Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 105 (2017) 2831 – 2836

Energy Procedia

The 8th International Conference on Applied Energy – ICAE2016

Driving cycle development for electric vehicle application using principal component analysis and k-means cluster: with the case of Shenyang,China

Wenyu Zhou,Ke Xu, Ying Yang* and Jiahuan Lu

School of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819,China;

Abstract

Driving cycle is essential for the investigation of power management in electrified vehicles. Building a representative driving cycle remains a challenge due to the complex driving conditions. In this paper, the principal component analysis (PCA) and k-means cluster are employed to develop the driving cycle with case of Shenyang, China. First of all, many road conditions are collected, which are made up of a series of data including driving time and the instantaneous velocity. Then, PCA is applied to extract the main components of overall road condition and the K-means cluster is used to select representative kinematic fragments for forming a new driving cycle. Finally, the proposed driving cycle is simulated and verified. The result shows that the proposed driving cycle matches to overall road condition well. Under certain driving conditions, the driving cycles can achieve an alleviation of errors with actual road condition.

Keywords: driving cycle ;principal component analysis (pca); k-means clustering method

1. Introduction

New energy automobile is an efficient measure to achieve energy saving and emission reduction. Energy management strategy, which has important influence on the vehicle performance, is one of the most significant parts in the research of new energy automobile [1-2]. Both the development and evaluation of energy management are usually dependent on some certain driving cycles. However, it remains a problem to be solved to construct a driving cycle.

Problem of driving cycle development has attracted a few attentions, for example, a Relative-proportion-based Route Adjustment Process overcomes many limitations of the driving cycle's study in Ref. [3]. Ref. [4] used a generalized car-following model to study the driving cycle and then using the full

* Corresponding author. Tel.: +86-024-83681095; fax: +86-024-83681095.
E-mail address: yangyang@mail.neu.edu.cn.

doi:10.1016/j.egypro.2017.03.620

velocity difference model to verify the analytical results. Ref. [5] showed that vehicle speed affected by several uncertain factors may have a large impact on driving cycle's research.

The above research methods have their own merits and demerits. Considering the PCA and k-means cluster can be easily implemented due to their rigorous research mentality, these two methods are investigated in this study. The remainder of this paper is arranged as follows: experimental study (consists of data collection, research theory and methods) is given in Section 2; the driving cycle and results validation are described in Section 3 while the conclusion is summarized in the end.

2. Experimental Study

2.1 Data Collection

The first step is to collect data that can represent the actual road conditions. Shenyang's second ring road is chosen to be the experimental object that results in its high utilization. Performance box (P-box) is a GPS based performance meter used to measure speed, braking distance and many more. Then P-box displayed in figure 2 is performed as the speed acquisition tool to obtain original data, which records instantaneous speed at the frequency of 0.1s/time. Kinematic fragment is defined as the duration of vehicles from the beginning of one idle stop to next beginning of idle stop. It is the basic unit to build driving cycle. As a result, the initial 40 minutes' data can be divided into 55 kinematic fragments shown on figure 1(expert).

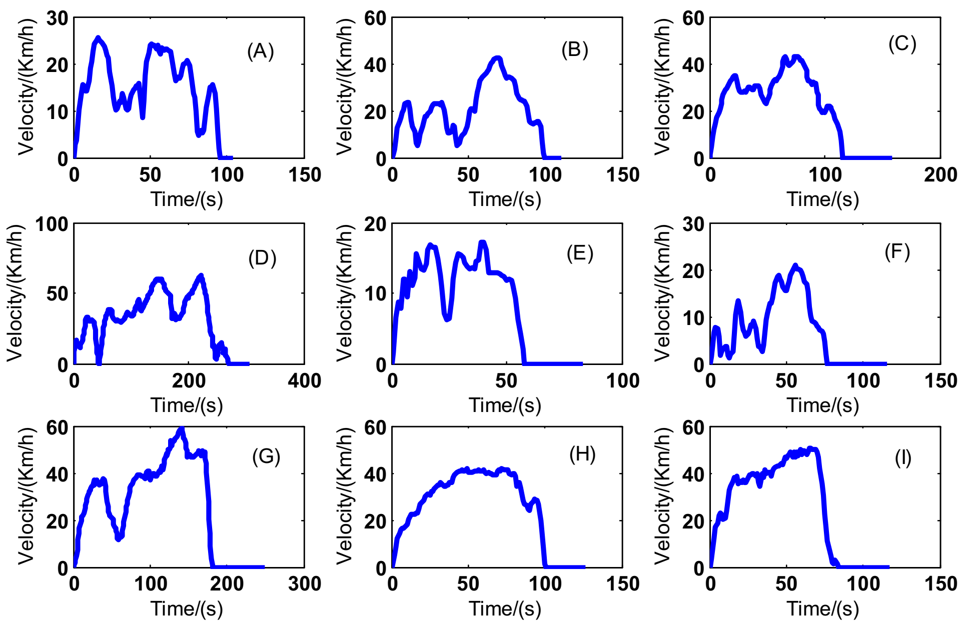

Figure 1 Kinematic fragments: (A)1st kinematic fragment; (B)2nd kinematic fragment; (C)3rd kinematic fragment; (D)4th kinematic fragment ;(E)5th kinematic fragment ;(F)6th kinematic fragment ;(G) 7th kinematic fragment; (H)8th kinematic fragment ;(I)9th kinematic fragment.

2.2 Theory of synthetic driving cycle

From the traffic conditions, the theory of mathematical statistics is applied to count driving parameters of the moving curve, then PCA and K-means cluster is used to study and classify kinematic fragments. Finally the working condition of driving cycle would be constructed. The selections of assessment criteria are shortlisted from the following driving parameters in Table 1.

There are four driving modes depicted in Table 2. Acceleration mode (A) is a continuous operation process that the vehicle's acceleration is greater than 0.05 m/s^2. Deceleration mode (B) refers to the vehicle whose deceleration is less than 0.05 m/s^2 .Idle mode (C) is acknowledged that the absolute value of vehicle's acceleration is less than 0.05 m/s^2 and the average speed is less than 3Km/h. The others belong to the cruise mode (D).

Table 1 Driving parameters

Driving parameters	Implication
①	average speed
②	average speed that not includes idle speed
③	average acceleration of acceleration mode
④	average deceleration of deceleration mode
⑤	time percentage of idle mode
⑥	time percentage of acceleration mode
⑦	time percentage of deceleration mode
⑧	time percentage of cruise mode

Table 2 Driving mode

Driving mode	Name
A	Acceleration mode
B	Deceleration mode
C	Idle mode
D	Cruise mode

Table 3 The classification of driving parameters

Class 1	① ②
Class 2	③ ④
Class 3	⑤ ⑥ ⑦ ⑧

2.3 The application of PCA

2.3.1 General introduction

PCA is a method that reduces the dimensionality of a dataset by finding a new set of variables which are smaller than the original set of variables [6]. This efficient reduction of the number of variables is achieved by obtaining orthogonal linear combinations of the original variables. Exactly they are principal components (PCs). The main steps of the PCA are portrayed in Figure 3.

Figure 2 P-box

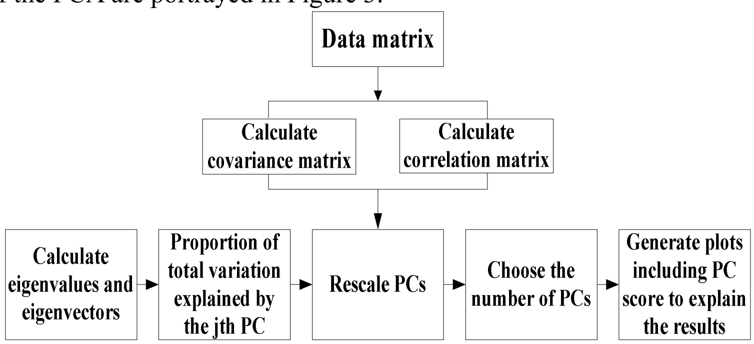

Figure 3 The main steps of PCA

2.3.2 The study process

SPSS is a widely used program for statistical analysis in social science. In addition to statistical analysis, data management (case selection, file reshaping, creating derived data) and data documentation (a metadata dictionary was stored in the data file) are features of the base software. It is used for PCA and K-means cluster on account of its abundant functions mentioned above.

Each component's cumulative value percentage calculated by SPSS is described in Table 4, the selected components are considered enough to explain the initial data when they are up to 80%. The first three components are taken as the PCs since their cumulative value percentage has reached 86.398%. A PCs' score matrix is depicted in Table 5 that would be used for K-means cluster in order to classify the total data sample. The main driving parameters of each component can be worked out and the consequences are shown in Table 6.

Table 4 The cumulative value percentage

Components	Cumulative value%
1	49.555
2	70.771
3	86.398
4	95.905
5	99.313
6	99.939
7	100
8	100

Table 5 The PCs' score matrix

Driving parameters' score	Components 1	2	3
①	0.938	-0.189	0.184
②	0.850	-0.221	0.396
③	-0.847	-0.032	-0.330
④	0.579	0.752	0.151
⑤	-0.738	0.132	0.661
⑥	0.537	-0.748	-0.358
⑦	0.452	0.654	-0.587
⑧	0.521	0.203	0.134

Table 6 The main driving parameters of each component

Component	Driving parameters
1	① ② ④ ⑥ ⑦ ⑧
2	④ ⑤ ⑦ ⑧
3	① ② ④ ⑤ ⑧

2.4 K - means cluster

Cluster analysis is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar in some sense or another to each other than to those in other groups (clusters). It is not only a main task of exploratory data mining but a common technique for statistical data analysis.K-means cluster is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

On the basis of the analysis above, K-means cluster can be designed for the extraction of excellent kinematic fragments. Through the comparison between the comprehensive scores (enumerated from SPSS) of kinematic fragments, the best ones are selected to represent the overall driving cycle portrayed in figure 4.

3. Results analysis

3.1 Composition of driving cycle

By the combination of the kinematic fragments extracted by K-means cluster according to the time sequence, the driving cycle of overall road condition could be concluded in figure 5.

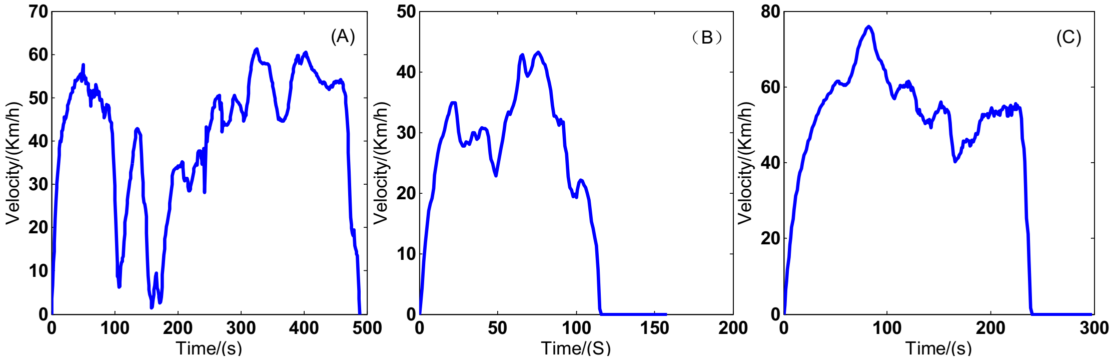

Figure 4 The selected kinematic fragments: (A) 14th kinematic fragment ;(B) 3rd kinematic fragment ;(C)11th kinematic fragment.

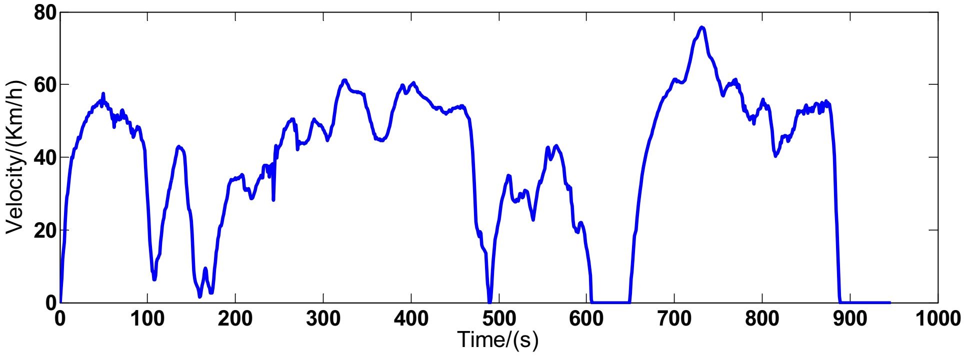

Figure 5 The development result of the Shenyang driving cycle

3.2 Results validation

The root mean square (RMS) is a statistical measure defined as the square root of the mean of the squares of a sample in statistics. RMS can be calculated for an integral of the squares of the instantaneous values during a cycle. It can be a characteristic of continuously varying driving parameters.

To ensure the accuracy of results, the 8 driving parameters are put into three classes shown in Table 3. Then 8 modes of compound driving cycles are taken to verify the veracities. The driving cycle composed by the two methods is served as the mode (1) and other seven driving cycles composed randomly are mode (2) to mode (8).In Table 7, RMS of each class belonging to its own mode are computed to compare the integrating degree between the driving cycles and the overall road condition.

From the values below, the result is considered to be reasonable since the RMSs of mode (1) are both small. Compared with mode (2), (3) and (7), it's distinct that the composed driving cycle achieves to alleviate the errors with actual road condition. However, it is not the greatest to fit the overall road condition as the RMSs of mode (4), (5) and (8) are less than mode (1).We will study for further research in the future.

Table 7 The RMSs of each mode

Mode / RMS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Class 1	1.577	1.583	2.942	1.271	0.532	1.773	1.172	0.577
Class 2	0.049	0.038	0.066	0.027	0.043	0.036	0.037	0.026
Class 3	0.026	0.047	0.082	0.028	0.018	0.009	0.046	0.013

Conclusions

In this paper, PCA and K-means cluster are used for the construction of driving cycle of Shenyang, China. The two methods investigated above are proved to be efficient for our study. By calculating the RMSs of the three classes of driving parameters attaching to eight composed driving cycle modes, a conclusion can be drawn safely that the driving cycle gotten from the two methods can represent the actual road conditions availably.

References

[1] N J Schoutena, M A Salmanb, N A Kheir. Energy management strategies for parallel hybrid vehicles using fuzzy logic. *Control engineering practice.*2003, 11, 171–177.
[2] Z Chen, R Xiong, J Cao. Particle swarm optimization-based optimal power management of plug-in hybrid electric vehicles considering uncertain driving conditions. *Energy*. 2016, 96, 197-208.
[3] W Zhu, S Ma, J Tian , G Li .Nonlinear relative-proportion-based route adjustment process for day-to-day traffic dynamics: modelling ,equilibrium and stability analysis. *Communications in nonlinear science and numerical simulation.* 2016, 129-137.
[4] T Tang , L Chen ,H Huang, Z Song.Analysis of the equilibrium trip cost without late arrival and the corresponding traffic properties using a car-following model. *Physical statistical mechanics and its applications.*2016, 348-360.
[5] AK Jagerbrand, J Sjobergh.Effects of weather conditions, light conditions, and road lighting on vehicle speed.*Springerplus.*2016,5.
[6] S Xu, X Jiang, J Huang, S Yang, X Wang. Bayesian wavelet PCA methodology for turbo machinery damage diagnosis under uncertainty. *Mechanical systems and signal processing.*2016, 1-18.

Biography

Dr.Ying Yang is a Professor of Northeastern University, where she received her M.Sc. degree in vehicle Engineering and Ph.D. degree in mechanical Engineering in 1989 and 2000, respectively. She has published more than 30 papers in vehicular technique indexed by SCI/EI. Her research interests include hybrid electric vehicle, visual navigation and intelligent active security system of vehicle.