Construction and research on the navigation condition of oil supply vessels

Shiding Hong [1], *Chaokui* Qin [1*], *Haifeng* Dai [2], *Xin* Lai [3]

[1] School of Mechanical and Energy Engineering, Tongji University, Shanghai 201804, China.

[2] School of Automotive Studies, Tongji University, Shanghai 201804, China.

[3] College of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Abstract. Constructing navigation condition of oil supply vessels is of vital importance to both the research on the main engine and the safety of navigation. This paper carries out in-depth research and analysis based on the half-year shipping data derived from Hai Gong You 31, the oil supply vessel of Shanghai Chimbusco Marine Bunker Co., Ltd. After classifying 3946 pieces of shipping data, and there are 119 valid sequences. As for 13 characteristic parameters ,which include time, distance, average acceleration ,etc., they are eigenvalues of kinematics sequences ,and there are principal component analysis and contribution rate analysis relevant to them. Instead of original eigenvalues, three principal components whose contribution ratios are the highest and the accumulative one is around 85% are selected to be analyzed then. By means of between-groups linkage, this paper calculates the distance among clusters ,and the nearest two clusters are regarded as a new cluster. Afterward, calculations and combinations are made over and over till there is only one cluster. Eventually, the navigation condition of oil supply vessels is successfully formed, that is the one with a total length of 9240s.

Key words: oil supply vessels, the construction of working condition, sequence analysis, eigenvalue analysis, principal component analysis, cluster analysis.

1 INTRODUCTION

China joined the World Trade Organization in 2001. Since then, this country has gradually become a global manufacturing base. Under the background of economic globalization, China's economy grows rapidly, posing considerable challenges to its shipping industry[1-2]. In a sense, a country's development is dependent on this industry. With domestic and international trade expanding, China's shipping industry achieves significant growth. Meanwhile, the demand for marine fuel oil rises sharply. To be more specific, statistics reveal that the internal trade demand and the external trade demand are around 9 million tons and 16.87 million tons ,respectively. In most cases, marine fuel oil is supplied by oil supply vessels. For the purpose of conducting further research into these vessels' main engines and safety, it is necessary to construct working condition of oil supply vessels.

Shanghai Chimbusco Marine Bunker Co., Ltd. is the largest supplier of marine fuel oil in Shanghai ,and its annual provision is approximately 1.3 million tons. Playing a key role in Shanghai International Shipping Center, this enterprise has almost 20 oil supply vessels and 5 of them belong to the enterprise itself. Owned by Shanghai Chimbusco Marine Bunker Co., Ltd., Hai Gong You 31 is 52.73 meters long, 9.20 meters wide, and can carry 750 tons of goods. Moreover, it is a relatively typical kind of oil supply vessel. As a consequence, this paper chooses the data obtained from Hai Gong You 31 to form working conditions, carries out in-depth research and analysis based on the half-year shipping data derived from it. And with the method of Factors Analysis, Principal components Analysis, Contribution rates Analysis and Cluster Analysis, finally, the navigation condition of oil supply vessels is successfully formed, that is the one with a total length of 9240s.

2 THE SCHEME FOR GATHERING DATA AND CONSTRUCTING WORKING

* Corresponding author: chkqin@tongji.edu.cn

CONDITION

Mainly undertaking the task of supplying bonded oil in Shanghai Chimbusco Marine Bunker Co., Ltd., Hai Gong You 31 works within districts shown in Fig.1, including Wai Gaoqiao, Baoshan and Wusong. Hai Gong You 31's annual supply is about 100,000 tons, the designed speed is 9 knots ,and the main engine power is 180 KW (dual main engine). In order to form its navigation condition, this paper makes use of Automatic Identification System (AIS) to collect its voyage data from September 2020 to February 2021. Acquiring and analyzing 3946 pieces of data, there are 119 valid sequences in the end. As is shown in Fig.2, methods such as eigenvalue analysis, principal component analysis and cluster analysis are employed to analyze these sequences and construct fitted navigation conditions of Hai Gong You 31 then.

Fig.1 Hai Gong You 31's course

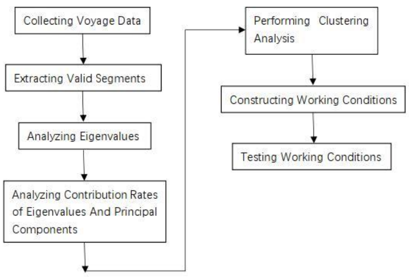

Fig.2 The procedure for constructing working condition

3 DATA PROCESSING METHODS

3.1 The division of kinematics sequences

In the course of sailing, ships can be in different conditions, including start, acceleration, deceleration ,and idle speed[3]. On the basis of the demand for data processing, this paper divides data into kinematics sequences. More specifically, a kinematics sequence refers to a continuous process that the ship experiences setting sail, acceleration, maintaining a steady speed, decelerating ,and stopping successively. Such classification is beneficial to the analysis of ship's condition. As is reflected in Table 1, after classifying 3946 pieces of data herein, 119 valid segments are gained in total.

Table 1. The summary of kinematics sequences.

Kinematics sequences（KM/h）						
1	2	3	……	117	118	119
0.00	0.00	0.00	……	0.00	0.00	0.00
4.00	5.33	7.78	……	3.46	4.58	4.17
9.33	11.33	15.00	……	6.61	16.53	10.67
……	……	……		……	……	……
3.33	4.17	5.00	……	5.82	2.92	1.83
1.83	1.83	2.96	……	2.52	0.83	1.33
0.00	0.00	0.00	……	0.00	0.00	0.00

3.2 Factor Analysis

After dividing data into kinematics sequences, speed and time fail to give a comprehensive description of such sequences. Consequently, it is essential to take advantage of other characteristic parameters[4]. For the purpose of describing the feature of kinematics sequences adequately, this paper concerns 13 characteristic parameters as eigenvalues of them, including running time, the time in steady running speed, acceleration time, etc. Information on each characteristic parameter is indicated in Table 2.

Table 2. Characteristic parameters used to describe kinematics sequences

Serial numbers	Eigen values	Definitions	Units
1	T	Running time	s
2	Tc	The time in steady Running speed	s
3	Ta	Acceleration time	s
…	…	…	…
11	aa	Average acceleration	m/s2
12	ad	Average deceleration	m/s2
13	asd	Standard deviation of acceleration	m/s2

Using Java language to design a program in My Eclipse, it is possible to make computations of 119 segments and then obtain eigenvalues of all of them. Taking a segment as an example, relevant results can be seen in Table 3.

Table 3. Eigenvalues of a segment

Running time	2520	s
The time in steady running speed	240	s
Acceleration time	840	s
…	…	…
Average acceleration	0.00206	m/s2
Average Deceleration	-0.00204	m/s2
Standard deviation of acceleration	0.005541	m/s2

3.3 Principal component analysis

The analysis reveals that these 13 characteristic parameters are not mutually independent. Instead, there is overlap among the information demonstrated by them. If all these parameters are used to form sailing conditions, the amount of data will be so large that it will be complicated to handle[5]. In order to make the calculation easier, this paper uses the method of principal component analysis to recombine previous indicators with certain correlations into new linear independent indicators through linear transformation, which could condense them into a few principal components with the least loss in information.

3.3.1 Calculation Process of Principal Component Analysis

3.3.1.1 Data Processing

Let n be the number of kinematics sequences and p be the number of characteristic parameters, the corresponding matrix is:

$$A = \begin{pmatrix} a11 & a12 \cdots & a1p \\ \dots & \dots & \dots \\ an1 & an2 \cdots & anp \end{pmatrix} \quad (1)$$

Now matrix A is called the original matrix, and then this original matrix is standardized to eliminate the differences between indicators[6].

Let:

$$X_{ij} = \frac{a_{ij}-\overline{a_j}}{\sqrt{var(a_j)}} \quad (2)$$

of which:

$$\overline{a_j} = \frac{1}{n}\sum_{i=1}^{n} a_i j \quad (3)$$

$$\mathrm{var}(a_j) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i j - \overline{a_j})^2 \quad (4)$$

The standardized matrix is expressed as:

$$X = \begin{bmatrix} x11 & x12 \cdots & x1p \\ \vdots & \dots & \vdots \\ xn1 & xn2 \cdots & xnp \end{bmatrix} = (X1 \quad X2 \quad \dots \ Xp) \quad (5)$$

3.3.1.2 Calculation of correlation coefficient matrix

Suppose the correlation coefficient matrix is R ,, then:

$$R = \begin{pmatrix} r11 & r12 & \cdots & r1p \\ \vdots & & \ddots & \vdots \\ rp1 & rp2 & \cdots & rpp \end{pmatrix} \quad (6)$$

where r_{ij} is the correlation coefficient between X_i and ,X_j,

$$r_{ij} = \frac{\mathrm{cov}(X_i,X_j)}{\sqrt{\mathrm{Var}(Xi)var(X_j)}} \quad (7)$$

$$cov(X_i,X_j) = \frac{1}{n-1}\sum_{ij=1}^{n}(Xij - \overline{X}_i)(X_{ij} - \overline{X}_j) \quad (8)$$

3.3.1.3 Calculation of the eigenvalue and eigenvector of R

$$R-\lambda E = 0 \quad (9)$$

In total, p eigenvalues of λ1, λ2…λp which satisfyλ1 > λ 2 > … > λp > 0 can be calculated, and the corresponding eigenvectors u1, u2 … up can be obtained. All eigenvectors can be constructed into an eigenmatrix U.

$$U = (u_1, u_2, \dots, u_p) \quad (10)$$

3.3.1.4 Calculation of the contribution rate of principal components

In terms of indicators recombined from the original indicators by a linear transformation in the process of principal component analysis, although these new indices are linearly independent of each other, their importance is unequal[7]. The eigenvalues obtained in the previous step, λ1, λ2…λp, represent the importance of the first to p_{th} principal components ,respectively. This paper utilizes the contribution rate of each principal component to show its importance, and the former can be calculated as follows:

$$f_k = \frac{\lambda_i}{\sum_{i=1}^{P}\lambda_i} \quad (11)$$

It is universally acknowledged that when the cumulative contribution rate of principal components exceeds 85%, this part of principal components contains the mostinformation concerning the original indices. Supposing the cumulative contribution rate is a_k, it can be believed that the first k principal components can replace the original index when $a_k \geq 85\%$[8].

$$a_k = \sum_{i=1}^{k} f_k \qquad (12)$$

3.3.1.5 Calculation of the value of principal components

According to 3.3.1.4, the value of the first k principal components requires computation. The method is as follows:

$$Z = \left(X_{u1}, X_{u2}, \ldots, X_{uk}\right) = (\varepsilon 1, \ \varepsilon 2, \cdots, \ \varepsilon k) \qquad (13)$$

Herein, $\varepsilon 1, \ \varepsilon 2, \cdots, \ \varepsilon k$ are the first to K_{th} constructed principal components respectively, and they are linearly independent with each other. f_1, f_2, $\cdots$, f_k are the corresponding contribution ratios, which decrease in turn.

By means of SPSS, this paper carries out the principal component analysis of 119 kinematics sequences and calculates all their eigenvalues. Ultimately, principal components' contribution rates and cumulative ones can be acquired. See Table 4 for related figures[9].

Table 4. Principal components' contribution rates and cumulative contribution rates

Compo-nents	Original eigenvalues		
	Overall	Contribution rates (%)	Cumulative contribution rates (%)
m1	6.125	47.119	47.119
m2	3.338	25.675	72.793
m3	1.484	10.798	83.592
…	…	…	…
m11	0.010	0.077	99.985
m12	0.002	0.015	100.000
m13	3.9E-17	3.0E-16	100.000

As is shown in Table 4, the cumulative contribution rate of the first three principal components is 83.592%, which is close to 85% and can reflect the original working condition adequately. Therefore, this paper chooses these three components to continue handling the data derived from Hai Gong You 31. Table 5 suggests the principle-component score matrix of the working condition segment.

Table 5. Principle-component score for Hai Gong You 31's working condition segment

Working condition segment	Principle component m1	Principle component m2	Principle component m3
1	1.31194	-0.04274	-0.20554
2	0.58536	0.19714	-0.51715
3	0.05514	0.7543	0.25503
……	……	……	……
117	-1.02179	0.14224	0.27458
118	-1.2167	-0.29464	1.70879
119	-0.94873	-0.17625	0.07079

3.4 Hierarchical Cluster Analysis

When performing hierarchical cluster analysis, this paper concerns each fragment in the whole as a cluster, the nearest two clusters as a new one. Afterward, distance calculation and clusters combination are made over and over till there is only one cluster. It is by means of between-groups linkage that the distance among clusters is calculated herein[10].

The program written in My Eclipse is able to compute the overall eigenvalues of not only each category of kinematics sequences but also the whole. The results are shown in Table 6.

Table 6. Eigenvalues of the whole and each category

Eigenvalues	The first category of segments	The second category of segments	The whole
Running time (s)	464880	8640	473520
The time in steady Running speed (s)	91560	2880	94440
Acceleration time (s)	188400	3120	191520
…	…	…	…
Constant speed ratio (%)	0.19695	0.33333	0.19944
Acceleration ratio (%)	0.40526	0.36111	0.40446
Deceleration ratio (%)	0.39778	0.30555	0.39609

According to Table 6, it can be realized that the running time of the first category segments is 464880s, while the second one is 8640s. Obviously, the latter only accounts for 1.82% of the overall running time. Taking the result of hierarchical cluster analysis into consideration, it is 118 segments in the first category rather than the second category of segments that the fitted working condition is extracted from.

4 CONSTRUCTION OF NAVIGATION CONDITION

4.1 The length of the working condition to be constructed.

Based on the method for determining the length of vehicle working conditions, this paper takes data as well as reality into consideration and then sets 9000 seconds as the length of ship navigation conditions. In other words, the working condition fitted with a certain number of kinematics sequences should fulfill this condition.

4.2 Calculation of correlation coefficients

Correlation coefficients indicate the degree of correlation between two vectors. The absolute value of the correlation coefficient gets closer to 1, and these two vectors are much more correlated with each other. The calculation formula of correlation coefficients is as follows:

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{var}(Y)}} \qquad (14)$$

Where Var(x) and Var(y) are variances of vector X and vector Y respectively, ρ is the correlation coefficient of vector X and vector Y.

4.3 Judgment about whether the absolute value of correlation coefficient meets the standard

When constructing ship navigation conditions, it is generally believed that the segments with a relatively high degree of correlation are those with the absolute value of correlation coefficient above 0.8. As a consequence, this paper removes the segments whose absolute value of correlation coefficient is less than 0.8. Then the remaining segments are randomly combined according to the required quantity calculated in the second step, and the characteristic parameter values of these randomly combined long segments are recalculated. In addition, we compute the correlation coefficient between these characteristic parameter values and the overall comprehensive characteristic value of such segments to find the combination with the largest correlation coefficient. Afterward, this combination is used to construct the final working condition[11].

4.4 Construction of ship navigation condition

From the first kind of characteristic value in Table 6, this paper extracts the average velocity, average acceleration, average deceleration, acceleration ratio, deceleration ratio and constant speed ratio, followed by calculating the correlation coefficient between the eigenvalue of every segment and the overall comprehensive eigenvalue with CORREL function in Excel. Eventually, it can be found that the absolute value of each coefficient is greater than 0.9.

In the first kinematics sequence herein, there are three segments with the highest correlation coefficient, whose length reaches 3480s, 2880s and 2880s ,respectively. They are selected to form navigation conditions with a total length of 9240s. The correlation coefficients between eigenvalues of these three segments and the overall eigenvalue are all above 0.99991, revealing that they have a close correlation and can represent the actual sailing condition. Fig.3 exactly shows the constructed

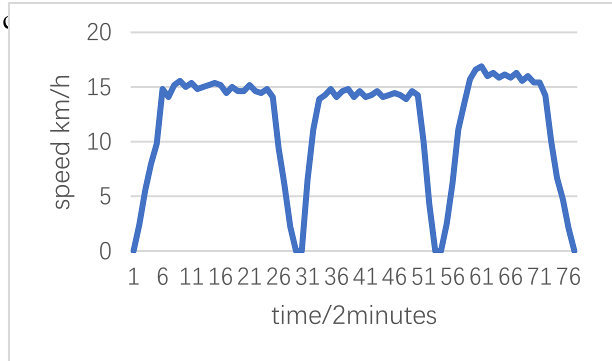

Fig.3 The fitted curve indictating Hai Gong You 31's navigation condition

5 CONCLUSION

Oil supply vessels, offering logistical support, play a prominent role in the development of the shipping industry. Nevertheless, there have not been references mentioning their real standard navigation condition. Consequently, it is crucial to undertake research into their voyage data as well as navigation condition. To be more specific, such research can assist in learning more about oil supply vessels, including the performance of their main engines, shipborne lithium battery and so on. This paper collects the data about Hai Gong You 31's voyage, carries out eigenvalue analysis, principal component analysis, contribution rate analysis as well as cluster analysis, and constructs fitted working conditions which can represent actual shipping eventually.

REFERENCES

[1] Andre M.Driving Cycles Development: Characterization of the Methods [J].Sae Technical Papers,1996.
[2] Nesamani K S, Subramanian K P. Development of a driving cycle for intra-city buses in Chennai, India [J].Atmospheric Environment,2011,45(31):5469-5476.
[3] Knez M, Muneer T, Jereb B, et al. The estimation of a driving cycle for Celje and a comparison to other European cities[J]. Sustainable Cities & Society,2014,11(2-3):56-60.
[4] Zhang Hong，Yao Yan-gang，Zhang Lu． Study on Construction of Light- duty Vehicle Driving Cycle in Hohhot City [J]. Journal of Highway and Transportation Research and Development，2019，36(11):104-110．
[5] Liu Ying-ji，Xia Hong-wen，Yao Yu. A Method of Vehicle Driving Cycle Development Based on Combined Principal Component Analysis and Fuzzy C Means Clustering [J]. Journal of Highway and Transportation Research and Development, 2018, 35(3): 79-85.
[6] Shi Qin，Zhen Yu-bo，Jiang Ping. A Research on Driving Cycle of City Roads Based on Microtrips [J]. Automotive Engineering, 2011, 33(3): 256-261.
[7] Li Ning. Construction and Study of Driving Cycle of Urban Road Vehicle [D]. Baoding: Hebei Agricultural University, 2013．
[8] Liu Zi-tan，Zhu Ping，Liu Xu-peng. Research on Improved K-means and Driving Cycle Construction [J]. Automobile Technology, 2019(11):57-62．
[9] ADAK P， SAHU Ｒ， ELUMALAI S P. Development of emission factors for motorcycles and shared auto-rickshaws using real-world driving cycle for a typical Indian city [J]. Science of the total environment, 2016, 544: 299-308.
[10] BRADY J, O' MAHON M. Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas [J]. Applied energy, 2016, 177: 165-178．
[11] LIN J, NIEMEIER D A. An exploratory analysis com-paring a stochastic driving cycle to California's regulatory cycle [J]. Atmospheric environment, 2002, 36(38): 5759-57