# Research on Processing Quality Control Model Based on Regression Analysis

Linji Li *, Tan Li, Yifeng Qiu, Junze Shao

School of Public Health, Southern Medical University, Guangdong, China, 510515

* Corresponding Author Email: 3217042035@i.smu.edu.cn

**Abstract.** This article mainly studies the mathematical relationship between system temperature and the quality of finished ore processing products, while the variables of raw ore parameters, processing voltage, and processing water pressure remain unchanged. By establishing daily unified production batch processing data and visually analyzing data characteristics through charts, and visualizing the interaction between different temperature systems, a multiple interaction regression analysis model based on system temperature and processing quality indicators is selected. And through goodness of fit p-value testing, interval prediction with confidence, and sensitivity analysis, the model is sensitivity optimized. An interactive effect diagram and a visual and controllable product quality control program are created for the raw ore quality, system temperature, and quality index qualification rate. The theoretical value with the highest ore processing efficiency is deduced, which can effectively improve resource utilization and yield rate in the ore processing process.

**Keywords:** Processing Quality Control, Multiple Regression, Interactive Regression Analysis, Goodness of Fit Test.

## 1. Introduction

In today's world, with a rapidly increasing population and an increasingly fierce conflict between population and resources, most developed countries are committed to maintaining efficient resource utilization and sustainable energy utilization. At the same time, China's industry is also undergoing transformation and upgrading, and the demand for mining resources is still rapidly increasing. The energy consumption generated by ore processing has long been at the forefront of global industries, Therefore, improving the rationality of the ore processing process and improving the yield of finished ore products has become an urgent task worldwide. The quality of ore processing will affect the utilization rate of mineral resources, energy utilization rate, and production efficiency of enterprises. Production technicians need to set different system temperature controls for raw materials with different parameters in advance to manufacture ores that meet sales conditions [1-5]. In the actual processing process, due to various factors, the real-time processing temperature may differ from the set temperature, and even under the same system temperature conditions, the quality of products produced may have significant differences. Therefore, in order to ensure processing quality, the production workshop should try to find the most possible mathematical relationship between raw ore parameters, system temperature, and processing results, and control production based on the predicted results to achieve higher output of high-quality products and reduce energy consumption, in order to further improve the economic benefits of the enterprise and contribute to environmental protection.

## 2. Research on the problem of predicting ore quality based on system temperature

### 2.1. Preparation of the model

To facilitate the establishment of subsequent models in this article, it is necessary to assume that every two hours is a production batch, and the product quality data at even number of times corresponds one-to-one to the production end time. In addition, due to various factors affecting the system temperature, there is a situation where the temperature changes significantly in the same production batch. If not processed, it will cause random errors in the subsequent prediction results.

Therefore, it is necessary to preprocess and analyze the data, and the final selected data must meet the following conditions:

1) Considering that scatter plots require a certain degree of discreteness between data, especially when the temperature values of a certain system are relatively rare. Delete temperature values with fluctuations exceeding ±1 in the same production batch, and only take the average of the remaining temperatures in that production batch.

2) Considering the refinement and convenience required for processing data, there are two types of deletion scenarios. If there are more temperature values in a certain system, and there are more than 2 temperature values in a production batch that fluctuate more than ±1, the production data of that batch can be directly discarded and the average value of other batches can be selected. Otherwise, only the values with excessive fluctuations will be deleted and the remaining temperature in the production batch will be averaged.

3) The average temperature values obtained within different units of production time will be used as the pre-set system temperature values for the production batch, providing a data basis for the establishment of the next model.

### 2.2. Modeling

This article needs to study the relationship between system temperature and ore processing quality indicators and conduct fitting analysis. Therefore, the influencing factors are two independent system temperatures T1 and T2, and the predictive variables are four independent quality indicators A, B, C, and D.

Due to the need to simultaneously examine the correlation between multiple variables, among which scatter plots are the leading tool for studying correlation. Although scatter plot matrices can observe the connections between multiple variables simultaneously, they may lose important information due to the observation of two or two variables on a planar scatter plot. A three-dimensional scatter plot is a method of studying the relationship between variables in a three-dimensional space determined by three variables. Therefore, in terms of information accuracy and visibility, 3D scatter plots are superior to scatter plot matrices. Therefore, after completing the preliminary data preparation, the system temperature I, system temperature II, and product quality were divided into x-axis, y-axis, and z-axis variables, and the scatter plots of the two types of system temperature with quality indicators A, B, C, and D were plotted. Through factor analysis, the correlation between multiple variables can be preliminarily and intuitively studied, preparing for the next step of determining multiple regression fitting.

Due to the need to predict the product indicators with the highest likelihood in this article, goodness of fit is used as the indicator for evaluating the fit function [6], and the function with the lowest fit among all fit functions is the optimal choice. By calculating the p-value, the closer the p-value is to 0, the more reliable the fitting function can be proven. In addition, if $P<0.01$, it can indicate that the parameter assumption is a very strong judgment result, so it can indicate that the obtained result is the most likely predicted value.

### 2.3. Solution of the model

By filtering through a series of constraint conditions, two system temperature settings and four quality indicator values corresponding to each production batch on a daily basis are obtained. Based on the preparation of the model in this article, the data was grouped and numerical cleaned to obtain the average system temperature and four quality index values corresponding to the production data, and the corresponding 3D scatter plot was depicted, as shown in Fig 1:

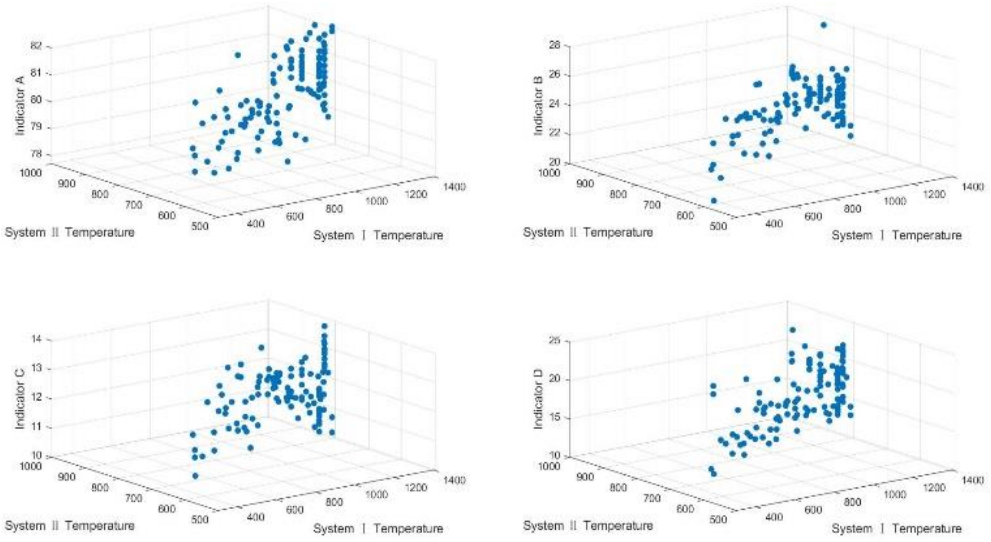



**Figure 1.** Scatter plots of system I and II temperatures corresponding to indicators A, B, C, and D, respectively

By using the poly fit curve fitting toolbox in MATLAB, the graph with the lowest fitting degree is selected to generate a multiple regression function [7]. System temperature I, system temperature II, and product quality indicators are used as x-axis, y-axis, and z-axis variables, respectively. The function with the lowest fitting degree among all available fitting functions is selected as the optimal choice, and the three-dimensional curve diagrams corresponding to the temperature and quality indicators A, B, C, and D of the two types of systems are depicted, as shown in Fig 2:

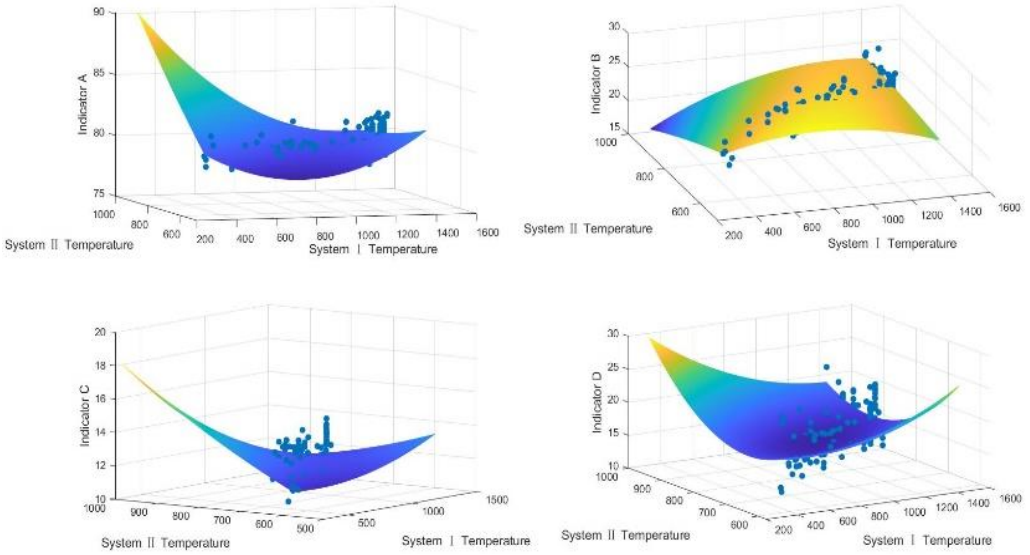



**Figure 2.** Curve plots of system I and II temperatures corresponding to indicators A, B, C, and D, respectively

The functional relationship between quality indicators and system temperature can be expressed by a binary quadratic polynomial equation:

$$Z = \alpha T_1^2 + \beta T_2^2 + \gamma T_1 + \delta T_2 + \varepsilon T_1 T_2 + \theta \quad (1)$$

According to MATLAB's testing of product quality indicators A, B, C, and D, the P-values of fit obtained are p1=0.0012, p2=0.0085, p3=0.00017, and p4=0.0041, respectively. Since p<0.01 can indicate that the parameter assumption has a very strong judgment result, based on the test results, it can be concluded that the selected function of this model has a very high degree of fit, which can indicate that the obtained result is the most likely predicted value.

# 3. Research on temperature issues in ore quality estimation systems

## 3.1. Preparation of the model

Due to the need to consider that the same set of ore quality can be obtained by multiple temperature control methods in this article, it is necessary to study the scatter plot between system temperature and indicators to clarify the interactive relationship between the two types of system temperatures. [8] Furthermore, it is necessary to consider that the temperature factors of the two types of systems have a completely different impact on the results when they act together and when they act separately. Therefore, it is necessary to add an interaction term to the original basic model to establish an interactive regression prediction model.

The scatter plot between the temperature of each system and the four indicators is shown in Fig 3 and Fig 4:

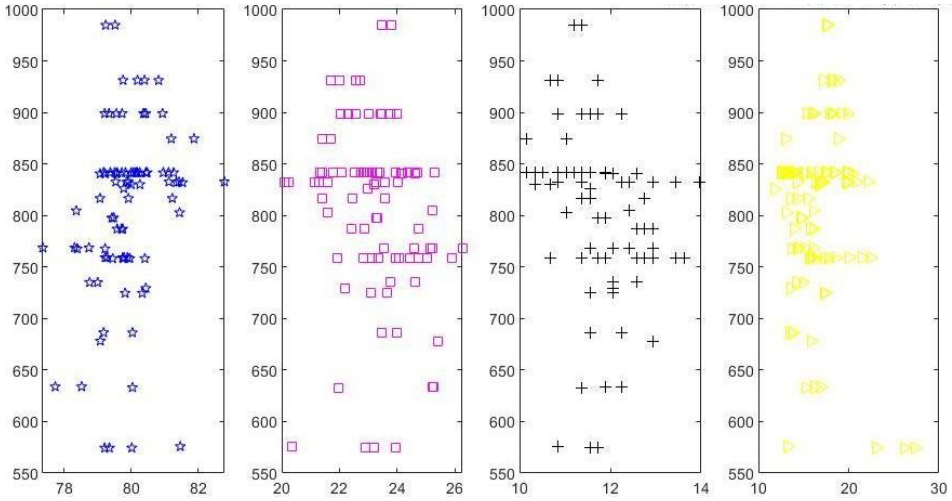


**Figure 3.** Scatter plot corresponding to system I temperature and indicators A, B, C, and D

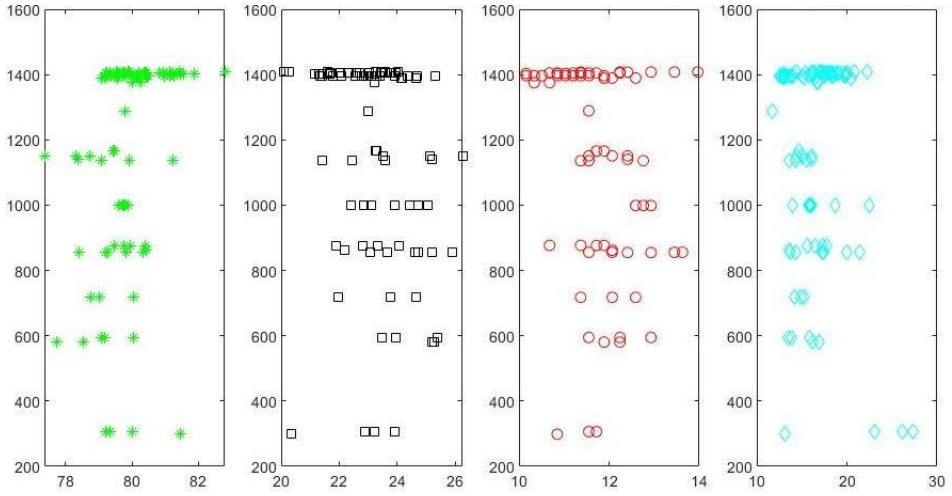


**Figure 4.** Scatter plot corresponding to system II temperature and indicators A, B, C, and D

In summary, it can be concluded that there is a significant similarity between the temperatures of the two types of systems. In the case where the quality of the same group of products can be obtained by multiple temperature control methods, the temperatures of the two types of systems affect each other. More importantly, based on the discovery of characteristic laws, the conclusion of this article can be verified from another fuzzy angle based on the high-frequency values of temperature and indicators.

### 3.2. Modeling

Due to the fact that the preparation of the model can reveal the mutual influence between the temperatures of two types of systems, and when there is an interaction, it is necessary to consider both the individual factors and the magnitude of the common factors separately. Therefore, it is necessary to add an interaction term to the model to make the fitting effect more superior. By obtaining a confidence interval, it is verified that the reliability of the solution is the most likely system temperature combination.

### 3.3. Solution of the model

In the preparation of the model in this article, through the aggregation, prominence, and similarity between the temperatures and indicators of the two types of systems, it can be concluded that there is a significant unified connection between the temperatures of the two types of systems, and that the temperatures of the two types of systems can affect each other when the quality of the same set of ore can be obtained through multiple temperature control methods.

Due to the fact that the quality of the same group of ore can be obtained by multiple temperature control methods, it is necessary to consider that the temperature factors of the two systems have a completely different impact on the results when they act together and when they act alone, that is, interaction [9]. Therefore, an interaction term should be added to the original basic model to build an interchange cross regression prediction model.

$$T = \alpha Z_a + \beta Z_b + \gamma Z_c + \delta Z_d + \epsilon \tag{2}$$

The corresponding interaction effects of the two types of systems are shown in Fig 5 and Fig6:

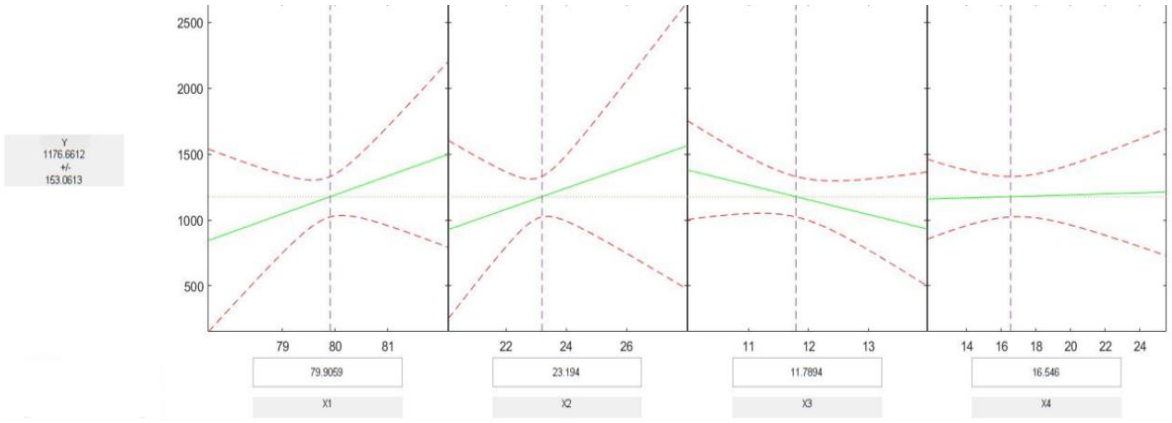



**Figure 5.** Interaction effect diagram between system I temperature and indicators A, B, C, and D

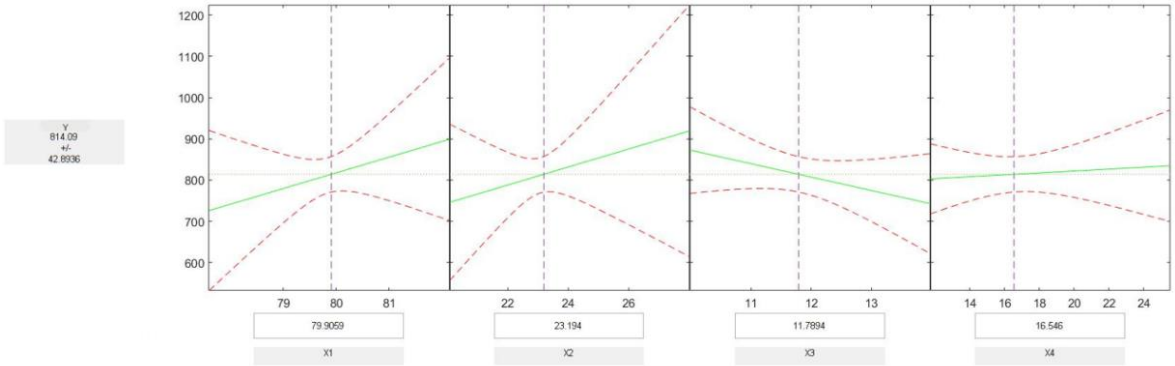



**Figure 6.** Interaction effect diagram between system II temperature and indicators A, B, C, and D

The fitting degree $R^2$=0.8 of the interactive regression prediction function [10] for the system temperature combination based on product indicators can be obtained from MATLAB, indicating that the parameter assumption has a very strong judgment result. Therefore, based on the test results, the selected function of this model has a very high fitting degree, which can indicate that the obtained

result is the most likely predicted value. Due to the confidence interval [11] referring to the error range between sample statistical values and overall parameter values at a certain confidence level, which verifies the strong reliability of the solution, the most likely system temperature combination is obtained.

## 4. Research on the qualified rate of ore quality indicators

### 4.1. Preparation of the model

Due to the large amount of data that needs to be processed in this article, in order to facilitate the further establishment of the model, it is necessary to normalize the data and establish a unified expression between the temperature of each system, product indicators, raw ore parameters, and process data [12]. Through preliminary observation, it was found that process data can reflect the quality of raw ore, and the effect of fitting product quality through process data is far better than that of fitting product quality through raw ore quality. At the same time, due to the different time nodes of each data, the average value of temperature, product quality, and process data for each day was calculated, and qualified products were selected based on the qualification conditions of the four indicators [11], in order to calculate the qualification rate for each day. Thus, providing a data foundation for the establishment of the next step model.

### 4.2. Modeling

Through the model preparation in this article, an interactive regression model that determines product quality based on raw ore quality and system temperature can be obtained. Similarly, it is necessary to consider that the temperature factors of the two systems, when acting together with raw ore quality, have a completely different impact on the results compared to when acting alone, that is, the interaction effect. Therefore, a new raw ore parameter C should be added to the original basic model to establish an interactive regression prediction model for the qualification rate G.

$$G = \alpha T_1 + \beta T_2 + \gamma C_1 + \delta C_2 + \theta C_3 + \vartheta C_4 + \rho \tag{3}$$

The interactive effect diagram of obtaining the qualification rate of quality indicators based on the raw ore quality and system temperature is shown in Fig 7:

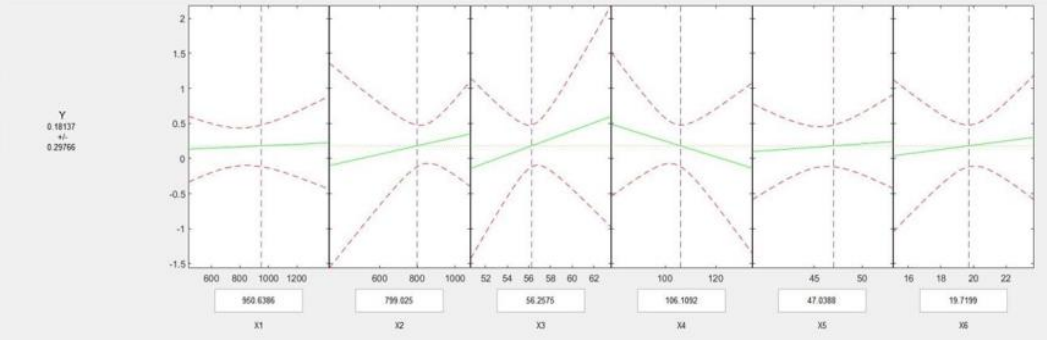



**Figure 7.** Interactive effect diagram of raw ore quality, system temperature, and quality index qualification rate

According to the MATLAB interactive regression prediction function, the confidence interval for obtaining the qualification rate can be obtained, which can evaluate the high accuracy and robustness of the selected interactive regression prediction model.

## 5. Conclusion

This article uses various mathematical modeling methods such as processing quality control, multiple regression analysis, and goodness of fit testing to conduct in-depth research on the control of workshop ore processing quality. The start and end times of daily production batches are set with

a production unit of two hours. While retaining data accuracy, a unified daily production batch is established to process a large amount of data. More intuitively established data characteristics under different system temperatures and product quality indicators, and clearly reflected the aggregation, prominence, and similarity of ore quality indicators at different system temperatures. Based on this, an interactive regression prediction function was obtained using MATLAB to establish a model. Multiple regression expressions were obtained between the system temperature and the ore quality indicators, as well as the qualification rate of the indicators. The goodness of fit P values was all less than 0.01. At the same time, a visual and predictable product quality control model was obtained with known raw ore parameters, processing voltage, and processing water pressure, in order to improve the qualification rate during the ore processing. Reduce resource waste and further improve the economic efficiency of enterprises.

## References

[1] Li Zhongjie. Application of Mathematical Models in Mineral Processing - Review of "Mathematical Models for Mineral Processing" [J]. Non ferrous Metals Engineering, 2021, 11 (03): 153-154.

[2] Yang Jide, Wang Dongli, Wu Xiaoli. Quality Process Control of Gold Ore Processing and Testing [J]. World Nonferrous Metals, 2018 (13): 289-290.

[3] Chen Yonghong, Meng Xianwei, Liu Zhenghong, Wang Lichen. Progress in the Analysis and Determination of Gold in China from 2017 to 2018 [J]. Gold, 2020, 41 (01): 82-90.

[4] Yang Jinyan. Research on Factors Influencing the Benefits of Mining Enterprises Based on Economic and Technical Indicators [J]. World Nonferrous Metals, 2023 (06): 223-225.

[5] Liao Shiwen. Innovative sintering process technology to improve ore quality and reduce production costs [J]. Metallurgy and Materials, 2023, 43 (07): 127-129.

[6] Hu Chunyan, Hu Liangping. How to Use It Correctly $\chi\sim$ 2-test - goodness of fit test and SAS implementation [J]. Sichuan Mental Health, 2021, 34 (05): 417-423.

[7] Guo Juan. A Multiple Linear Regression Model for Predicting Automotive Fuel Consumption Based on the Existence of Interaction Terms [J]. Guangxi Quality Supervision Guide, 2019, (11): 138-140.

[8] Zhang Shuang. Estimation of Multiple Interaction Terms in Nonlinear Models [J]. Computer Knowledge and Technology, 2010, 6 (21): 58225823.

[9] Ma Meng. Fundamentals and Practice of Quantitative Financial Analysis in MATLAB [M] MATLAB Engineering Application Library, 2018, 06.

[10] Zhang Yuanyuan. Comparison and Application of Several Simultaneous Confidence Intervals [D]. Lanzhou University, 2021.

[11] Liu Yaohui, Xu Huiying, Chen Qipeng, Zhan Peida. Problem solving ability measurement and data analysis method based on process data [J]. Progress in Psychological Science, 2022, 30 (03): 522-538.

[12] Luo Xuehui, Zhang Yong, Chen Zhansheng, Li Xuanhui, Chen Xue. Quality Process Control of Gold Ore Processing and Testing [J]. Gold, 2012, 33 (03): 61-64.