Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

# Modeling and Prediction on Octane Number Loss in Marine Fuel Production

Weiming Yang, Fen Liu*

College of Information Engineering, Shanghai Maritime University, China

* liufeng@shmtu.edu.cn

---

*Abstract: Environmental shipping is related to the development prospects of the shipping economy, as well as the harmonious coexistence of man and nature. Petroleum is still the main fuel for large ships, the tail gas from combustion pollutes the environment. With increasingly stringent environmental protection policies at home and abroad, it is indispensable to use data analysis to create smart shipping. In this paper, the principal component analysis method is used to reduce the dimensionality of the data obtained in the production of marine oil to obtain the main variables. Then the neural network model is used to establish the octane number loss model, and the relationship between the octane number loss and variables is obtained. While ensuring the reduction of sulfur and olefin content in gasoline, the loss of its octane number should be minimized. To achieve the purpose of improving combustion efficiency and protecting the ecological environment.*
*Keywords: Environmental Shipping, Data Analysis, Neural Network, Octane Number Loss.*

---

## 1. INTRODUCTION

With the acceleration of economic globalization and information globalization, the role of ship transportation has become increasingly significant. The shipping industry not only attaches equal importance to technology and management, but also relates to the harmonious development of man and the natural environment, which puts forward higher requirements for the environmental protection and efficiency of marine fuel oil[1]. The existing technology also loses the octane number of the fuel when desulfurization and olefin reduction are performed in the production of ships. Therefore, in order to ensure the efficiency of fuel production, companies will lower the standards for fuel desulfurization and olefin reduction, which will increase fuel combustion emissions. Therefore, to reduce the loss of octane number in the process of desulfurization and olefin reduction, higher desulfurization and olefin reduction effects can be achieved under the same requirements[2].

Based on this situation, this paper investigated the Sinopec Gaoqiao Petrochemical real-time database (Honeywell PHD) and the LIMS experimental database, and obtained 325 data samples collected by the catalytic cracking gasoline refining unit to establish the octane loss in the production of marine oil The prediction model of each sample is given, and the optimal operating conditions for each sample are given.

## 2. MAIN VARIABLES OF THE MODEL

In this paper, 325 original data are processed with invalid eigenvalues to eliminate abnormal data. What is obtained is the mean value of 354 manipulated variable values at each time point[3]. By standardizing the original data, the characteristic value and characteristic vector of the correlation coefficient of the data sample are obtained. Use the obtained data to establish a principal component model, calculate the cumulative variance contribution rate, and determine the number of principal components.

By analyzing the data of the principal component calculation results, the influence of the top ten contributing factors on the octane loss is obtained[4]. It can be seen from Figure 2 that the first 30 principal components can represent 88.01% of the original data, so only the first 30 principal components need to be extracted in the meaning of 88.01% to achieve the purpose of principal component extraction. And the final cumulative contribution of the first 10 principal components can be obtained.

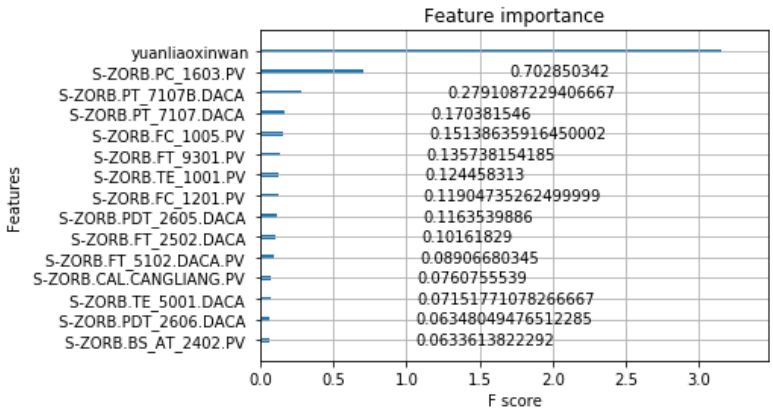


Fig. 1 Contribution of main components

Through the analysis of the 367 data indicators in the figure, combined with the analysis of other factor indicators, it can be concluded that the focus of different operating variables is different, and the principal components of the top 30 can be obtained through principal component analysis as follows:

Table 1 Principal component analysis

| Variable index | Variable name | Variable index | Variable name |
|---|---|---|---|
| 1 | D-109 loose wind flow | 16 | D-109 bottom |
| 2 | Hot nitrogen filter ME-113 differential pressure | 17 | K-103B exhaust temperature |
| 3 | Pressure difference between inlet and outlet of the tube side | 18 | E-205 tube side inlet tube temperature |
| 4 | D-123 Condensate inlet flow rate | 19 | Fuel gas inlet device temperature |
| 5 | ME-101 backflush gas header pressure | 20 | D-203 fuel gas inlet pipe temperature |

| | | | |
|---|---|---|---|
| 6 | EH-101 heating element temperature | 21 | D-122 top outlet pipe temperature |
| 7 | Backflush hydrogen pressure | 22 | Backflush gas concentrator |
| 8 | F-101 circulating hydrogen outlet pipe temperature | 23 | D101 raw material buffer tank pressure |
| 9 | 1.1 Steps PIC2401B.OP | 24 | S_ZORB AT-0008 |
| 10 | D-114 liquid level | 25 | Emergency hydrogen main |
| 11 | D105 temperature | 26 | Back blow gas pressure |
| 12 | K-102B intake temperature | 27 | D-110 top pressure |
| 13 | K-103B intake air temperature | 28 | ME-104 inlet and outlet |
| 14 | E-106 pipe pass outlet pipe temperature | 29 | refined gasoline outlet device line pressure |
| 15 | D-124 top outlet pipe temperature | 30 | degassing flow rate of gasoline products |

## 3. MODEL OF OCTANE NUMBER LOSS

In this paper, neural network model is used to predict octane number loss[5]. The topology of 30-59-1 is established. There are 30 neurons in the input layer, representing 30 variables affecting octane number loss; 59 neurons in the hidden layer and 1 neuron in the output layer, representing octane number loss. The neural network topology is shown in Figure 2

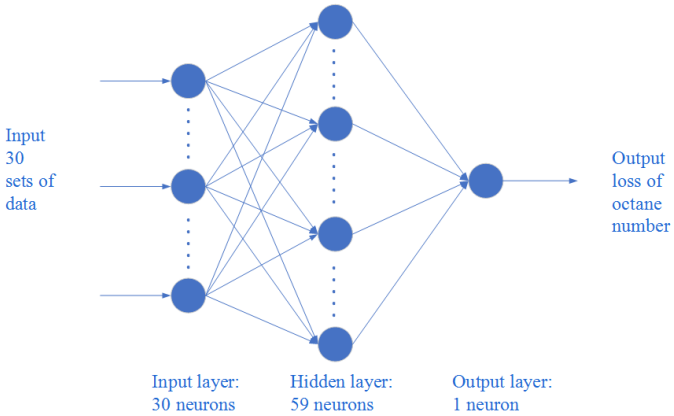


Fig. 2 Construction of BP neural network topology

The original data is divided into training samples and test samples, which are substituted into the trained BP neural network to predict the octane number loss[6]. Then the results are evaluated. It can be seen from Figure 3 that r = 0.52206. It can be considered that the neural network model is close to the prediction of octane number[7].

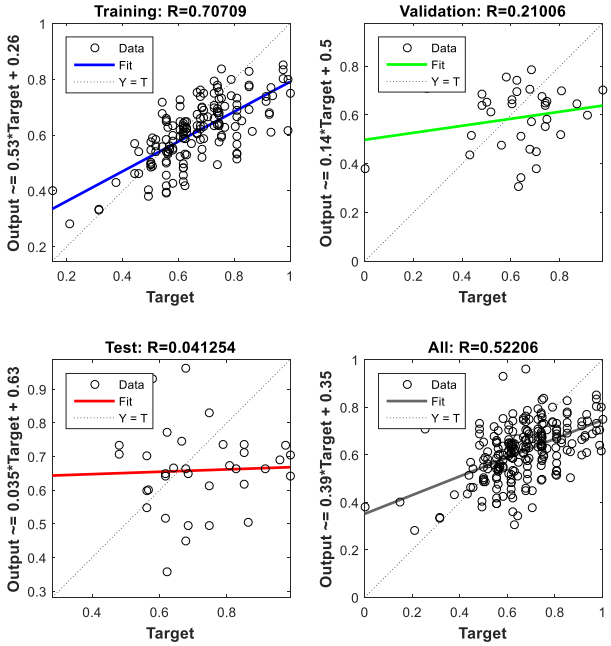


Fig. 3 Fitting effect analysis

The weights and thresholds in the neural network model are very important, and their accuracy will directly affect the prediction output of the test sample, so when solving the weights and thresholds, the gradient descent method is generally used[8]. Through continuous iteration, the prediction error in the neural network is reduced, and the optimal solution is finally obtained, making the prediction error relatively small. It can be seen in Figure 4 that the gradient is 0.0010527 when iterating to the 9th generation, and then the values of MU and val fail are integrated, and the prediction error is also very close to 0.

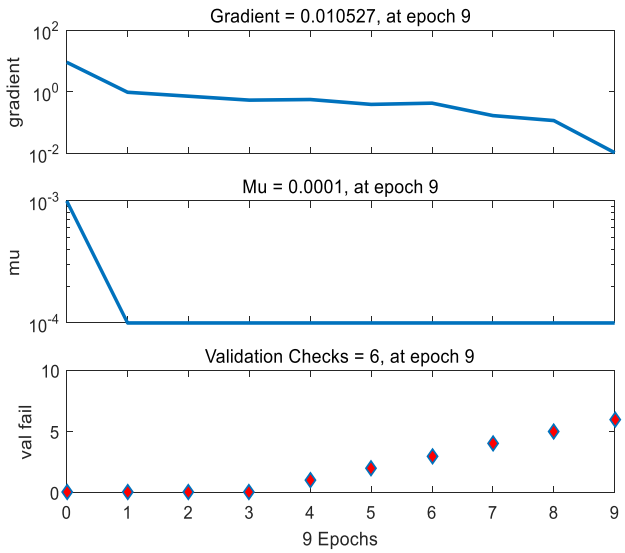


Fig. 4 Change of gradient with number of iterations

## 4. OPTIMIZATION OF MAIN VARIABLES

In order to visualize the octane loss model in ship fuel production, this article uses the model to obtain 325 data samples under the premise that the product sulfur content is not greater than 5μg/g, and the octane loss decreases by more than 30% Requirements to measure the pros and cons of the model.

This paper uses a data fitting model based on the least squares method to optimize the main variables[9].

For the convenience of calculation, it is assumed that $p$ dependent variable $y_1, y_2, y_3, \cdot; y_p$ and $m$ independent variable $x_1, x_2, x_3, \cdots\cdots, x_m$ are all standardized variables, and they are recorded as[10]:

$$F_0 = \begin{pmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{np} \end{pmatrix}, \quad E_0 = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \tag{1}$$

Extract the first pair of components of the two variable groups separately, and maximize the correlation.

Establish regression equations of $y_1, y_2, y_3, \cdots\cdots, y_p$ and $x_1, x_2, x_3, \cdots\cdots, x_m$ versus $t_1$

Suppose the regression model is[11]

$$\begin{cases} E_0 = \hat{t}_1 \alpha_1^T + E_1 \\ F_0 = \hat{u}_1 \beta_1^T + F_1 \end{cases} \tag{2}$$

Let the rank of $n \times m$ data matrix $E_0$ be $r \le (n-1, m)_{\min}$, then there are $r$ components $t_1, t_2, t_3, \cdots\cdots, t_r$ such that

$$\begin{cases} E_0 = \hat{t}_1 \alpha_1^T + \hat{t}_2 \alpha_2^T + \cdots + \hat{t}_r \alpha_r^T + E_r \\ F_0 = \hat{t}_1 \beta_1^T + \hat{t}_2 \beta_2^T + \cdots + \hat{t}_r \beta_r^T + F_r \end{cases} \tag{3}$$

The partial least squares regression equation for the $p$ dependent variables is:

$$y_j = a_{j1} x_1 + a_{j2} x_2 + \cdots + a_{jm} x_m \left( j = 1, 2, 3, \cdots, m \right) \tag{4}$$

The constraints established here are as follows:

$$\begin{cases} S \le 5 \\ \dfrac{A_0}{A_1} \le 70\% \end{cases} \tag{5}$$

Firstly, the functional relationship between the octane number loss and 30 main operating variables is established, and then a fitting is performed through the data of the octane number loss and the 30 main variables. In the process of fitting, the data is sieved twice as necessary, and then the data is fitted after sieving[12]to obtain a 30-yuan equation.

The calculation equation can get the coefficients before 30 operating variables. The values of the coefficients are shown in Table 2. These coefficients represent the weights in the principal component analysis. According to the parameters produced by the fitting, the quality of the fitting effect is evaluated[13]. Here, the RMSE result is 0.2004, which is very close to 0, and the fitting error is relatively small.


Fig. 5 Data fitting process of main variables

Table 2 Coefficient before manipulated variable

| Coefficient | Value | Coefficient | Value |
|---|---|---|---|
| C1 | -0.0008999 | C16 | -0.08503708 |
| C2 | 1.278303 | C17 | 0.001060527 |
| C3 | 0.0346863 | C18 | 0.001532335 |
| C4 | -0.1460983 | C19 | -0.4039221 |
| C5 | -0.0130704 | C20 | -0.00366823 |
| C6 | 0.000459 | C21 | -0.02619136 |
| C7 | 4.28E-05 | C22 | 0.02334884 |
| C8 | 0.0032523 | C23 | -0.00483198 |
| C9 | -0.5136655 | C24 | 1.07E-06 |
| C10 | 0.0145612 | C25 | 0.000595062 |
| C11 | -0.0022216 | C26 | 0.002015721 |
| C12 | -0.0003261 | C27 | 0.004491746 |
| C13 | 0.7763253 | C28 | -0.8792055 |
| C14 | 3.911284 | C29 | -6.14E-07 |
| C15 | -0.0357907 | C30 | -0.00686233 |

In the process of solving, these constraints can be added to 30 operation variables for screening, and then fit according to the screening results. Finally, the octane number loss can be reduced to 70% of the original value, and the fitting results are in line with the actual requirements.

## 5. CONCLUSION

While developing fast shipping, develop smart shipping and use data processing as a means to create environmentally friendly shipping. This paper uses data analysis as a means to establish a model of octane number loss in marine fuel oil. In this paper, the scientific mathematical model is used to improve the octane number of ship oil, ensure the effective reduction of harmful substances in ship oil, and make the development of shipping data and environmental protection intelligent.

## ACKNOWLEDGMENTS

This research was funded by the National Natural Science Foundation of China (U1701265) and the Natural Science Foundation of Shanghai (20ZR1423200).

## REFERENCES

[1] Xiaojin Wang, “Strive to build environmentally friendly shipping”[J], World Shipping, 2015:p51-56.
[2] Haijun Wei, Feng Yin, Jizhe Wang, “Use and management of marine fuel oil”[M], Dalian:Dalian Maritime University Press, 2017:24.
[3] Zongli Li, Kaimin Liu, “Feasibility Analysis on Development of High Performance Medium Speed Diesel Engine”[J], Shandong Internal Combustion Engine, 2019,95(1):1-3.
[4] Junfeng Qin, Lin Lai, “Analysis on Optimization Measures to Reduce Octane Number Loss of Gasoline Hydrogenation Unit”[J]. China Chemical Trade, 2018, 10 (1):179.
[5] Yi Zhang, Jianyong Wang, “Neural Network Algorithm for Big Data Analysis”[J], Engineering Science and Technology, 2017, 49(1):9-18.
[6] Shengli Jiang, “Research on Feature Selection and Feature Extraction of High Dimensional Data”[D], Xi'an, Xi'an University of Electronic Science and technology, 2011.
[7] Liangjun Zhang, Junde Chen, Mingjun Liu, Rong Chen, “Data Mining Practical Case Analysis”[M], Beijing, Mechanical Industry Press, 2016:195-200.
[8] Dan Zheng, “New Course Under The Low Sulphur Order”[M], Shanxi, Industry Press, 2019:28-39.
[9] Shengwei Yu, “MATLAB Mathematical Modeling Classic Case Practice”[M], Beijing, Tsinghua University Press, 2015:201-209.
[10] Ruwen Chen, Ren Huang, “Research and Application of Nonlinear Autoregressive Time Series Model”[J], System Engineering Theory And Practice, 2015 (09).
[11] Haifeng Fang, “Analysis of Ship Oil under the New Environmental Protection Policy”[J], Modern Chemical, 2015(68):32-34.
[12] Xiaohua Zhao, Ying Yao, Yiping Wu, Chen Chen, Jian Rong, “Study on Combined Prediction Model of Driving Energy Consumption Based on Principal Component Analysis and BP Neural Network”[J], Transportation System Engineering and Information, 2016 (03).
[13] Cong Li, “Research and Application of Logistics Information Big Data Analysis Method”[D], Wuhan University of Technology, 2014.
[14] Shanbao Yu, Hong Guo, “Strengthening Process Management to Reduce Octane Number Loss of Refined Gasoline in s Zorb Unit”[J], Petroleum Refining and Chemical Industry, 2018, 49 (10):15-19.