy of 0.718. The graph of eco-driving behaviours during lane change was established. The conclusions provide theoretical support for developing eco-driving intervenes.
(Received in March 2022, accepted in May 2022. This paper was with the authors 2 weeks for 1 revision.)

Key Words: Eco-Driving, Driving Behaviour, Time Series Segmentation, Piecewise Linear Representation, Random Forest

1. INTRODUCTION

Many countries are adopting policies to reduce gas emissions and fuel consumption in response to the impact of global warming. The emission in the transportation has become a hot issue of energy conservation and emission reduction. CO_2 emissions from road transport account for 17 % of total CO_2 emissions in main countries in the world and are expected to increase in the future [1]. Eco-driving is an effective way to increase energy efficiency and has attracted widespread research interest [2]. Previous studies have found that environmentally friendly driving can reduce fuel consumption by 5-10 %, and by up to 20-50 % [3]. Thus, a method that could identify eco-driving behaviours and the description of eco-driving behaviours are necessary.

Modelling and simulation techniques are being widely applied to transportation research [4, 5]. Compared with field tests driving simulation experiments present obvious advantages such as risk-free, cost-effective, and repeatability, also which are easy for data collection [6-8]. The driving simulation also was widely used in eco-driving behaviour research. For example, Wu et al. and Wu and Zhao [9, 10] obtained relevant driving operation behaviour indicators (the depth of acceleration pedal, the depth of clutch pedal, the depth of brake pedal, the gear position, and steering wheel angle) to identify eco-driving behaviour and established the eco-driving graph based on a driving simulator. It can verify the efficacy of eco-driving assistance systems (EDAS) and evaluate the fuel-saving potential [11].

The methodology of eco-driving behaviour identification was various. The application of the models including consumption models, linear models, and neural networks is popular. For example, several methods based on dynamic programming (DP) and the joint optimization model were to predict near-optimal control trajectories; this method reduces computational complexity and enables real-time prediction of optimal trajectories [12, 13]. Meanwhile,

driving operation indicators have achieved remarkable results in driving behaviour recognition and prediction [14-16]. It includes the combined eco-driving behaviour prediction models of energy consumption based on output component analysis and back-propagation (BP) neural network [10, 17]. The eco-driving behaviours prediction was also made by constructing models based on consumption models including comprehensive modal emission model (CMEM) models and the models of EIMT [18-20].

Meanwhile, the eco-driving behaviour rules specify what the driver should do in specific situations [21]. The eco-driving behaviour standards are different in different driving scenarios [22]. Therefore, the definition of eco-driving behaviour should be comprehensive. It should provide a parsimonious and reliable interpretation of the driving behaviours that constitute eco-driving rules in the corresponding topography and scenarios.

Few studies identified eco-driving behaviours based on sequence features of driving behaviours. Previous studies lacked an intuitive representation and detailed description of the eco-driving behaviour time series. This paper was organized into the following step. First, a driving simulation platform was established to collect the multivariate time series data of the lane-changing driving behaviours. Then, the significance analysis based on Kruskal-Wallis statistical tests was used to evaluate the effect of driving behaviour variables on fuel consumption. Third, the characteristics indexes of multivariate time series composed of significant driving behaviours were extracted as model input based on PLR, and a Random Forest model was established to identify lane change eco-driving behaviours. The study also established other two eco-driving identification models based on the KNN algorithm and SVM algorithm for comparison. Meanwhile, the influence of the number of segments used for the linear representation of time series on the recognition model was explored. Finally, the eco-driving behaviour graph was established.

2. RESEARCH MODEL

2.1 Segmentation and linear representation of multivariate time series

The time series of lane change driving behaviours are high-dimensional and massive. In order to quickly extract the significant features of the time series, Piecewise Linear Representation (PLR) method was applied to fit multivariate time series. PLR can compress data, and was widely used for the advantages of intuitive data representation and good noise filtering effect [23]. The main idea of this method is to approximate the original time series with several adjacent straight-line segments. The PLR pattern of time series data $X = \{x_1, x_2, ..., x_n\}$ divided into K segments can be expressed as the following Eq.

$$X(t) = \begin{Bmatrix} f_1(t, w_1) + e_1(t), & t \in (1, t_1] \\ f_2(t, w_1) + e_2(t), & t \in (t_1, t_2] \\ \vdots & \vdots \\ f_K(t, w_m) + e_m(t), & t \in (t_{m-1}, t_m] \end{Bmatrix} \quad (1)$$

where w_j is the j^{th} segment of the original time series after piecewise linear representation, and the time interval is $[w_{j-1}, w_j]$; $f_j(t, w_j)$ is the linear function connecting the two endpoints of w_j; $e_j(t)$ is the fit error between the sequence data by the linear representation of this segment and the original data.

In order to make the error between the linear representation result and the original time series data as small as possible, the key point of the PLR linear representation method is the determination of segment points. In this study, principal component analysis (PCA) was used to map data into a low-dimensional space and improve the efficiency of linear representation [24]. The algorithm creates a fine approximation of the time series, and interactively merges the lowest cost pair of segments until a stopping criterion is met. When the S_i and S_{i+1} pair of

adjacent segments are merged, the cost of merging the new segment with its right neighbour and the cost of merging the S_{i-1} segment with its new larger neighbour must be calculated. The pseudocode of the algorithm is shown in Table I.

Table I: Bottom-up segmentation algorithm.

Bottom-up segmentation algorithm
● Create initial fine approximation. ● Find the cost of merging for each pair of segments: $Mergecost(i)=cost(S(a_i,b_{i+1}))$ ● While $min(mergecost)<maxerror$ • Find the cheapest pair to merge: $i=argmin_i(mergecost(i))$. • Merge the two segments, update the a_i, b_i boundary indices, and recalculate the merge costs. $mergecost(i)=cost(S(a_i,b_{i+1}))$ $mergecost(i-1)=cost(S(a_{i-1},b_i))$ end

2.2 Random Forest classification model

The Random Forest algorithm is an ensemble learning algorithm of vote classification that combines multiple classification algorithms. It is a combination of Bagging algorithm and Random Subspace algorithm [25]. The basic unit is a decision tree. The accuracy of classification is improved through multiple decision trees $h_1(x)$, $h_2(x)$, ..., $h_{nTree}(x)$, and then a Random Forest classifier is obtained. As shown in Fig. 1. If the regression problems are solved, the average results of all decision trees are output. Suppose the original dataset $T=\{x_{i_1}, x_{i_2}, ..., x_{i_M}, y_i\}_{i=1}^{N}$ contains N samples, where the number of attribute features of $x_{i_1}, x_{i_2}, ..., x_{i_M}$ is M, $Y=\{y_i\}_i^N$ is the category label, and the number of values of y_i is $c \geq 2$, which represents the c category.

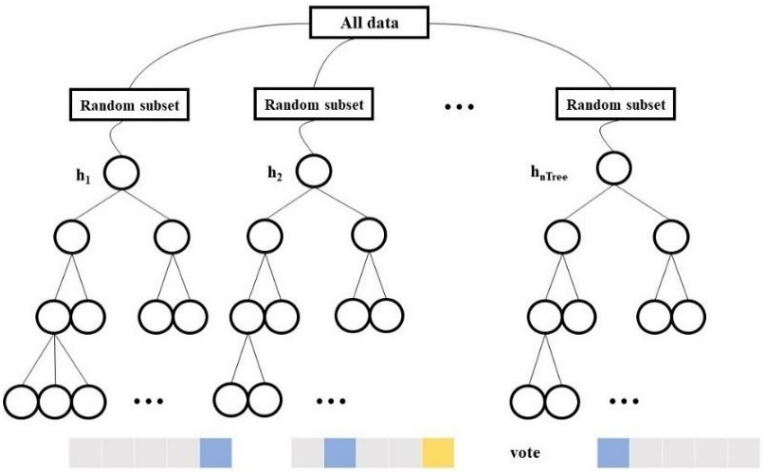


Figure 1: Training and classification process of Random Forest classifier.

Taking the driving behaviour time series features as the model input and the fuel consumption levels as the model output, an eco-driving behaviour identification model was established based on the Random Forest algorithm.

2.3 Evaluation criteria for eco-driving behaviours identification model

It is necessary to select a suitable method to evaluate the accuracy of the identification. A wrapper method based on the area under the ROC curve (*AUC*) is used in the study. The *AUC* is the integrated indicator that reflects the specificity and sensitivity of variables [26]. Also, the prediction performance of the eco-driving behaviour identification model was evaluated by the

following four indicators: recall, precision, and accuracy in the study. These indicators were calculated as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{Precison} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \tag{4}$$

where *TP* is the number of samples that are actually positive and predicted to be positive; *FP* is number of samples predicted to be positive for actual negative; *TN* is the number of samples that are actually negative and predicted to be negative; *FN* is the number of samples that are actually positive predicted to be negative.

3. SIMULATION AND RESULTS

3.1 Simulation platform construction

In this study, an experiment was performed based on a driving simulation platform. The driving simulator equipment was equipped with the Logitech G29 driving set, one 40" 2K HP monitors, and a stereo (shown in Fig. 2 a). And the driving operation software environment was set up with Unreal Engine 4 and the simulation platform was rendered in CALAR. The Logitech G29 driving set included a steering wheel, pedals (clutch, brake, accelerator), and play-seat; the steering wheel of the Logitech G29 has a 900-degree steering range and offers programmable buttons and directional controls. The driving simulator equipment can simulate the real simulated driving environment and provide the real driving feelings for the participants. The data collection system could collect driving behaviour variables such as speed and acceleration, the depth of clutch pedal, the depth of brake pedal, and steering wheel angle at a frequency of 10 Hz.

30 participants, including 20 males and 10 females were recruited from different groups to complete the experiment. In 2020 the number of motor vehicle drivers in China will reach 456 million, of which 308 million are male drivers, accounting for 67.57 %, and there were 148 million female drivers, accounting for 32.43 %; there are 327 million drivers between the ages of 26-50, accounting for 71.79 % of the total number of drivers, and about 60.86 million drivers aged 51 to 60 [27]. Thus, more male drivers were employed; the ages of participants were between 24 and 60 (with mean = 42.36 years). All the participants held valid driver's licenses and drove an average of 6.9 hours per week, and had normal physical conditions.

a) driving simulator b) simulated driving scenario

Figure 2: Experiment details.

The four-lane passing through the urban area was selected as the test route to conduct the experiment. Considering the energy consumption and emissions of vehicles are greatly affected by traffic flow and weather, there were no other vehicles on the road except the tested vehicle, and the weather condition was pretty nice with good visibility. A bird's-eye view of the experimental scene is shown in Fig. 2 b. As shown in the figure, green is the test route. Four driving tasks of lane change were set for the route; the lane change points in test route are marked in red. Participants were permitted to accelerate and decelerate randomly.

The experiments were conducted on campus between August 2021 and December 2021. All experiments were conducted in the same laboratory, and the temperature and lighting conditions in the lab were set the same. Two investigators administered the experiment and each participant was trained to follow an identical procedure. As shown in Fig. 3.

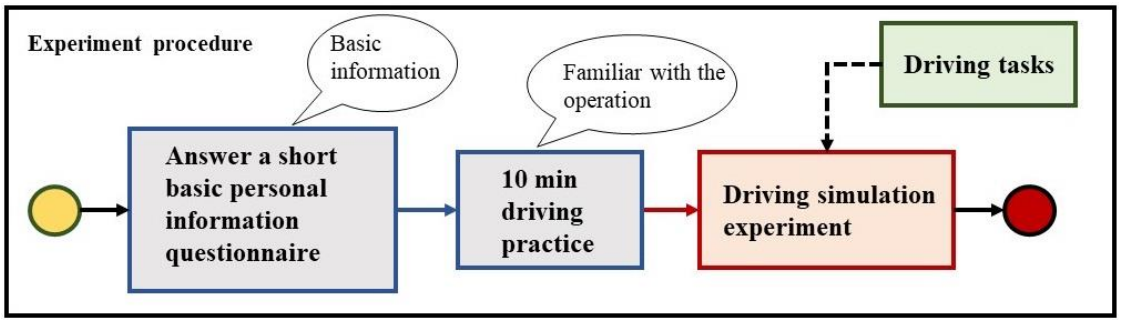


Figure 3: Experiment procedure.

After the pre-experiment, participants were asked to fill out a short questionnaire of basic personal information, which includes age, sex, driving experience, and so on. Then the investigators introduced the experiment procedures and reviewed the inclusion criteria after the participants signed the consent form. Next, the participants practiced driving in the experiment scenario for 10 minutes to be familiar with the operation of the driving simulation device.

Then the formal experiment began. Participants were asked to finish the experiments which included simulated lane change tasks on the test routes. When the participant conducted each lane change task, he was requested to press the red button on the steering wheel; and the system would record the timing of pressing the button as the start timing of the lane change task. The participants were only told about the starting point, destination, and the description of lane change driving task during the experiment.

3.2 Fuel consumption of lane change

This study estimated vehicle exhaust emissions based on the microscopic emissions model according to vehicle specific power (*VSP*) [28, 29], and then calculated vehicle fuel consumption using carbon balance method [30].

Most of the fuel consumption of 100 km in the lane change process was 6 L/100 km ~ 8 L/100 km, with as many as 68 groups. The number of fuel consumption of 100 km that were higher than 8 L/100 km or lower than 6 L/100 km is relatively small, 30 groups and 12 groups, respectively. Further analysis of the fuel consumption of 100 km of the 110 groups of lane change processes showed that the average fuel consumption was 6.81 L/100 km, which conformed to the actual fuel consumption. Therefore, this study divided the fuel consumption of 100 km during the lane change into three levels (low fuel consumption, normal fuel consumption, high fuel consumption). Set as in Table II.

Table II: The symbols and range of driving operation behaviours.

Fuel consumption level	Fuel consumption	Range	Frequency
1	Low fuel consumption	<6	30
2	Normal fuel consumption	[6, 8)	68
3	High fuel consumption	>8	12

3.3 Feature selection of eco-driving behaviours based on Kruskal-Wallis

Based on 110 sets of lane change driving behaviour time series data, this study analysed seven driving behaviours (the depth of acceleration pedal, the depth of clutch pedal, the depth of brake pedal, steering wheel angle, gear position, vehicle lateral position and lane position). Kruskal-Wallis statistical tests were used to assess the effect of driving behaviour variables on fuel consumption. The detailed results are shown in Table III.

Table III: The symbols and range of driving operation behaviours.

Driving behaviour	Level-level	Std. error	Std. test statistic	Adj. sign.
The depth of acceleration pedal	1-2	89.080	-12.618	0.000
	1-3	88.416	-50.262	0.000
	2-3	54.852	-60.525	0.000
The depth of clutch pedal	1-2	50.555	-47.519	0.000
	1-3	50.178	-52.104	0.000
	2-3	31.130	-6.817	0.000
The depth of brake pedal	1-2	41.527	53.359	0.000
	1-3	41.218	57.493	0.000
	2-3	25.571	6.018	0.000
Steering wheel angle	1-2	86.972	-3.699	0.001
	1-3	86.324	-6.225	0.000
	2-3	53.554	-4.026	0.000
Gear position	1-2	73.775	-31.931	0.000
	1-3	73.225	-34.566	0.000
	2-3	45.428	-3.862	0.000
Vehicle lateral position	1-2	89.371	19.607	0.000
	1-3	88.705	20.347	0.000
	2-3	55.005	0.957	1.000
Lane position	1-2	82.241	5.401	0.000
	1-3	81.629	5.967	0.000
	2-3	50.641	0.847	0.800

The test results show that the adjustment significances of the acceleration pedal depth, clutch pedal depth, brake pedal depth, gear position, and steering wheel angle under different fuel consumption levels are all less than 0.05, so there are significant differences between the above five driving behaviour indicators in the three fuel consumption levels. Therefore, five driving behaviour indicators (the depth of acceleration pedal, the depth of clutch pedal, the depth of brake, gear position, and steering wheel angle) were selected to construct the time series that was employed to be analysed.

3.4 Feature selection of multivariate driving behaviour time series based on PLR

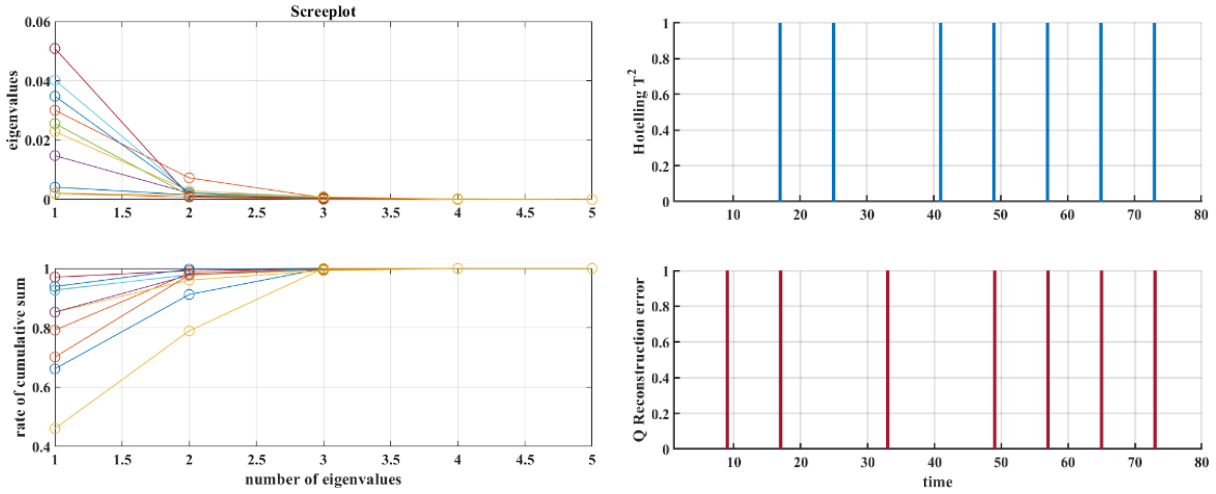
The multivariate time series were composed with the extracted significant driving behaviour indicators (the depth of acceleration pedal, the depth of clutch pedal, the depth of brake pedal, gear position, and steering wheel angle). According to the definition of multivariate time series [31], the multivariate time series consists of vehicle driving behaviour characteristic indicators (vehicle lateral position, gear position, lane position, the depth of acceleration pedal, the depth

of clutch pedal, the depth of brake pedal, steering wheel angle) was obtained as the following Eq. (5):

$$\begin{bmatrix} A \\ C \\ B \\ G \\ W \end{bmatrix} = \begin{bmatrix} a(1) & a(2) & \cdots & a(i) \\ c(1) & c(2) & \cdots & c(i) \\ b(1) & b(2) & \cdots & b(i) \\ g(1) & g(2) & \cdots & g(i) \\ w(1) & w(2) & \cdots & w(i) \end{bmatrix} \tag{5}$$

where $a(i)$ ($i = 1, 2, \ldots, i$) is the depth of acceleration pedal at time i, $c(i)$ ($i = 1, 2, \ldots, i$) is the depth of clutch pedal at time i, $b(i)$ ($i = 1, 2, \ldots, i$) is the depth of brake pedal at time i, $g(i)$ ($i = 1, 2, \ldots, i$) is the gear position at time i, $w(i)$ ($i = 1, 2, \ldots, i$) is the steering wheel angle at time i.

The bottom-up algorithm based on PCA was used to segment the multivariate driving behaviour time series, and the number of principal components was determined by the scree plot shows in Fig. 4 a. The number of principal components of the driving behaviour sequence indicators were input into the bottom-up algorithm. Based on the above works, the multivariate driving behaviour time series composed of significant indicators were divided into segments.



a) screen plots of the driving behaviours b) bottom-up segmentation result

Figure 4: Segmentation result of multivariate time series.

Determined the principal component number (the principal component number of the driving behaviour time series is 3 in Fig. 4 a by the scree plot. Input the number of principal component q into the bottom-up algorithm, and divided each set of time series into 8 subsections to obtain the segmentation results of the multivariate driving behaviour time series. As shown in Fig. 4 b, the eight non-repetitive subsections of the multivariate driving behaviour time series are: $S_1(1, 17)$, $S_2(17, 25)$, $S_3(25, 41)$, $S_4(41, 49)$, $S_5(49, 57)$, $S_6(57, 65)$, $S_7(65, 73)$, $S_8(73, 80)$.

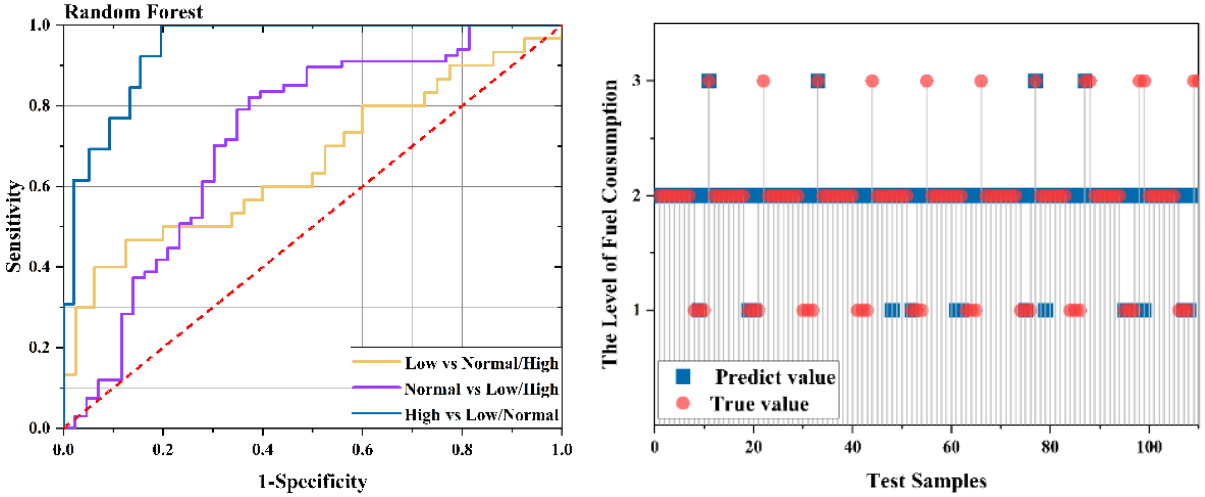
Based on PLR, extracted its slope, time interval, maximum value, minimum value, and expected value as the features to identify the fuel consumption level, then the feature sequence composed of all subsections is as follows:

$$\begin{bmatrix} A^F \\ C^F \\ B^F \\ G^F \\ W^F \end{bmatrix} = \begin{bmatrix} \left(\rho_1^a, \omega_1^a, max_1^a, min_1^a, mean_1^a\right), & \left(\rho_2^a, \omega_2^a, max_2^a, min_2^a, mean_2^a\right), & \cdots, & \left(\rho_m^a, \omega_m^a, max_m^a, min_m^a, mean_m^a\right) \\ \left(\rho_1^c, \omega_1^c, max_1^c, min_1^c, mean_1^c\right), & \left(\rho_2^c, \omega_2^c, max_2^c, min_2^c, mean_2^c\right), & \cdots, & \left(\rho_m^c, \omega_m^c, max_m^c, min_m^c, mean_m^c\right) \\ \left(\rho_1^b, \omega_1^b, max_1^b, min_1^b, mean_1^b\right), & \left(\rho_2^b, \omega_2^b, max_2^b, min_2^b, mean_2^b\right), & \cdots, & \left(\rho_m^b, \omega_m^b, max_m^b, min_m^b, mean_m^b\right) \\ \left(\rho_1^g, \omega_1^g, max_1^g, min_1^g, mean_1^g\right), & \left(\rho_2^g, \omega_2^g, max_2^g, min_2^g, mean_2^g\right), & \cdots, & \left(\rho_m^g, \omega_m^g, max_m^g, min_m^g, mean_m^g\right) \\ \left(\rho_1^w, \omega_1^w, max_1^w, min_1^w, mean_1^w\right), & \left(\rho_2^w, \omega_2^w, max_2^w, min_2^w, mean_2^w\right), & \cdots, & \left(\rho_m^w, \omega_m^w, max_m^w, min_m^w, mean_m^w\right) \end{bmatrix} \tag{6}$$

where ρ_j^{φ} ($\varphi = a, c, b, g, w$, $j = 1, 2, ..., m$) are the slope of the j^{th} subsection; ω_j^{φ} ($\varphi = a, c, b, g, w$, $j = 1, 2,..., m$) are the time length of the j^{th} subsection; max_j^{φ}, min_j^{φ}, $mean_j^{\varphi}$ ($\varphi = a, c, b, g, w$, $j = 1, 2, ..., m$) are the maximum values, minimum values, and average values of the j^{th} subsection.

3.5 The eco-driving behaviour identification model

Based on the extracted 8-segment multivariate time series feature indicators and the Random Forest algorithm, the eco-driving behaviour identification model for lane change was established. 10-fold cross validation was applied to evaluate the classification performance of the eco-driving behaviour identification model, and the predicted and true values are shown in Fig. 5 a. It can be seen from the figure that the accuracy of the eco-driving behaviour identification model is 0.718. 79 samples out of 110 samples are correctly predicted,



a) prediction accuracy b) the ROC curve of Random Forest

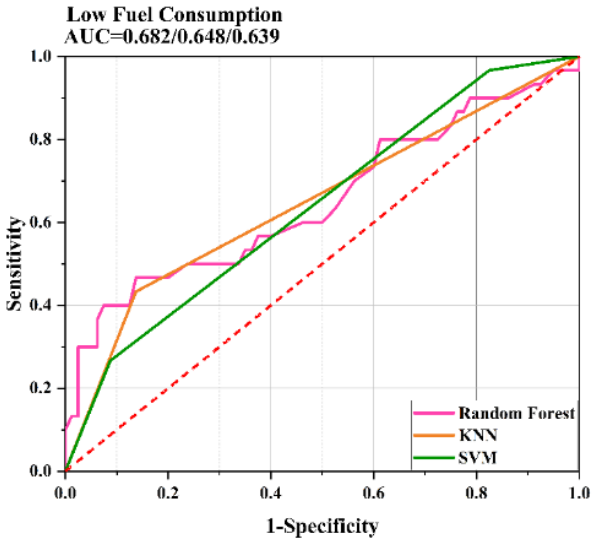
Figure 5: Prediction results of Random Forest.

As shown in Fig. 5 b, The *AUC* values of the area under the ROC curve corresponding to each classification result were 0.682, 0.737, and 0.941, respectively. The results show that the eco-driving behaviour identification model of the lane change process based on the Random Forest model has a good prediction performance on the three different fuel consumption levels.

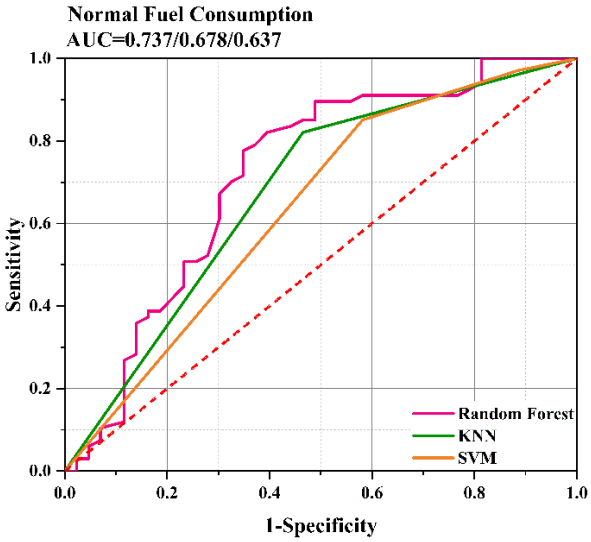
4. DISCUSSION AND CONCLUSION

4.1 Model classification performance analysis

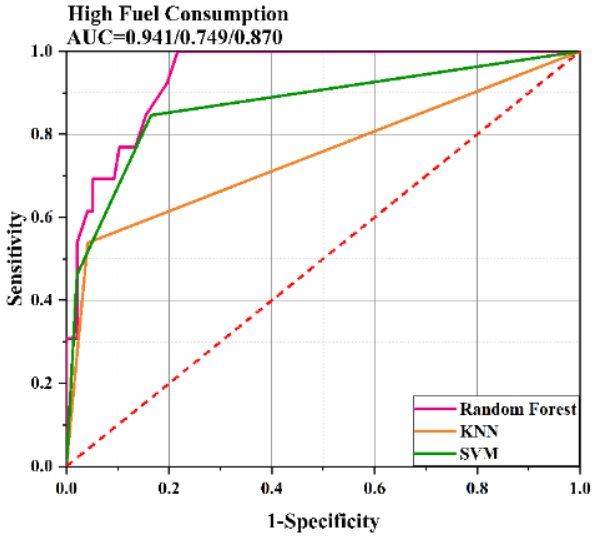
The classifiers based on the Random Forest, KNN, and SVM algorithms also have good prediction performance on the lane change process of the three fuel consumption levels. Among them, the classifier based on the Random Forest algorithm has optimal prediction performance on the lane change processes of the three fuel consumption levels. As shown in Fig. 6, the fuel consumption prediction results of different classifiers for the lane change process of the three fuel consumption levels are respectively drawn. It can be seen from the Fig. 6 d that the Random Forest model was optimal among the three algorithms and was effective enough to be used to identify the eco-driving behaviour of lane change.



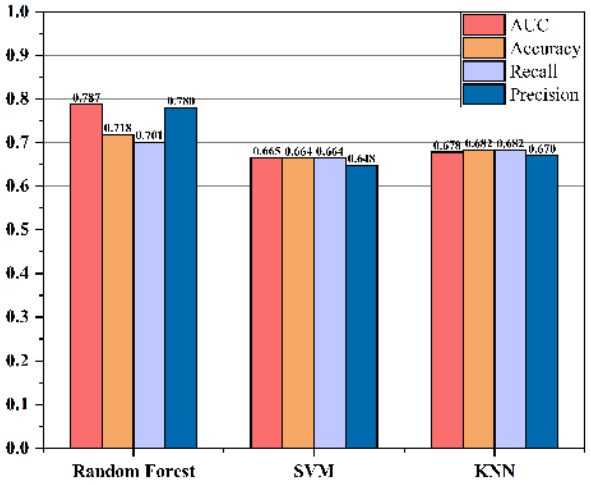
a) the ROC curve of low fuel consumption



b) the ROC curve of normal fuel consumption



c) the ROC curve of high fuel consumption



d) the comparison based on different models

Figure 6: The results of different classification algorithms.

4.2 Analysis of the number of segmentations

The same sample set is segmented, and it can be seen from Fig. 7 that the linear representation of the number of different segmentations of the multivariate time series has a significant impact on the classification performance of the Random Forest model. When the number of segments is 8, Random Forest's classifier works best, and its *AUC*, accuracy, recall, and precision are 0.787, 0.718, 0.701, and 0.780, respectively. Where the accuracy is 0.018, 0.036, and 0.054 higher than that of 5-segment, 10-segment, and 15-segment respectively. Although the accuracy value of the 8-segment linear representation has no obvious advantage over the 5-segment representation, its *AUC*, recall, and precision is in the Random Forest classifiers based on the different segment linear representation features of driving behaviour multivariate time series Highest. Therefore, if the number of time series segments is too large or too small, time series features cannot be extracted well; the lane change eco-driving behaviour identification model based on 8-segment linear representation features and Random Forest has the optimal classification performance.

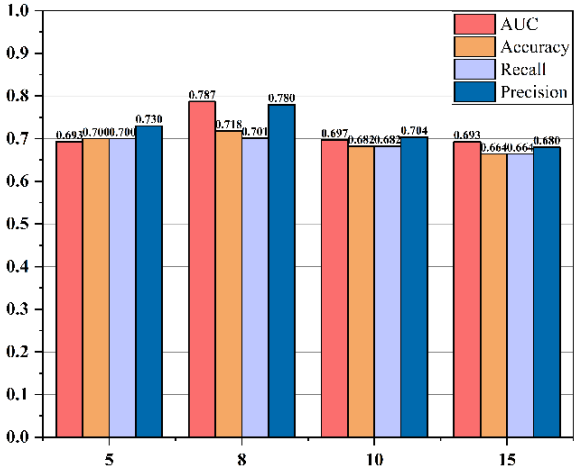


Figure 7: Performance of Random Forest classifier for different numbers of segments.

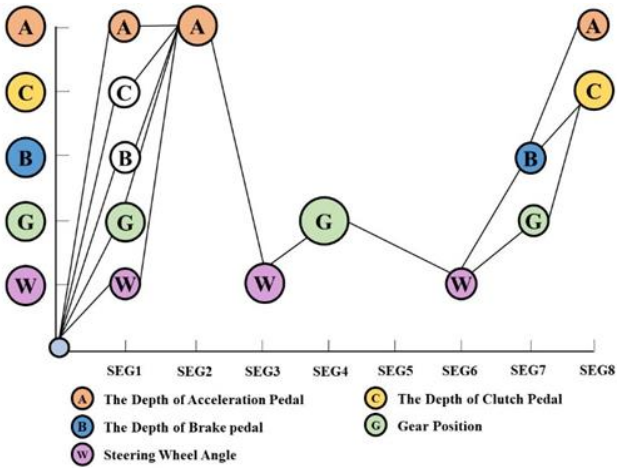


Figure 8: Lane change driving behaviour graph for low fuel consumption.

Combined with the above driving behaviour analysis in each segment and the theory of graph [8], a feature graph of low fuel consumption driving behaviour is drawn, as shown in Fig. 8. The horizontal axis of the coordinate axis is the time axis, and the eight scale points represent eight time series segments in turn; the vertical axis of the coordinate axis represents the driving behaviour. The red solid circle is the acceleration pedal depth symbol, the yellow solid circle is the clutch pedal depth symbol, the blue solid circle is the brake pedal depth symbol, the green is the gear position symbol, and the purple is the steering wheel angle symbol. It can be seen from the figure that during the low fuel consumption lane change driving processes, the vehicle accelerates smoothly, there is no operation that presses the pedal suddenly, the gear position changes smoothly, and the steering wheel angle deflection is stable.

Based on the driving simulation platform, this study extracted the driving behaviour indicators (the depth of acceleration pedal, the depth of clutch pedal, the depth of brake pedal, gear position and steering wheel angle) that have a significant impact on the fuel consumption level and the driving behaviour time series features (slope and time interval) with the best classification performance; and linear representations of driving behaviours base on PLR were used as the input of the model. And the lane change eco-driving behaviour identification model was established based on the Random Forest algorithm. This study also established the eco-driving behaviour identification model based on the SVM algorithm and KNN algorithm for comparison and evaluated the classification performance of the model with the use of four indicators (*AUC*, accuracy, recall, precision). At the same time, established Random Forest models based on different time series segmentation numbers for comparison and selected the optimal segmentation numbers. The result shows that: 1) the variables that have a significant impact on the fuel consumption are the depth of acceleration pedal, the depth of clutch pedal, the depth of brake pedal, gear position, and steering wheel angle; 2) the three algorithms can effectively identify lane change eco-driving behaviours and the accuracies are all above 0.6; among them, the Random Forest model is superior to the other two algorithms; 3) The performance of the fuel consumption identification model based on the 8-segment linear representation feature is superior to that of the small segment number (5-segment) and the larger segment numbers (10-segment, 15-segment). The number of segments has a greater impact on the performance of the identification model. 4) During the low fuel consumption lane change driving process, the drivers do not have the operation that press the pedals urgently, the accelerations and the gear changes are smooth, and the steering wheel angles are deflected smoothly.

The limitations of this research are as follows. The results of this study show that the performance of eco-driving identification model based on the 8-segment linear representation

feature is superior to the small number of segments (5-segment) and the larger number of segments (10-segment, 15-segment). The conclusion indicates that the number of segments of PLR linear representation has a great impact on the classification performance of the model, and the adaptive algorithm for determining the number of segments can be explored in the future. Secondly, this study only established an eco-driving behaviour identification model of the lane change process. In the future, the eco-driving behaviour identification methods of the various driving process can be explored.

ACKNOWLEDGEMENT

This research is supported by the National Nature Science Foundation of China (52162049, 51805169, 52062014, 52062015), Natural Science Foundation of Jiangxi Province (20202BABL212009). This research is also jointly supported by Jiangxi Provincial Major Science and Technology Project – 5G Research Project (Grant No. 20212ABC03A07).

REFERENCES

[1] Button, K.; Vega, H.; Nijkamp, P. (2011). *A Dictionary of Transport Analysis*, Edward Elgar Publishing, Cheltenham
[2] Gense, N. L. J. (2000). *Driving Style, Fuel Consumption and Tail Pipe Emissions: Final Report*, TNO Automotive, Delft
[3] Martin, E.; Nelson, C.; Susan, S. (2012). How eco-driving public education can result in reduced fuel use and greenhouse gas emissions, *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2287, No. 1, 163-173, doi:10.3141/2287-20
[4] Lu, Q.; Tettamanti, T. (2021). Impacts of connected and automated vehicle on freeway with increased speed limit, *International Journal of Simulation Modelling*, Vol. 20, No. 3, 453-464, doi:10.2507/IJSIMM20-3-556
[5] Grznar, P.; Gregor, M.; Gaso, M.; Gabajova, G.; Schickerle, M.; Burganova, N. (2021). Dynamic simulation tool for planning and optimization of supply process, *International Journal of Simulation Modelling*, Vol. 20, No. 3, 441-452, doi:10.2507/IJSIMM20-3-552
[6] Zic, J.; Zic, S. (2020). Multi-criteria decision making in supply chain management based on inventory levels, environmental impact and costs, *Advances in Production Engineering & Management*, Vol. 15, No. 2, 151-163, doi:10.14743/apem2020.2.355
[7] Freile, A. J.; Mula, J.; Campuzano-Bolarin, F. (2020). Integrating inventory and transport capacity planning in a food supply chain, *International Journal of Simulation Modelling*, Vol. 19, No. 3, 434-445, doi:10.2507/IJSIMM19-3-523
[8] Zhao, X. F.; Liu, H. Z.; Lin, S. X.; Chen, Y. K. (2020). Design and implementation of a multiple AGV scheduling algorithm for a job-shop, *International Journal of Simulation Modelling*, Vol. 19, No. 1, 134-145, doi:10.2507/IJSIMM19-1-CO2
[9] Wu, Y.; Zhao, X. H.; Chen, C.; Yao, Y.; Rong, J. (2017). Research on eco-driving behavior characteristics identification and feedback optimization method, *Journal of Transportation Engineering*, Vol. 17, No. 4, 1-6, doi:10.13986/j.cnki.jote.2017.04.001
[10] Wu, Y.; Zhao, X. H. (2018). A graph based method to describe individual driving behaviour, *Journal of Transportation Engineering*, Vol. 18, No. 1, 13-17, doi:10.13986/j.cnki.jote.2018.01.003
[11] Daun, T. J.; Braun, D. G.; Frank, C.; Haug, S.; Lienkamp, M. (2013). Evaluation of driving behavior and the efficacy of a predictive eco-driving assistance system for heavy commercial vehicles in a driving simulator experiment, *Proceedings of the 16th IEEE Conference on Intelligent Transportation Systems*, 2379-2386, doi:10.1109/ITSC.2013.6728583
[12] Huang, A. Q.; Zhang, Y. Q.; He, Z. F.; Hua, G. W.; Shi, X. L. (2021). Recharging and transportation scheduling for electric vehicle battery under the swapping mode, *Advances in Production Engineering & Management*, Vol. 16, No. 3, 359-371, doi:10.14743/apem2021.3.406

[13] Wahl, H.-G.; Holzpfel, M.; Gauterin, F. (2014). Approximate dynamic programming methods applied to far trajectory planning in optimal control, *Proceedings of the 2014 Intelligent Vehicles Symposium*, 1085-1090, doi:10.1109/IVS.2014.6856459
[14] Wang, W. H.; Zhang, W.; Guo, H.; Bubb, H.; Ikeuchi, K. (2011). A safety-based behavioural approaching model with various driving characteristics, *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 6, 1202-1214, doi:10.1016/j.trc.2011.02.002
[15] Pauwelussen, J.; Feenstra, P. J. (2010). Driver behavior analysis during ACC activation and deactivation in a real traffic environment, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 11, No. 2, 329-338, doi:10.1109/TITS.2010.2043099
[16] Lindgren, A.; Angelelli, A.; Mendoza, P. A.; Chen, F. (2009). Driver behaviour when using an integrated advisory warning display for advanced driver assistance systems, *IET Intelligent Transport Systems*, Vol. 3, No. 4, 390-399, doi:10.1049/iet-its.2009.0015
[17] Zhao, X. H.; Yao, Y.; Wu, Y. P.; Chen, C.; Rong, J. (2016). Prediction model of driving energy consumption based on PCA and BP network, Vol. 16, No. 5, 185-191, doi:10.3969/j.issn.1009-6744.2016.05.028
[18] Dantzig, G. B.; Wolfe, P. (1960). Decomposition principle for linear programs, *Operations Research*, Vol. 8, No. 1, 101-111, doi:10.1287/opre.8.1.101
[19] Desrochers, M.; Soumi, F. (1989). A column generation approach to the urban transit crew scheduling problem, *Transportation Science*, Vol. 23, No. 1, 1-13, doi:10.1287/trsc.23.1.1
[20] Potthoff, D.; Huisman, D.; Desaulniers, G. (2008). Column generation with dynamic duty selection for railway crew rescheduling, *Transportation Science*, Vol. 44, No. 4, 493-505, doi:10.1287/trsc.1100.0322
[21] Sanguinetti, A.; Kurani, K.; Davies, J. (2017). The many reasons your mileage may vary: toward a unifying typology of eco-driving behaviors, *Transportation Research Part D: Transport and Environment*, Vol. 52, Part A, 71-84, doi:10.1016/j.trd.2017.02.005
[22] Yang, X.; Wu, C.; He, Y.; Lu, X. Y.; Chen, T. (2022). A dynamic rollover prediction index of heavy-duty vehicles with a real-time parameter estimation algorithm using NLMS method, *IEEE Transactions on Vehicular Technology*, Vol. 71, No. 3, 2734-2748, doi:10.1109/TVT.2022.3144629
[23] Chang, P.-C.; Fan, C.-Y.; Liu, C.-H. (2009). Integrating a piecewise linear representation method and a neural network model for stock trading points prediction, *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, Vol. 39, No. 1, 80-92, doi:10.1109/TSMCC.2008.2007255
[24] Abonyi, J.; Feil, B.; Nemeth, S.; Arva, P. (2005). Modified Gath-Geva clustering for fuzzy segmentation of multivariate time-series, *Fuzzy Sets and Systems*, Vol. 149, No. 1, 39-56, doi:10.1016/j.fss.2004.07.008
[25] Gislason, P. O.; Benediktsson, J. A.; Sveinsson, J. R. (2006). Random forests for land cover classification, *Pattern Recognition Letters*, Vol. 27, No. 4, 294-300, doi:10.1016/j.patrec.2005.08.011
[26] Feng, G.; Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data, *Accident Analysis & Prevention*, Vol. 61, 3-9, doi:10.1016/j.aap.2012.06.014
[27] Ministry of Public Security of the People's Republic of China. In 2020, 33.28 Million New Energy Vehicles Were Registered Nationwide, Reaching 4.92 Million, from *https://app.mps.gov.cn/gdnps/pc/content.jsp?id=7647257*, accessed on 15-04-2022
[28] Lu, H.; Song, G.; Yu, L. (2016). A comparison and modification of car-following models for emission estimation, *Proceedings of the Transportation Research Board 95th Annual Meeting*, Paper 16-3329, 23 pages
[29] Guo, D.; Wang, Y.; Zou, G.; Gao, S.; Tan, D.; Shao, J. (2012). A study on the correlation between VSP and vehicle emission based on on-board emission test, *Automotive Engineering*, Vol. 34, No. 1, 18-21, doi:10.3969/j.issn.1000-680X.2012.01.005
[30] Zhao, X.; Wu, Y.; Rong, J.; Zhang, Y. (2015). Development of a driving simulator based eco-driving support system, *Transportation Research Part C: Emerging Technologies*, Vol. 58, Part D, 631-641, doi:10.1016/j.trc.2015.03.030
[31] Michael, E. (2011). Graphical modelling of multivariate time series, *Probability Theory and Related Fields*, Vol. 153, No. 1, 233-268, doi:10.1007/s00440-011-0345-8