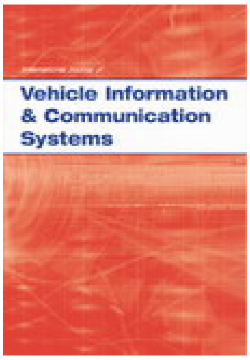


International Journal of Vehicle Information and Communication Systems

ISSN online: 1741-8208 - ISSN print: 1471-0242
https://www.inderscience.com/ijvics

Driving style recognition of highway-driving semi-trailer at different altitudes

Ke Liang, Huasheng Chen, Yu Ye, Song Zhang, Mingzhang Pan

DOI: 10.1504/IJVICS.2023.10058625

Article History:

Received:	02 December 2022
Last revised:	14 April 2023
Accepted:	20 April 2023
Published online:	25 January 2024

Driving style recognition of highway-driving semi-trailer at different altitudes

Ke Liang

College of Mechanical Engineering,
Guangxi University,
Nanning, Guangxi, China
and
Guangxi Key Laboratory of Manufacturing
System & Advanced Manufacturing Technology,
College of Mechanical Engineering,
Guangxi University,
Nanning, Guangxi, China
Email: 20200035@gxu.edu.cn

Huasheng Chen

College of Mechanical Engineering,
Guangxi University,
Nanning, Guangxi, China
Email: 2011391005@st.gxu.edu.cn

Yu Ye and Song Zhang

Guangxi Yuchai Machinery Group Co., Ltd.,
Yulin, Guangxi, China
Email: yeyu@yuchai.cn
Email: tjuzhs@126.com

Mingzhang Pan*

College of Mechanical Engineering,
Guangxi University,
Nanning, Guangxi, China
Email: pmz@gxu.edu.cn
*Corresponding author

Abstract: Driving style provides information about driving behaviour and the driving environment, which reflects the driver's operation while driving. High altitudes can significantly influence the human body, thereby affecting driving ability. Consequently, accurately recognising driving styles at different altitudes has significant implications for driving safety, road design and fuel economy. This paper proposes a method that incorporates data processing, feature selection, a Bi-LSTM autoencoder and spectral clustering to address this issue. Based on the analysis of real-driving data experiments, three driving styles were identified as calm, moderate and aggressive. These styles accounted for 46%, 19% and 36% in plateau driving and 33%, 29% and 38% in plain driving. The results demonstrate how the proposed method can effectively recognise driving styles at different altitudes with fewer features. Additionally, driving styles remained relatively consistent for the same driver driving at varying altitudes, despite changes in vehicle performance.

Keywords: driving style recognition; whale optimisation algorithm; feature selection; Bi-LSTM; autoencoder; spectral clustering.

Reference to this paper should be made as follows: Liang, K., Chen, H., Ye, Y., Zhang, S. and Pan, M. (2024) 'Driving style recognition of highway-driving semi-trailer at different altitudes', *Int. J. Vehicle Information and Communication Systems*, Vol. 9, No. 1, pp.34–59.

Biographical notes: Ke Liang received the BS, MS and PhD degrees in Mechanical Engineering from Tianjin University, Tianjin, China, in 2007, 2011 and 2020, respectively. He currently teaches at the School of Mechanical Engineering, Guangxi University. His research interests include medical minimally invasive surgical robotics and artificial intelligence.

Huasheng Chen received the MS degree in Mechanical Engineering from Guangxi University, Nanning, China, in 2023. His research interests include artificial intelligence, intelligent networked vehicle.

Yu Ye is an engineer from Guangxi Yuchai Machinery Group Co., Ltd. His research focuses on intelligent networked vehicle.

Song Zhang received the PhD degree in School of Automotive Studies from the Tongji University, Shanghai, China, in 2012. From 2012 to 2021, he was a senior engineer in Guangxi Yuchai Machinery Co., Ltd. Since 2021, he has been engaged in research and development of new energy power with Yuchai Synland New Energy Power Technology Co., Ltd. His research interests mainly include intelligent control, integrated development of powertrain about HEV/EV.

Mingzhang Pan received the PhD degree in Power Machinery and Engineering from the Tianjin University, Tianjin, China, in 2014. From 2014 to 2016, he worked as a postdoctoral researcher at Tianjin University. From 2017 to 2019, he worked as a postdoctoral researcher at Zhejiang University. Since 2016, he has been engaged in teaching and scientific research at the School of Mechanical Engineering, Guangxi University. His research interests mainly include intelligent equipment control, medical robot, and advanced power technology.

1 Introduction

Driving style is a complex concept influenced by several factors, leading to various definitions. Typically, driving style consists of driving abilities and driving behaviour. More specifically, driving abilities pertain to the capacity of drivers to manage their vehicles, which is linked to their psychological health, knowledge, skills and experience. Driving abilities are related to the driving experience the driver has (Jachimczyk et al., 2018; Milleville-Pennel and Marquez, 2020). There are interrelationships between driving behaviour and factors such as road environment, real-time traffic situation, etc. (Bellini et al., 2020; Wang et al., 2018). Road safety, fuel efficiency and passenger comfort can be adversely affected by driving behaviour. Determining driving behaviour enables an increase in safety awareness, energy conservation and passenger comfort. As a result, research on driving style has gained significant attention in recent years.

With increasing altitude, air turns thin, resulting in a certain extent of psychological impact on the driver. Additionally, lessened vision, dynamism shifts and reaction times are liable to affect human behaviour and health (Zhang et al., 2022). Low-oxygen levels at high altitudes can affect the visual sensitivity and saccade amplitude of drivers. Measuring heart rate variability via LF/HF was employed to assess changes in the driver's heart rate while driving in a plateau and it was found that the tension of the sympathetic nerve was affected by high altitudes and prolonged driving (Li and Eli, 2015). As a matter of fact, driving at high altitudes decreased the drivers' stress reaction ability, leading to increased errors in judgement and a weaker stress response to driving decisions. Furthermore, examining the fluctuating rate of heartbeat intervals indicated local drivers had greater fatigue resistance than non-local drivers amidst rising altitudes (Chen and Eli, 2016; Liu et al., 2016a). Therefore, the study of the driver's behaviour and performance of the vehicle is necessary in light of the influence on the human body during driving in plateau.

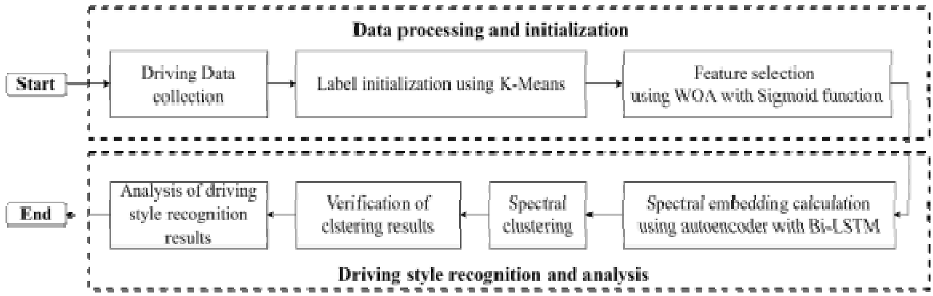
Diesel engines exhibit several phenomena such as a reduction in power and an increase in fuel consumption in plateau, which greatly limit their performance. At altitudes exceeding 4000 m, the power output of diesel engines declines by 14% while fuel consumption increases by approximately 10%, compared to lower altitudes (Guo et al., 2011; Liu et al., 2016b). Owing the reduction in air pressure and oxygen content, combustion was retarded and incomplete. At high altitudes, the ignition delay is prolonged, causing the rate of pressure to increase even exceeding the maximum allowed limit. The engine's lifespan may have been shortened as a result. Moreover, the reduction of excess air ratio and gas density in the cylinder, poor spray information, and insufficient preparation of the mixture resulted in increased emissions and power loss in the engine (Liu and Liu, 2022). Numerous studies have found that driving in high-altitude regions has a negative impact on the human body, specifically on control abilities and vehicle performance, which is distinct from driving in plain. Recognising and distinguishing between these different driving styles is crucial in evaluating road design and performance. There was a wide use of real-time data analysis in the recognition of driving styles. The unsupervised machine learning method, *K*-Means clustering, is commonly used by researchers exploring driving style recognition. Driving manoeuvres were grouped using *K*-means clustering, and risk indexes were calculated to determine the propensity towards aggressive driving behaviour after classifying road types into highways and urban areas (Ma et al., 2021; Martinelli et al., 2018). Further, *K*-Means was employed to classify driving conditions into hilly roads, start/stop and turning cycles and flat roads in order to evaluate the effect of driving conditions on driving behaviour (Si et al., 2018). As part of a study aimed at determining aggressive driving profiles, a two-stage clustering approach based on *K*-Means was used to first distinguish between aggressive and non-aggressive drivers, and then to identify distraction and risk taking in both stages (Mantouka et al., 2019). An algorithm based on hierarchical clustering and Principal Component Analysis (PCA) was used to classify driving styles into five categories using a GPS-based vehicle tracking system called Gipix (Constantinescu et al., 2010). By combining algorithms and data analysis, driving style was recognised efficiently. A supervised method that includes learning and prediction from labelled data enables vehicle owners to distinguish potential impostors by using attributes stored in the car's embedded sensor (Martinelli et al., 2020). A semi-supervised machine learning method called Semi-Supervised Support Vector Machine (S3VM) has been developed to identify aggressive and normal driving styles based on a few labelled data points (Wang et al., 2017). The use of neural networks, however, is widespread in a variety of tasks, such as classification, image recognition, voice recognition and behaviour prediction. For driving style recognition, an Artificial Neural Network (ANN) model was developed to identify longitudinal and lateral driving manoeuvres using data collected from inertial sensors. Based on a score calculated by this model, the driving style was classified into one of five categories (Brombacher et al., 2017). The advancement of machine learning has simplified the process of distinguishing between diverse driving styles. However, large amounts of data affect the calculation and accuracy of a model, feature selection can be used to address this issue.

Feature selection is the process of selecting uniform, non-redundant and essential features related to machine learning models that helps to reduce data set complexity while conserving information hidden within it. PCA was a typical method that used for feature selection (Si et al., 2018; Wang et al., 2022; Xie et al., 2018) With this method, features were extracted from 383 dimensions of the electronic vehicle big data based on the percentage of intervals, three different joint distribution characteristics in which the cumulative contribution rate of the first 35 principal components was over 85% and sufficient for representing the driving style (Xia and Kang, 2021).

Driver behaviour and vehicle status exhibit significant differences at varying altitudes, making it necessary to distinguish their respective driving styles. Since machine learning can be inefficient and expensive when used with large quantities of data, it may not be practical in some situations. Natural driving data consists of continuous time series information, meaning that a driver's behaviour at a particular moment is linked to their behaviour before and after that moment. As a result, it is necessary to consider the states preceding and following each instance when identifying driving style.

This paper proposes a driving style spectral clustering recognition method using autoencoder with Bi-LSTM. In the initial stage, the original labels are determined from the cleaned data set using the *K*-Means clustering algorithm. Subsequently, the whale optimisation algorithm combined with Sigmoid function is employed to reduce the size of the data set while selecting essential features. To learn the structure of spectral embedding, the auto-encoder with Bi-LSTM is used and the final driving style label is determined using spectral clustering. Finally, the feasibility of the proposed methodology is demonstrated through experimental analysis of semi-trailer driving at different altitudes. Figure 1 depicts the framework of the proposed method.

Figure 1 Framework of driving style recognition method



2 Method

2.1 Label initialisation and feature selection

The *K*-Means method, an unsupervised machine learning technique, is extensively utilised in data mining and cluster analysis. It aims to group data points with similar features, dividing the data set into non-overlapping clusters, where each data point belongs to only one group. This algorithm seeks to maximise inter-cluster differentiation while minimising intra-cluster differences. The method assigns data points to clusters based on the shortest Euclidean distance between their centroids and the data points. In this study, *K*-Means was employed to establish the initial labels that served as input for feature selection.

This study evaluates the effectiveness of *K*-Means clustering by using the Calinski-Harabasz Index (CHI) (Chamidah and Wasito, 2015; Mewada et al., 2020). The core calculation of CHI, as illustrated in equation (1), involves the assessment of both the inter-cluster and intra-cluster variances to determine the score. The larger the CHI value, the better the clustering effect.

$$s(k) = \frac{tr(B_k)}{tr(W_k)} \cdot \frac{(m-k)}{(k-1)} \tag{1}$$

where B_k is the between-cluster covariance matrix and W_k is the within-cluster covariance matrix, tr is the trace of a matrix, m is the total number of observations and k is the total number of clusters.

One of the meta-heuristic algorithms, the Whale Optimisation Algorithm (Mirjalili and Lewis, 2016) is inspired by the hunting strategy of whales. The algorithm could be divided into three phases which are discussed in the following.

Encircling prey is the first phase of the algorithm. Initially, the algorithm selects the first prey through a random search process. The WOA algorithm presumes that the first prey is the best candidate or the one closest to it, referred to as the target prey.

Subsequently, the swarming behaviour adjusts the positions of the prey towards the candidate solution, which can be expressed as follows:

$$\mathbf{D} = \left|\mathbf{C} \cdot \mathbf{X}^{*}(t) - \mathbf{X}(t)\right| \tag{2}$$

$$\mathbf{X}(t+1) = \mathbf{X}^{*}(t) - \mathbf{A} \cdot \mathbf{D} \tag{3}$$

where t is the current iteration, X is the position vector, X^{*} is the position vector coincide to the best solution found, A and C are coefficient vectors which are defined as following:

$$\mathbf{A} = 2\mathbf{a} \cdot \mathbf{r} - \mathbf{a} \tag{4}$$

$$\mathbf{C} = 2 \cdot \mathbf{r} \tag{5}$$

where a is linearly decreased from 2 to 0 over the course of iterations (in both exploration and exploitation phases) and r is a random vector in [0,1].

The second phase is exploitation phase called Bubble-net attacking. A shrinking encircling mechanism is applied at the beginning of this phase. This behaviour is achieved by reducing the value of A in the equation (4). Subsequently, the whale's distance to its prey is calculated using the spiral updating position method. The spiral equation mimics the helix-shaped movement of the whale and is expressed as follows:

$$\mathbf{X}(t+1) = \mathbf{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \mathbf{X}^{*}(t) \tag{6}$$

where l is a random number in the range [−1, 1] and b is a constant. 50% of the probability was assumed to choose between either shrinking encircling mechanism or the spiral model. The mathematical model is following:

$$X(t+1) = \begin{cases} X^{*}(t) - A \cdot D & p < 0.5 \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X^{*}(t) & p \geq 0.5 \end{cases} \tag{7}$$

where p is a random number in [0, 1].

As the final phase of the algorithm, the exploration phase utilises values represent by A that do not fall within the range of [–1, 1]. This forces the search agent to move a considerable distance from its current position. Mathematically expressed as follows:

$$\mathbf{D}_{rand} = \left|\mathbf{C} \cdot \mathbf{X}_{rand} - \mathbf{X}\right| \tag{8}$$

$$\mathbf{X}(t+1) = \mathbf{X}_{rand} - \mathbf{A} \cdot \mathbf{D}_{rand} \tag{9}$$

where X_{rand} is a random position vector (a random whale) chosen from the current population. D_{rand} denotes the distance from a randomly selected individual whale to its prey.

As for the feature selection, the continuous WOA has to be transforming to their corresponding binary space (Hussien et al., 2019; Mafarja et al., 2020). The conversions proposed in the study is the sigmoidal transfer function which forces the search agent moving in the binary space. The transfer equation used to get the continuous form and define it as given in equation (10).

$$S\left(\Delta X_t\right) = \frac{1}{1+e^{-\Delta X_t}} \tag{10}$$

where ΔX_t represents the step vector of search space at t. Then, the equation for current search agent position updating is given in equation (11).

$$X_{t+1}^{d}\left(t+1\right) = \begin{cases} 1 & \text{if} \quad rand < S\left(\Delta X_{t+1}\right) \\ 0 & \text{if} \quad rand \geq S\left(\Delta X_{t+1}\right) \end{cases} \tag{11}$$

where the $rand$ indicates a random number in (0, 1).

The objective of feature selection is to identify the minimum number of features that should be selected while achieving the highest possible classification accuracy. The two objectives are combined and converted into a single objective problem as shown in equation (12). The minimum fitness score is determined by adding the minimum error rate in classification to the minimum number of features that are selected.

$$Fitness = \lambda \times E_r + \eta \frac{\left|S_l\right|}{\left|F_l\right|} \tag{12}$$

$$\lambda + \eta = 1 \tag{13}$$

where E_r is the classification error rate, S_l and F_l are the length of the selected feature subset and the number of all features, respectively. λ and η are the classification accuracy and the importance degree of the length of the feature subset, respectively. In this paper, a value of $\lambda = 0.99$ is used to satisfy the fitness function

During the iterative process, the fitness value of each solution is continually calculated. The subset that has the lowest fitness value is the optimal solution. The classification accuracy can be calculated based on the optimal solution using the following formula:

$$Accurancy = 1 - E_r \tag{14}$$

To avoid highly linear-correlated features that may lead to overfitting in subsequent calculations, this paper uses the Pearson correlation coefficient to measure the linear correlation between pairs of features. By identifying and removing highly correlated features within the data set, the model's generalisation ability and performance are improved, thereby allowing it to concentrate on the most informative features. The Pearson correlation coefficient is calculated using equation (15).

$$\begin{aligned} \rho_{X,Y} &= corr\left(X,Y\right) = \frac{cov\left(X,Y\right)}{\sigma_X \sigma_Y} \\ &= \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \end{aligned} \tag{15}$$

where, $cov(X,Y)$ is the covariance of X and Y, σ_X and σ_Y are the standard deviation, $\overline{X}$ and $\overline{Y}$ are the mean value of X_i and Y_i, respectively. The range of $\rho_{X,Y}$ falls within the interval [–1, 1]. The strength of the linear relationship between features is positively correlated with a value of $\rho_{X,Y}$ closer to 1 or –1, and negatively correlated with a value of $\rho_{X,Y}$ closer to 0.

2.2 Autoencoder with Bi-LSTM and spectral clustering

A Recurrent Neural Network (RNN) is a specialised type of neural network that excels in processing sequence data. In RNN, the previous information is retained and used to inform the current output, through a series of organically connected units between hidden layers. However, the long-distance dependence on prior information, determined by the computational characteristics of RNN, results in its learning ability declining over time, which makes it difficult to achieve the expected training targets. To counteract this, Long Short-Term Memory (LSTM) utilises memory modules such as input gates, output gates and forget gates within the hidden layers, allowing for the storage and transmission of information over extended periods of time. As a result, the model can store and transfer information over extended periods, which will be further explained in the subsequent paragraph.

In the LSTM unit, the forget gate chooses which information to discard, while the input gate decides which new information to store. The output gate determines which information to retain and get passed to the next layer. The calculation of gates can be described as follows:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{16}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{17}$$

$$O_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{18}$$

The following equations describe the information that stored in the gates and update the state of cells.

$$\tilde{C}_t = \tanh\left(W_C \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{19}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{20}$$

Finally, the output gate produces the result of the multiplication between the output candidate and the current cell state.

$$h_t = o_t * tanh\left(C_t\right) \tag{21}$$

The Bidirectional Short-Term Memory (Bi-LSTM) stores information in both forward and backward directions based on the LSTM model, at each sequence step. The Bi-LSTM maintains two hidden states, one from past to future, and another from future to past, enabling preservation of complete past and future sequence information at any point in time.

The autoencoder model consists primarily of encoders and decoders, complemented by the nonlinear feature extraction capability of deep neural networks, whose purpose is to transform input into intermediate variables, convert those variables into output and then compare input and output to minimise the differences between them. This approach is particularly suited for Bi-LSTM.

Spectral clustering algorithms are derived from the spectral graph partition theory (Von Luxburg, 2007). The key idea of the algorithm is that it uses the eigenvalues of special matrices built from the graph or the data set. An essential component of spectral clustering is the use of optimal partition of graphs in order to solve the clustering problem, solving the singularity problem associated with high-dimensional feature vectors because the size of the data set is the only determinant, and the dimension of the data set does not play a role. In different algorithms, the splitting criterion function and spectral mapping method may differ, but the basic framework remains the same.

3 Experiment and discussion

To validate practicability of the proposed algorithm, a semi-trailer was chosen for analysis purposes. Based on the GPS data, the vehicle was mostly driven in plains while only a few days were spent driving in plateau. Data was collected by On-Board Diagnostic (OBD), and no communication was made with the driver about the study. To distinguish the driving styles at different altitudes, two separate routes, one for plains and another for plateau regions, are employed as depicted in Figure 2. Both routes were selected based on highways-dominated roads.

Figure 2 Route of experiment at different altitudes

(a) Plateau (b) Plain

3.1 Raw data preparation

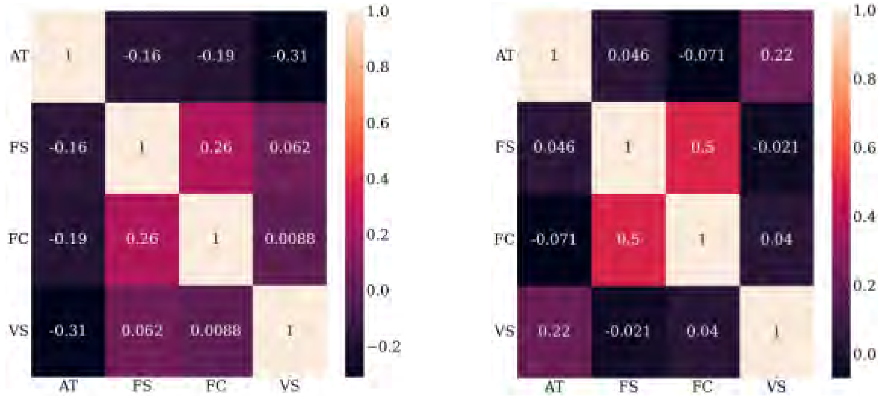
To conduct this study, the fields correlated with driving style in Table 1 were selected for analysis. The raw data were partially lost due to the instabilities of the sensors, and a box plot method was employed to mitigate the impact of data exceptions (Frigge et al., 1989; Hubert and Vandervieren, 2008). The study excluded the continuous and long-term parking periods, where stopping exceeds 180 seconds, from our analysis to ensure minimal negative impact on the accuracy and effectiveness of the driving style recognition. The raw data were captured at a 1 Hz sampling rate, depicting driving actions that change within seconds. The data set for plateau and plain regions comprised 26,747 and 49,358 valid driving data, respectively.

Table 1 Fields chose for driving style recognition.

Fields	*Definition*	*Range*
Altitudes (m)	Altitude of the vehicle	[0, 5000]
Boost Pressure (hPa)	Average boost air pressure at the booster air outlet	[0, 5000]
Throttle Position (%)	Throttle opening degree	[0, 100]
State of brake	Brake switch status	0: deactivate 3: activate
Ambient Pressure (hPa)	Atmospheric temperature at the site of engine operation.	[0, 1000]
Ambient Temperature (°C)	Atmospheric pressure value at the site of engine operation.	[–45, 55]
Engine Speed (rpm)	Number of revolutions per minute at which the engine crankshaft turns.	[0, 3500]
Fan Speed (rpm)	Speed of fan around the tank	[50, 5000]
Fuel consumption (L)	Amount of fuel consumed per unit time.	\
Mileage (m)	Distance travelled per unit time.	\
Injected fuel quantity (mg/hub)	The amount of fuel injected for each operating process of a single cylinder of the engine.	[0, 250]
Gear Position	Position of gear lever	[0, 12]
Vehicle speed (km/h)	Speed of the vehicle	[0, 120]
Slope (rad)	The change in slope of the vehicle travelled per unit time.	$[0, \frac{\pi}{2}]$

As mentioned above, *K*-Means method was applied in this study to initialise the label of the raw data, and three clusters were assigned to the algorithm, indicating the driving styles of calm, moderate and aggressive (Mohammadnazar et al., 2021; Deng et al., 2022). The labelled data was imported into the whale optimisation algorithm combined with a sigmoid function for the purpose of feature selection using 16 search agents and 70 iterations. The accuracy for feature selection in the plateau data was 97.34%, while for the plain data, it was 98.19%. To ascertain the correlation between the selected features, we used the Pearson correlation coefficient method. The results indicated that engine speed and vehicle speed had a correlation coefficient of 0.55 and 0.64, respectively, for plateau and plain regions. Furthermore, the throttle position and fuel consumption had a correlation coefficient of 0.88 in plateau and 0.96 in plain. Therefore, engine speed and throttle position were eliminated from the model operation to improve its efficiency. After eliminating, Ambient Temperature (AT), Fan Speed (FS), Fuel Consumption (FC) and Vehicle Speed (VS) were selected as features both in plateau and plain data analysis. Figure 3 shows the correlation matrix of selected features after eliminating with heatmap at different altitudes, which can be used to analyse the impact on the driving styles.

Figure 3 Correlation matrix of selected features with heatmap at different altitudes (see online version for colours)



(a) Plateau (b) Plain

Two Bi-LSTM layers with 32 units were implemented in the autoencoder architecture, acting as both encoders and decoders in order to mitigate the issue of overfitting caused by excessive layering. The model was trained using Tensorflow with 80 epochs and a batch size of 32 with the selected features. Spectral clustering was then applied. Subsequently, three eigenvalue spaces were selected to describe the data, each corresponding to a specific driving style. This is a final step, the embedded vectors were clustered using *K*-means and the results of the clustering were mapped back to the original data set.

To verify the validity of the driving style recognition results using Bi-LSTM for the parameter calculation of spectral clustering. Self-Organising Map (SOM) (Lakshminarayanan, 2020), widely used unsupervised neural network, was used to evaluate the CHI clustering effect with the proposed model. To avoid the generation of empty clusters and to ensure the usability of SOM model clustering results for driving style clustering, this study utilised five SOM topologies, specifically 1×3, 1×2, 1×4, 1×5 and 2×2. The comparison results of SOM and Bi-LSTM spectral clustering shows in Table 2.

Table 2 Validation of clustering method

Model		*CHI*	
		Plateau	*Plain*
SOM	1×2	3151.923	9490.043
	1×3	3381.910	9015.197
	1×4	3816.304	9853.867
	1×5	3267.544	13318.178
	2×2	3812.867	9853.867
Bi-LSTM- spectral clustering		109449.682	15772.193

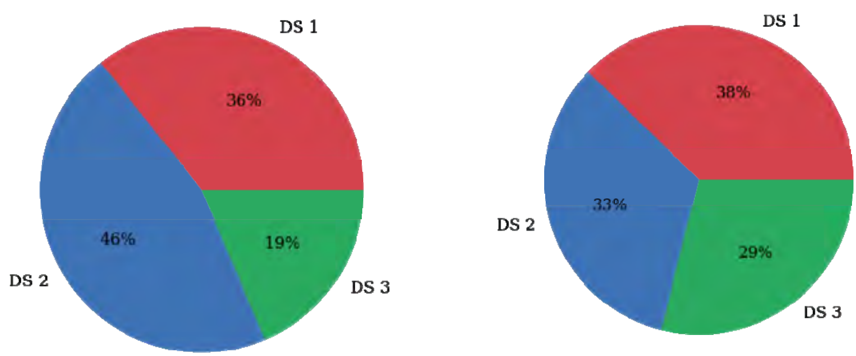
The experimental results show that the effectiveness of SOM clustering is significantly lower than the Bi-LSTM clustering. The primary reason for this is that SOMs often generate anomalies in the map, where two similar groupings appear in different areas of the map. However, modifying the topology of SOM reduces the interpretability of clustering results. In contrast, spectral clustering can find well-separated clusters even when the data is not well-structured. Besides, the Bi-LSTM can use past and future information of driving data, which is useful for driving style classification in continuous driving behaviour. Therefore, this paper used the Bi-LSTM method to determine the spectral embedding and apply spectral clustering in driving behaviour analysis.

As previously mentioned, the proposed algorithm divides driving style into three sections, specifically DS1, DS2 and DS3. Table 3 illustrates the mean values of the selected features for each driving style, while Figure 4 represents the distribution of each category at different altitudes. To analyse driving styles, multiple fields were compared and evaluated.

Table 3 Numerical definition of each cluster in plateau (plain)

Cluster	*Definition*	*Ambient temperature (℃)*	*Fan speed (rpm)*	*Fuel consumption (L)*	*Vehicle speed (km/h)*
DS1	Aggressive	23.07(28.3)	1238.74(465.6)	119.17(101.48)	75.91(72.23)
DS2	Calm	23.3(29.21)	1098.39(426.09)	58.2(89.69)	72.56(71.3)
DS3	Moderate	22.86(25.44)	1171.89(476.49)	70.63(98.87)	58.53(71.57)

Figure 4 Results of driving style recognition at different altitudes (see online version for colours)



(a) Labels of driving in plateau (b) Labels of driving in plain

3.2 Analysis of driving style in the whole route

The distribution of gear positions in Figure 5 indicates that the driving route was on the highway, as the gear position was mainly 12. Regardless of the vehicle's location, there were no significant variations in driving styles when low gear positions were being used. The distribution of Throttle Position (TP) in Figure 6 shows that [0, 10) of throttle position was dominant due to gear shifting and vehicle cruising associated with highway driving. Specifically, the distribution of TP is more symmetrical when compared to driving in plateau, which indicates that the experienced driver driving in familiar surroundings had a greater ability to control the throttle. Driving on plateau had an adverse impact on the driver's body function due to high altitudes, causing a decrease in the control ability of the throttle. As a result, the driving style tends to be less aggressive when driving at high altitudes. Additionally, the TP was mainly below 60% while driving in plateau, implying that drivers tend to be more cautious when driving in high altitudes.

Figure 5 Distribution of gear position at different altitudes (see online version for colours)

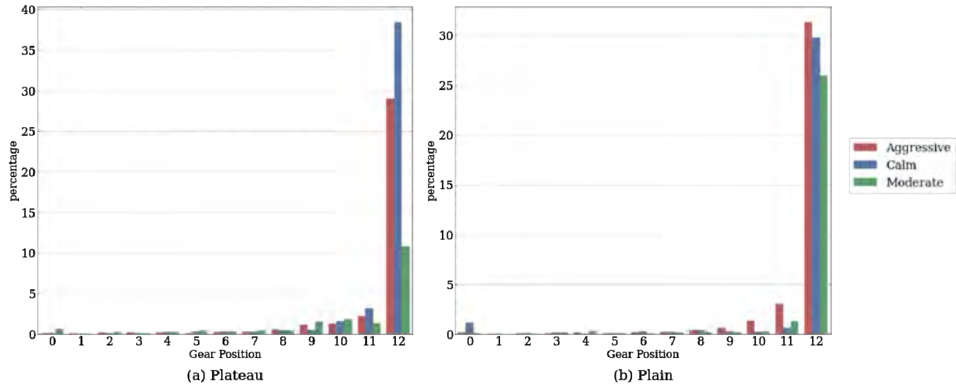


Figure 6 Distribution of throttle position at different altitudes (see online version for colours)

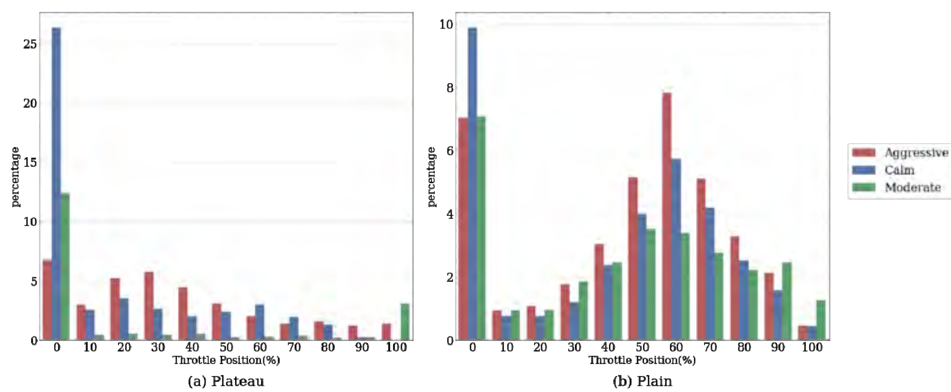


Figure 7 shows the distribution of gear position when holding brake at different altitudes. The main difference between driving on plains and plateau when holding the brake was that driving styles in plateau tended to be calm or moderate, with the majority of gear position being either 0 or 12. When driving in plain, aggressive gear shifting behaviour mostly occurred at higher gear positions, indicating that the vehicle remained relatively unsteady when holding the brake during the trip. This instability was due to the complicated highway conditions encountered, such as changing lanes and changing slopes.

Figure 8 shows the distribution of engine speed with label of driving style in different altitudes. The results indicate that calm and moderate driving styles were dominant regardless of the altitude. Aggressive driving styles accounted for 35.72% and 37.3% of the driving time in the plateau and plain, respectively, which were related to adjustments to throttle position and gear position. Only a small quantity of engine speed was found to be within the range of [1800, 2000), indicating that the diesel engine produced different performance in plateau driving situations.

Figure 7 Distribution of gear position with brake holding at different altitudes (see online version for colours)

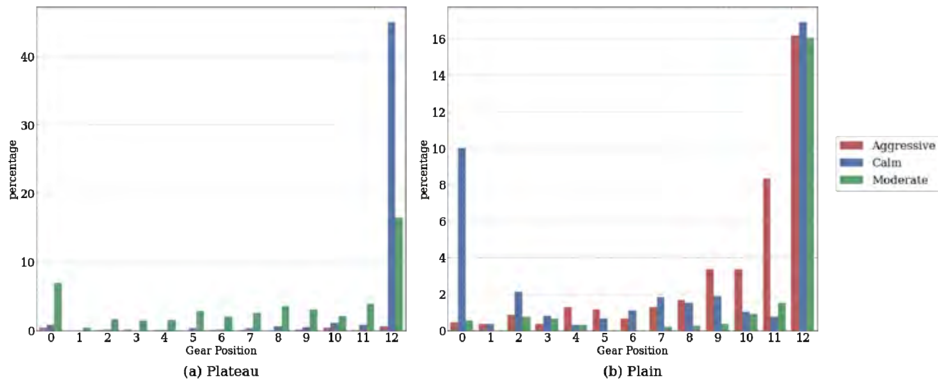


Figure 8 Distribution of engine speed at different altitudes (see online version for colours)

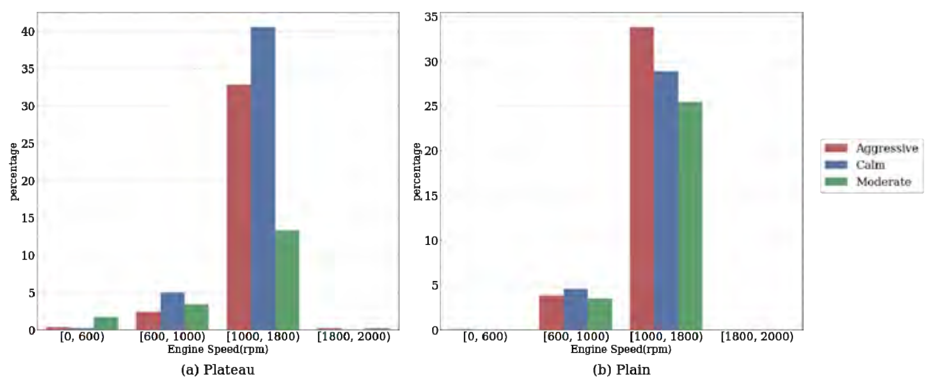


Figure 9 shows the distribution of vehicle speed with labels indicating driving styles at different altitudes. In total, with the speed limit of 100 km/h for semi-trailer, 82.68% of vehicles driving in plain maintained speeds above 60 km/h, while only 75.37% of vehicles did so in the plateau. These results indicate that drivers were more conservative when driving in high altitudes, particularly when they are not familiar with the driving environment. It is believed that small numbers of overspeed indicate a reduction in the control ability of the driver in plateau or the driver may be relaxed when a well-conditioned highway is applied with fewer vehicles (Hu and Yang, 2010). However, the aggressive driving style mainly appeared at high speed regardless of the location, suggesting that high-speed driving increases the driver's intensity, especially in plateau. Additionally, the complexity of road conditions leads to changes in throttle position, gear shifting, and vehicle speed, which can affect the recognition of driving styles.

Figures 10(a) and 10(b) show the correlation between vehicle speed and brake state at different altitudes. No matter what altitudes the vehicle was at, the distribution of vehicle speed with brake releasing is similar with the distribution of vehicle speed in the whole trip which demonstrated the state of holding brake was the minority in the trip. When the brake was held, the distribution of vehicle speed was more dispersed and drivers tended to be more conservative when driving in plateau compared to plain. Specifically, the aggressive driving style rarely appeared in plateau driving, while holding the brake resulted in a 39.23% incidence rate of aggressive driving style in plain driving. Notably, holding the brake typically occurred at high-vehicle speeds when driving on the plains, indicating that these vehicles were handling complicated road conditions such as overtaking.

Figure 9 Distribution of vehicle speed at different altitudes (see online version for colours)

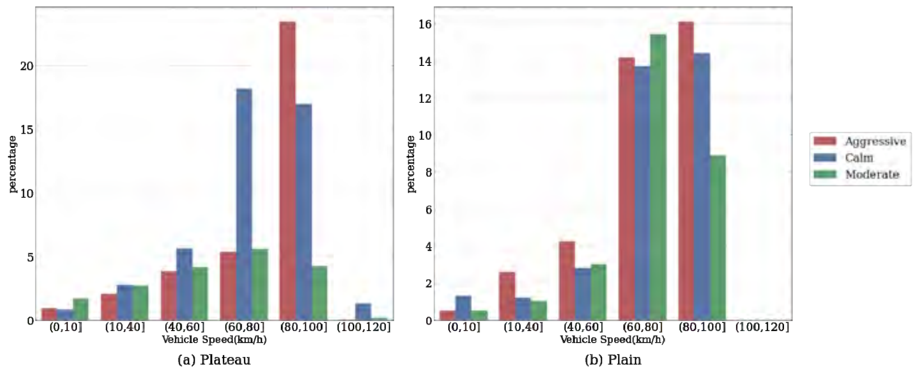
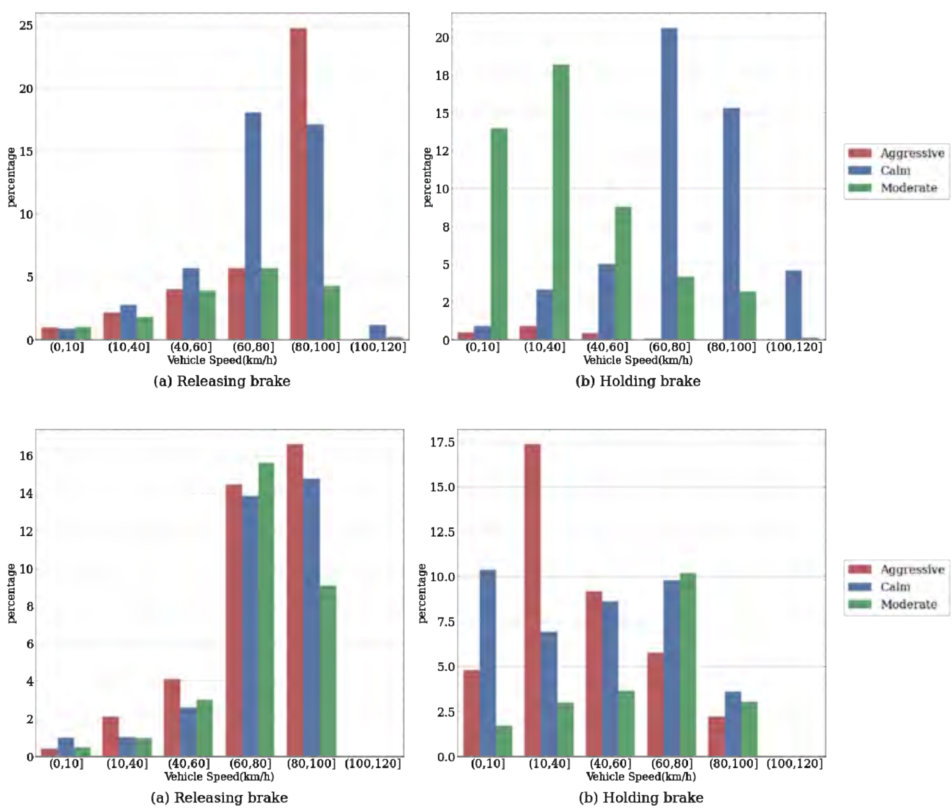


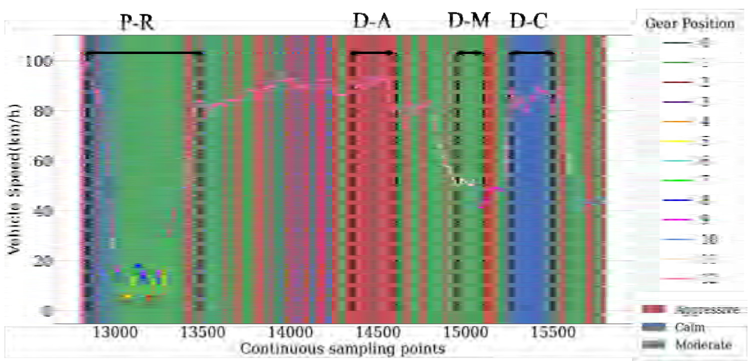
Figure 10 Distribution of vehicle speed with brake state in plateau (above), and distribution of vehicle speed with brake state in plain (below) (see online version for colours)



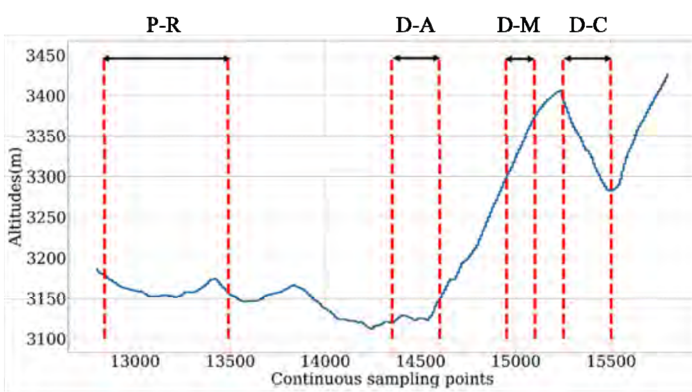
3.3 Analysis of driving style in continuous samples

To validate the recognition and definition of driving styles, 3000 continuous sample points were selectively examined to demonstrate the relationship between the data and the driving style. Figure 11(a) shows the relations among gear state, vehicle speed and driving style. When there is a drastic change in vehicle speed along with high gear position, driving style is often transformed into the aggressive one. On the contrary, with smooth changing of vehicle speed, gear shifting changes with gentle slope and is categorised as calm or moderate, regardless of the location. Three typical fragments were chosen in order to illustrate the differences in driving styles. D-A, D-M, D-C and P-R represent periods that are dominated by driving style of aggressiveness, driving style of moderate, driving style of clam and parking and restarting, respectively.

Figure 11 (a) Change on relations in gear state, vehicle speed and driving style when driving in plateau (b) Changes of altitudes on continuous sampling points when driving in plateau (c) Changes of state of brake on continuous sampling points when driving in plateau (d) Changes of throttle position on continuous sampling points when driving in plateau (e) Changes of engine speed on continuous sampling points when driving in plateau (f) Changes of ambient temperature on continuous sampling points when driving in plateau (g) Changes of fan speed on continuous sampling points when driving in plateau (see online version for colours)

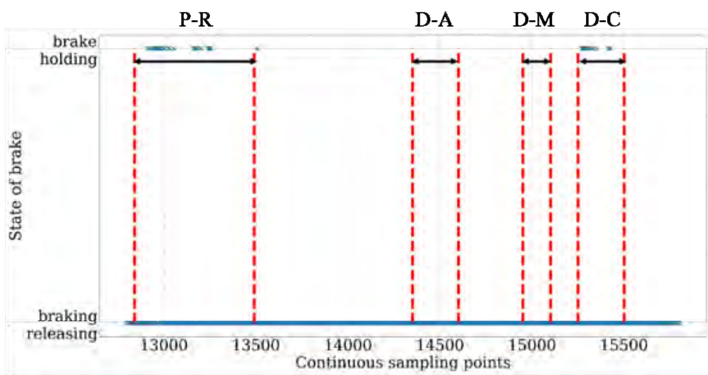


(a)

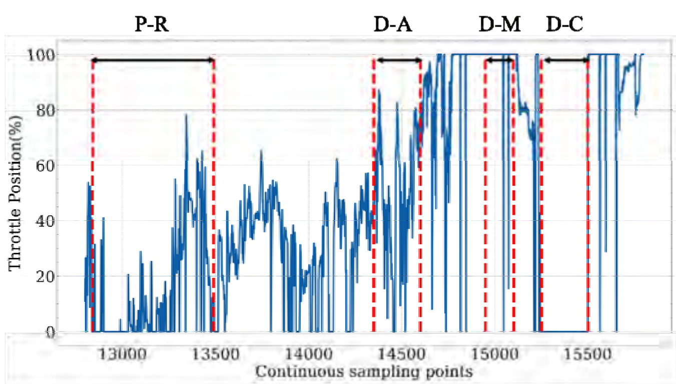


(b)

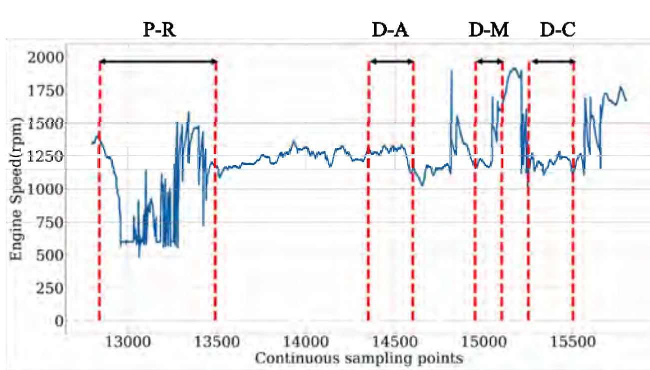
Figure 11 (a) Change on relations in gear state, vehicle speed and driving style when driving in plateau (b) Changes of altitudes on continuous sampling points when driving in plateau (c) Changes of state of brake on continuous sampling points when driving in plateau (d) Changes of throttle position on continuous sampling points when driving in plateau (e) Changes of engine speed on continuous sampling points when driving in plateau (f) Changes of ambient temperature on continuous sampling points when driving in plateau (g) Changes of fan speed on continuous sampling points when driving in plateau (continued) (see online version for colours)



(c)

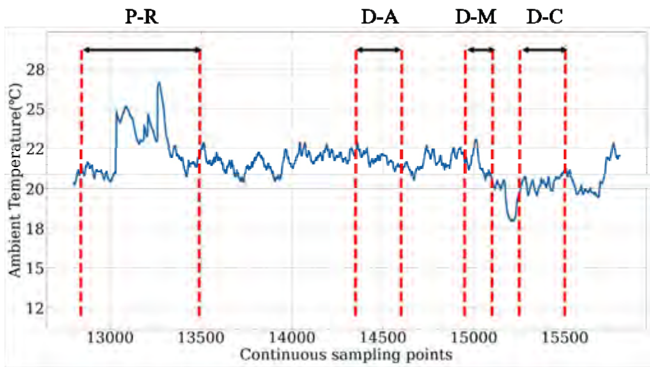


(d)

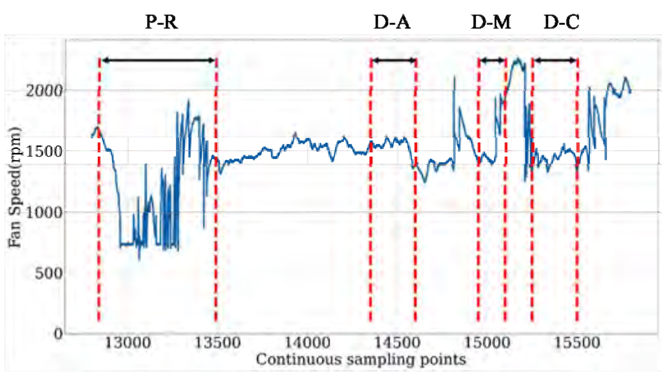


(e)

Figure 11 (a) Change on relations in gear state, vehicle speed and driving style when driving in plateau (b) Changes of altitudes on continuous sampling points when driving in plateau (c) Changes of state of brake on continuous sampling points when driving in plateau (d) Changes of throttle position on continuous sampling points when driving in plateau (e) Changes of engine speed on continuous sampling points when driving in plateau (f) Changes of ambient temperature on continuous sampling points when driving in plateau (g) Changes of fan speed on continuous sampling points when driving in plateau (continued) (see online version for colours)



(f)

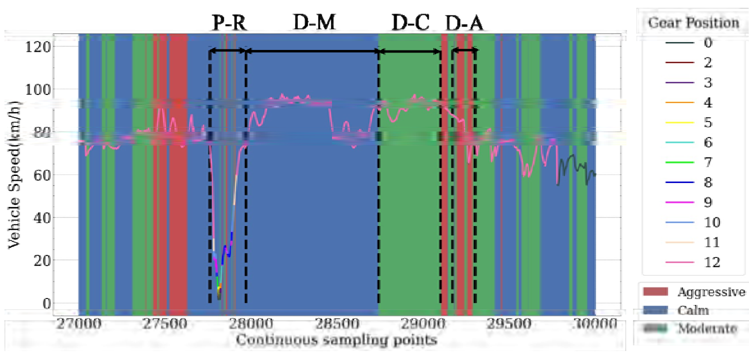


(g)

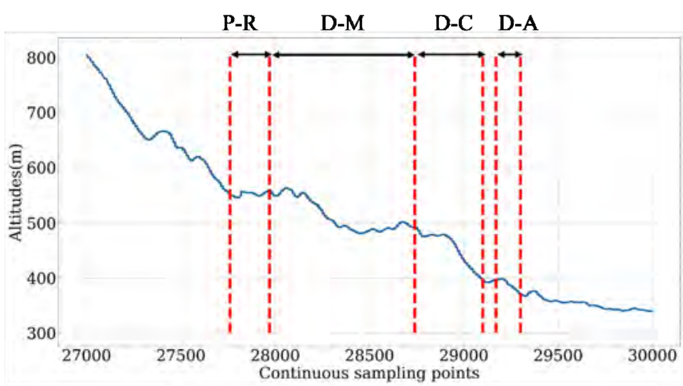
For driving in plateau, the vehicle speed changes considerably, ranging from a minimum of 39.41 km/h to a maximum of 96.26 km/h disregarding P-R periods. In D-A, the driving style frequently becomes aggressive when the vehicle speed undergoes rapid changes in high-gear positions. However, during D-C and D-M, the driving style mostly remains stable despite changes in vehicle speed through proper gear shifting. Figures 11(b) to 11(g) illustrate the changes in altitude, brake status, throttle position, engine speed, ambient temperature and fan speed while driving in plateau. For D-C, it can be inferred that the vehicle was moving downhill based on Figure 11(b). In such circumstances, to prevent overdriving, the driver regulated the speed of the vehicle by controlling the braking pedal, while maintaining a relatively steady engine speed, as illustrated in Figures 11(c), 11(d) and 11(e). Furthermore, the ambient temperature remained relatively constant, causing the fan speed to fluctuate less frequently, as demonstrated by Figures 11(f) and 11(g). For D-M, the vehicle was getting upslope according to Figure 11(b). To complete this progress, the driver executed a smooth gear shifting and lowered the vehicle speed to 41.47 km/h. The engine speed varied with gear-shifting, and it is worth noting that the exceptionally high-speed of 1697 rpm resulted in a moderate driving style. Moreover, the peak ambient temperature and fluctuations in fan speed contributed to the recognition of driving styles. For D-A, the vehicle experienced minimal changes in slope, as demonstrated in Figure 11(a). To sustain vehicle movement, the driver kept the gear position at 12 while decreasing speed from 94.26 km/h to 79.37 km/h. During this period, the throttle position was changing rapidly without gear shifting, resulting in severe engine jitter towards the end of this period. This suggests that the driver aimed to maintain a specific speed, as shown in Figures 11(d) and 11(a). For P-R, the majority of the driving style was calm and moderate, while the end of the restarting stage exhibited an aggressive driving style. The sudden increase in ambient temperature occurred as a result of reduced air flowability and the proximity of the temperature sensor to the tank, which caused a significant increase in the fan's speed. Additionally, gear shifting is required for substantial fluctuations in engine speed when adjusting vehicle speed. This operation did not have an adverse impact on the vehicle's performance at low speed, resulting in a comparatively moderate driving style. Taking into account the vehicle's previous and current speeds, the high-acceleration rate at the end of the period was identified as aggressive driving style.

Changes in altitudes, state of brake, throttle position, engine speed, ambient temperature and fan speed when driving in plain were shown in Figures 12(a) to 12(f), respectively. Among drivers driving in plain, the calm and moderate driving style was prevalent, indicating a tendency to drive in a stable state in a familiar driving environment. For D-C, the vehicle was getting downslope according to Figure 12(b). Despite complex highway conditions, the driver competently manoeuvred the semi-trailer during D-C. The vehicle maintained a consistent speed throughout, reaching a speed of 97.76 km/h at maximum and 72.57 km/h at minimum speeds, with the gear remaining fixed at position 12. Although frequent changes in throttle position with less control of brake pedal, the gear position remained steady and the engine speed fluctuated between 1000 rpm and 1500 rpm, as shown in Figures 12(c), 12(d) and 12(e). This suggests that the driver was able to retain control of the vehicle even getting downslope. The driver tended to be calm and adjusted the performance of the vehicle in time to adapt the change of road condition. Moreover, when the ambient temperature did not fluctuate much, the calm driving style displayed no abrupt changes. However, there was a sudden peak in the fan speed, as depicted in Figures 12(f) and 12(g), indicating the effectiveness of the proposed method. For D-M, the gentle rate of slope change that allowed the driver to adjust the throttle position continuously, thereby maintaining a consistent vehicle speed while matching the 12th gear position and without encountering sudden changes in engine speed. However, the most significant difference between D-M and D-C was the three peaks in fan speed. Compared with D-C, the altitudes decreased quickly during D-A. It's worth noting that driving style was divided into aggressive with a relatively rapid decrease in speed without gear shifting during D-A. The engine speed also changed rapidly from 630 to 1339 rpm without gear shifting, and the vehicle speed suddenly decreased. This indicates that the road conditions had changed.

Figure 12 (a) Changes of altitudes on continuous sampling points when driving in plain (b) Changes of altitudes on continuous sampling points when driving in plain (c) Changes of state of brake on continuous sampling points when driving in plain (d) Changes of throttle position on continuous sampling points when driving in plain (e) Changes of engine speed on continuous sampling points when driving in plain (f) Changes of ambient temperature on continuous sampling points when driving in plain (g) Changes of fan speed on continuous sampling points when driving in plateau (see online version for colours)

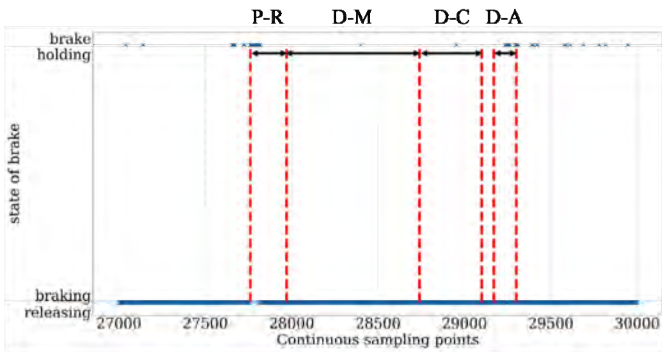


(a)

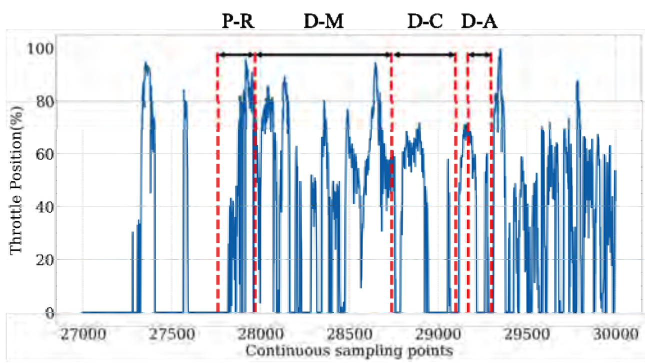


(b)

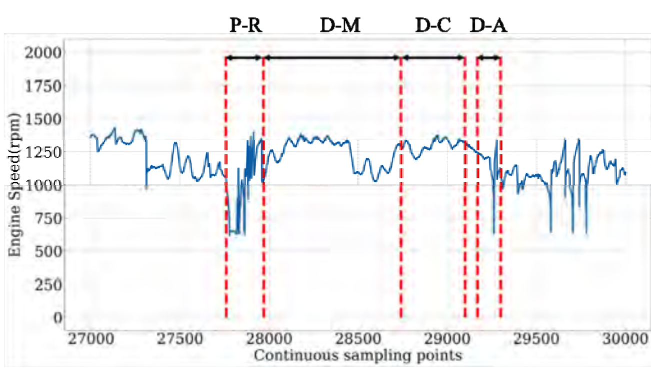
Figure 12 (a) Changes of altitudes on continuous sampling points when driving in plain (b) Changes of altitudes on continuous sampling points when driving in plain (c) Changes of state of brake on continuous sampling points when driving in plain (d) Changes of throttle position on continuous sampling points when driving in plain (e) Changes of engine speed on continuous sampling points when driving in plain (f) Changes of ambient temperature on continuous sampling points when driving in plain (g) Changes of fan speed on continuous sampling points when driving in plateau (continued) (see online version for colours)



(c)

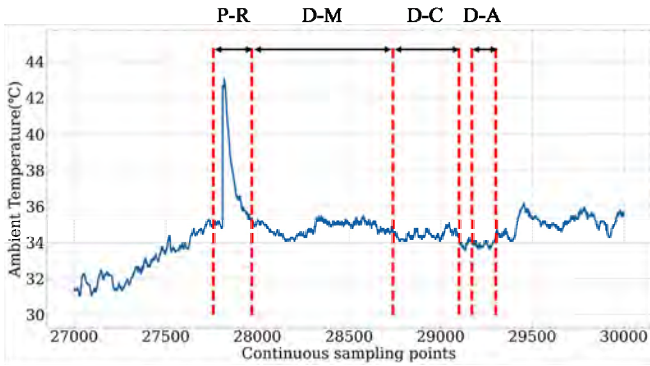


(d)

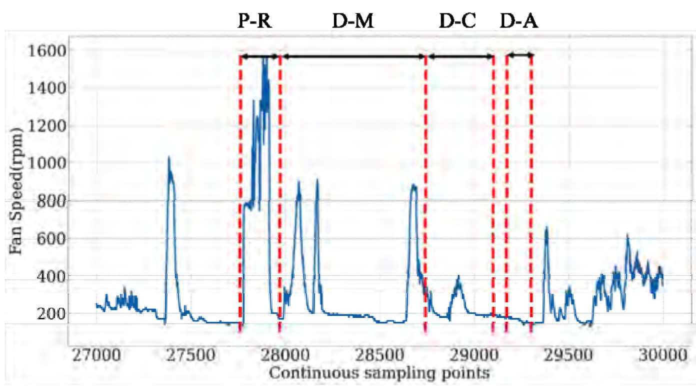


(e)

Figure 12 (a) Changes of altitudes on continuous sampling points when driving in plain (b) Changes of altitudes on continuous sampling points when driving in plain (c) Changes of state of brake on continuous sampling points when driving in plain (d) Changes of throttle position on continuous sampling points when driving in plain (e) Changes of engine speed on continuous sampling points when driving in plain (f) Changes of ambient temperature on continuous sampling points when driving in plain (g) Changes of fan speed on continuous sampling points when driving in plateau (continued) (see online version for colours)



(f)



(g)

To be specific, Table4 shows the difference between P-R periods at various altitudes. The mean ambient temperature of driving in plateau was 22.60°C, a 63.85% reduction compared to that in plain, while the mean fan speed was 984.61 rpm, only 1.13% lower than in plain, which could be attributed to the calm driving style. In plateau, the mean throttle position remained low and the engine speed and vehicle speed remained high, whereas the gear position change rate was lower, resulting in a moderate driving style, rather than a calm one. For safety reasons, the mean and change rate of the throttle position in the plateau were 14.71 and 2.87, respectively, which is significantly lower than in plain. Overall, during P-R period, the driving style was assigned to calm with proper control ability over vehicle speed, engine speed and gear shifting which correspond with the recognition results in Table 3.

Table 4 Difference in P-R at different altitudes

	Altitudes(m)		Throttle position (%)		State of brake		Ambient temperature ( )		Fan speed (rpm)		Engine speed (rpm)		Gear position		Vehicle speed (km/h)	
	Value	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*	*Value*	*Change rate*
Driving in plateau(duration=650s)																
Mean	3160.43^{++}	0.11	14.71	2.87	0.41	0.17	22.60	0.06	984.61	30.93	1200.08^{+}	38.83	8.58	0.22	39.36	0.46
std*	7.46^{+}	0.44	19.23	6.06	1.03	0.69	1.59	0.21	312.82^{+}	65.51	375.32	78.91	3.29^{+}	1.02	30.62^{+}	0.61
Minimum	3151^{++}	0	0	0	0	0	20.46	0	486	0	600^{+}	0	0	0	1	0
Maximum	3179^{++}	3	78.2715	44.7021	3	3	26.66	3.9	1582^{+}	421.5	1920^{+}	510	12	12	98.37^{+}	6.06
Range	28	3	78.27	44.7	3	3	6.2	3.9	1096^{+}	421.5	1320	510	12	12	97.37^{+}	6.06
Driving in plain(duration=210s)																
Mean	551.36	0.17^{+}	40.03^{+}	6.97^{+}	0.77^{+}	0.07^{+}	37.03^{+}	0.10^{+}	995.75+	35.43+	772	41.52+	9.50^{+}	0.33	40.12^{+}	0.86^{+}
STD	3.72	0.68^{+}	36.28^{+}	13.99^{+}	1.31^{+}	0.46^{+}	2.41^{+}	0.54^{+}	246.48	70.31+	471.24^{+}	80.48+	2.64	1.31	23.55	0.95^{+}
Minimum	545	0	0	0	0	0	34.76^{+}	0	618^{+}	0	130	0	0	0	1.70^{+}	0
Maximum	558	7.27^{+}	95.96	81.84^{+}	3	3	43.06^{+}	7.80^{+}	1398	459+	1570	580+	12	12	74.52	6.03^{+}
Range	13	7.27^{+}	95.96	81.84^{+}	3	3	8.30^{+}	7.80^{+}	780	459+	1440^{+}	580+	12	12	72.82	6.03^{+}

Notes: Std represents standard deviation;
+ represents the larger one between the same fields at different altitudes;
++ represents the much larger one between the same fields at different altitude

4 Conclusion

The paper presents a method that uses data processing, feature selection and spectral clustering with a neural network to recognise different driving styles at various altitudes, namely plateau and plain. The proposed method employs *K*-Means to initialise the label of the real-world data, and an improved WOA to identify the most relevant feature of the data. The experimental effectiveness of the WOA with Sigmoid function was 97.34% in plateau data, and 98.19% in plain data. An autoencoder with Bi-LSTM is applied to train the model to obtain eigenvalues and eigenvectors that are essential for the spectral embedding. Then, spectral clustering is used to classify driving styles. Based on the real-time data experiment, the driving progress was recognised into three driving styles, which accounted for 46%, 19% and 36% in plateau driving and 33%, 29%, 38% in plain driving. By analysis of the clustering result, the driving styles were recognised as calm, moderate and aggressive respectively except for the parking and restart stage. To validate the reliability of the recognition, the difference in driving styles was demonstrated based on the results of recognition and performance in each collected field. The proposed method can reflect changes in the vehicle while drivers perform different actions, allowing for the inference of driving behaviour, particularly in cases of significant differences in vehicle performance between driving in a plateau and in plain. The method could be further improved by exploring more accurate methods of initialising original data labels and experimenting with diverse highway conditions such as city and countryside areas, as well as various vehicle types such as trucks and coaches in plateau and plain.

Acknowledgements

This work was supported in part by Guangxi Major Scientific and Technological Special Project (Grant No. AA22068103), Guangxi Key Research and Development Program (Grant No. AB201220059), Guangxi Science and Technology Base and Talent Special Project (Grant No. 2021AC19324) and Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology (Grant No. 20-065-40S006).

References

Bellini, D., Iaconis, M.C. and Traettino, E. (2020) 'Speed limits and road warning signs as aid for driving behavior', *Transportation Research Procedia*, Vol. 45, pp.135–142. Doi: 10.1016/j.trpro.2020.02.100.

Brombacher, P., Masino, J., Frey, M. and Gauterin, F. (2017) 'Driving event detection and driving style classification using artificial neural networks', Paper presented at *the IEEE International Conference on Industrial Technology (ICIT)*, Toronto, Canada.

Chamidah, N. and Wasito, I. (2015) 'Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine', Paper presented at *the International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, Depok, Indonesia.

Chen, P. and Eli, I. (2016) 'Analysis of plateau highway driver stress response capability', *China Safety Science Journal*, Vol. 26, No. 4. Doi: 10.16265/j.cnki.issn1003-3033.2016.04.002.

Constantinescu, Z., Marinoiu, C. and Vladoiu, M. (2010) 'Driving style analysis using data mining techniques', *International Journal of Computers Communications and Control*, Vol. 5, No. 5, pp.654–663.

Deng, Z., Chu, D., Wu, C., Liu, S., Sun, C., Liu, T. and Cao, D. (2022) 'A probabilistic model for driving-style-recognition-enabled driver steering behaviors', *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 52, No. 3, pp.1838–1851. Doi: 10.1109/TSMC.2020.3037229.

Doi: 10.1186/s12544-020-00421-x.

Doi: 10.15837/ijccc.2010.5.2221.

Frigge, M., Hoaglin, D.C. and Iglewicz, B. (1989) 'Some implementations of the Boxplot', *The American Statistician*, Vol. 43, No. 1, pp.50–54. Doi: 10.1080/00031305.1989.10475612.

Guo, M., Wang, X. and Chang, Q. (2011) 'Combustion characteristics of turbocharged diesel in different atmosphere environment', *Diesel Engine*, Vol. 33, No. 1, pp.9–13.

Hu, J. and Yang, Y. (2010) 'Driver factor characteristics of traffic safety on the low traffic volume Qinghai-Tibet highway in China', Paper presented at *the International Conference on Optoelectronics and Image Processing*, Haikou, China.

Hubert, M. and Vandervieren, E. (2008) 'An adjusted boxplot for skewed distributions', *Computational Statistics and Data Analysis*, Vol. 52, No. 12, pp.5186–5201. Doi: 10.1016/j.csda.2007.11.008.

Hussien, A.G., Hassanien, A.E., Houssein, E.H., Bhattacharyya, S. and Amin, M. (2019) 'S-shaped binary whale optimization algorithm for feature selection', *Recent Trends in Signal and Image Processing*, Springer, pp.79–87.

Jachimczyk, B., Dziak, D., Czapla, J., Damps, P. and Kulesza, W.J. (2018) 'IoT on-board system for driving style assessment', *Sensors*, Basel, Switzerland, Vol. 18, No. 4. Doi: 10.3390/s18041233.

Lakshminarayanan, S. (2020) 'Application of self-organizing maps on time series data for identifying interpretable driving manoeuvres', *European Transport Research Review*, Vol. 12, No. 1, pp.1–11.

Li, T. and Eli, I. (2015) 'Experiment analysis of driver's HRV on Plateau highway', *Science Technology and Engineering*, Vol. 15, No. 19, pp.6. Doi: 10.3390/s19224982.

Liu, J., Ma, X., Zhang, Z., Guo, Z. and Liu, B. (2016a) 'Fatigue characteristics of driver in Qinghai-Tibet Plateau based on electrocardiogram analysis', *Journal of Traffic and Transportation Engineering*, Vol. 16, No. 4. Doi: 10.19818/j.cnki.1671-1637.2016.04.016.

Liu, S., Chen, Y., Wang, J., Zeng, J. and Han, W. (2016b) 'Effects of altitude on fuel consumption and thermal efficiency of naturally aspirated diesel engines', *Chinese Internal Combustion Engine Engineering*, Vol. 37, No. 6. Doi: 10.13949/j.cnki.nrjgc.2016.06.004.

Liu, Z. and Liu, J. (2022) 'Effect of altitude conditions on combustion and performance of a turbocharged direct-injection diesel engine', *Proceedings of the Institution of Mechanical Engineers: Part D, Journal of Automobile Engineering*, Vol. 236, No. 4, pp.582–593. Doi: 10.1177/09544070211026204.

Ma, Y., Li, W., Tang, K., Zhang, Z. and Chen, S. (2021) 'Driving style recognition and comparisons among driving tasks based on driver behavior in the online car-hailing industry', *Accident Analysis and Prevention*, Vol. 154. Doi: 10.1016/j.aap.2021.106096.

Mafarja, M., Qasem, A., Heidari, A. A., Aljarah, I., Faris, H. and Mirjalili, S. (2020) 'Efficient hybrid nature-inspired binary optimizers for feature selection', *Cognitive Computation*, Vol. 12, No. 1, pp.150–175. Doi: 10.1007/s12559-019-09668-6.

Mantouka, E.G., Barmpounakis, E.N. and Vlahogianni, E.I. (2019) 'Identifying driving safety profiles from smartphone data using unsupervised learning', *Safety Science*, Vol. 119, pp.84–90. Doi: 10.1016/j.ssci.2019.01.025.

Martinelli, F., Mercaldo, F., Nardone, V., Orlando, A. and Santone, A. (2018) 'Cluster analysis for driver aggressiveness identification', Paper presented at *the Proceedings of the 4th International Conference on Information Systems Security and Privacy*, Funchal, Portugal.

Martinelli, F., Mercaldo, F., Orlando, A., Nardone, V., Santone, A. and Sangaiah, A.K. (2020) 'Human behavior characterization for driving style recognition in vehicle system', *Computers and Electrical Engineering*, Vol. 83, pp.102504–102516. Doi: 10.1016/j.compeleceng.2017.12.050.

Mewada, H., Al-Asad, J.F. and Khan, A.H. (2020) '*Landscape Change Detection Using Auto-optimized K-means Algorithm*', Paper presented at *the IEEE International Conference on Industrial Technology (ICIT),* Marrakech, Morocco.

Milleville-Pennel, I. and Marquez, S. (2020) 'Comparison between elderly and young drivers' performances on a driving simulator and self-assessment of their driving attitudes and mastery', *Accident Analysis and Prevention*, Vol. 135. Doi: 10.1016/j.aap.2019.105317.

Mirjalili, S. and Lewis, A. (2016) 'The whale optimization algorithm', *Advances in Engineering Software (1992)*, Vol. 95, pp.51–67. Doi: 10.1016/j.advengsoft.2016.01.008.

Mohammadnazar, A., Arvin, R. and Khattak, A.J. (2021) 'Classifying travelers' driving style using basic safety messages generated by connected vehicles: application of unsupervised machine learning', *Transportation Research: Part C, Emerging Technologies*, Vol. 122. Doi: 10.1016/j.trc.2020.102917.

Si, L., Hirz, M. and Brunner, H. (2018) 'Big data-based driving pattern clustering and evaluation in combination with driving circumstances', *SAE Technical Paper*, Vol. 1087, No. 1.

Von Luxburg, U. (2007) 'A tutorial on spectral clustering', *Statistics and Computing*, Vol. 17, No. 4, pp.395–416. Doi: 10.1007/s11222-007-9033-z.

Wang, J., Besselink, I. and Nijmeijer, H. (2018) '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*, Vol. 232, No. 11, pp.1528–1542.

Wang, S., Yu, P., Shi, D., Yu, C. and Yin, C. (2022) 'Research on eco-driving optimization of hybrid electric vehicle queue considering the driving style', *Journal of Cleaner Production*, Vol. 343. Doi: 10.1016/j.jclepro.2022.130985.

Wang, W., Xi, J., Chong, A. and Li, L. (2017) 'Driving style classification using a semisupervised support vector machine', *IEEE Transactions on Human-Machine Systems*, Vol. 47, No. 5, pp.650–660. Doi: 10.1109/THMS.2017.2736948.

Xia, L. and Kang, Z. (2021) 'Driving style recognition model based on NEV high-frequency big data and joint distribution feature parameters', *World Electric Vehicle Journal*, Vol. 12, No. 3. Doi: 10.3390/wevj12030142.

Xie, H., Tian, G., Chen, H., Wang, J. and Huang, Y. (2018) 'A distribution density-based methodology for driving data cluster analysis: a case study for an extended-range electric city bus', *Pattern Recognition*, Vol. 73, pp.131–143. Doi: 10.1016/j.patcog.2017.08.006.

Zhang, D., Chen, F., Zhu, J., Wang, C., Cheng, J., Zhang, Y., Bo, W. and Zhang, P. (2022) 'Research on drivers' hazard perception in plateau environment based on visual characteristics', *Accident Analysis and Prevention*, Vol. 166. Doi: 10.1016/j.aap.2021.106540.