Analytical Chemistry: A Journal (2022)
Clausius Scientific Press, Canada
DOI: 10.23977/analc.2022.010108
ISSN 2616-1958 Vol. 1 Num. 1

Study on Composition Analysis and Identification of Ancient Glass Products

Shengqiang Yao[1], Yang Nie[1], Wei Zhu[1], Xiaojing Yang[2]

[1]*School of International Business, Zhejiang Yuexiu University, Shaoxing, 312069, China*
[2]*School of Information Technology, Hebei University of Economics and Business, Shijiazhuang, 050061, China*

Keywords: Grey correlation, k-means clustering, contour, chi-square tes

Abstract: Affected by the weathering process, the internal elements in ancient glass will be exchanged with the external environmental elements, resulting in a change in the proportion of chemical components, but there will be a certain correlation between the exchange of elements. In this paper, a mathematical model for component analysis and prediction using the K-means clustering model and BP neural network is established.

1. Introduction

The Silk Road was a road of cultural exchange between China and the West in ancient times, in which glass was valuable material evidence of early trade and friendly exchanges. At that time, glass was generally made into exquisite ornaments in West Asia and Egypt and introduced into China. After learning glass-making technology, the ancients of China used local materials to make glass products, so they looked similar to glass products in other regions, but their chemical composition was different.

With more and more ancient glass products unearthed, the determination of glass properties is a key point in the study of ancient glass. Because the difference in chemical composition will affect the type of glass and lead to the difference in texture, in order to determine the type and decoration of glass, it is necessary to analyze the properties of unknown glass unearthed according to some existing glass types, decorations and chemical compositions.

In referenc [1], the data on the chemical composition of more than 100 ancient glass samples unearthed in Chongqing, Sichuan, Guangxi, Guizhou, Guangdong and other regions of China were analyzed and processed by a multivariate statistical method, and the glass production status of each region was pointed out by factor analysis. This experimental data processing method provides a new way with the study ancient glass.

However, there are many factors affecting the glass, and more appropriate models need to be introduced. This paper focuses on the correlation between the chemical composition and the external properties of glass products to analyze and predict the properties of glass before weathering and establishes a chemical composition analysis model to analyze and predict the chemical composition of ancient glass by using Matlab to draw three-dimensional curved surface, contour lines and BP neural network prediction methods. Various chemical costs of the ancient glass can be efficiently and accurately analyzed and predicted. This paper provides a new way to study ancient glass.

2. Model building and solving

2.1. Weathering chemical composition model

2.1.1. Frequency analysis

For four variables, the degree of surface weathering is plotted against the remaining three variables. In this paper, the frequency analysis histograms of the variables, which are drawn by importing the given data through Matlab, and the correlation between the three variables and the surface weathering degree is preliminarily analyzed. [3]By observing the three frequency analysis charts and comparing the frequency before and after weathering, it can be seen that the change of weathering degree and the fluctuation of colour are relatively large, so it can be concluded that the glass colour is related to the surface weathering degree.as shown in Figure 1 and Figure 2.

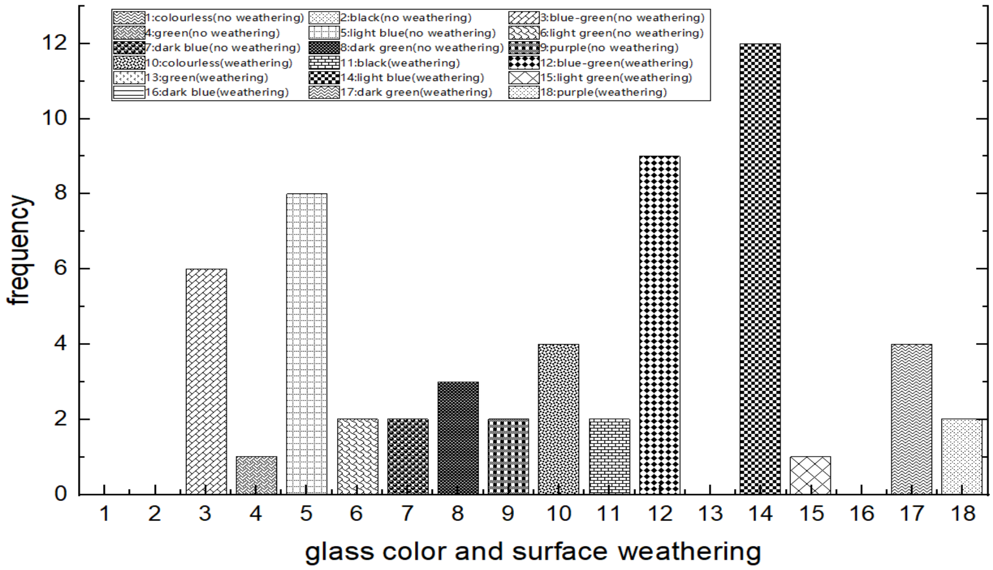

Figure 1: Glass color and surface weathering

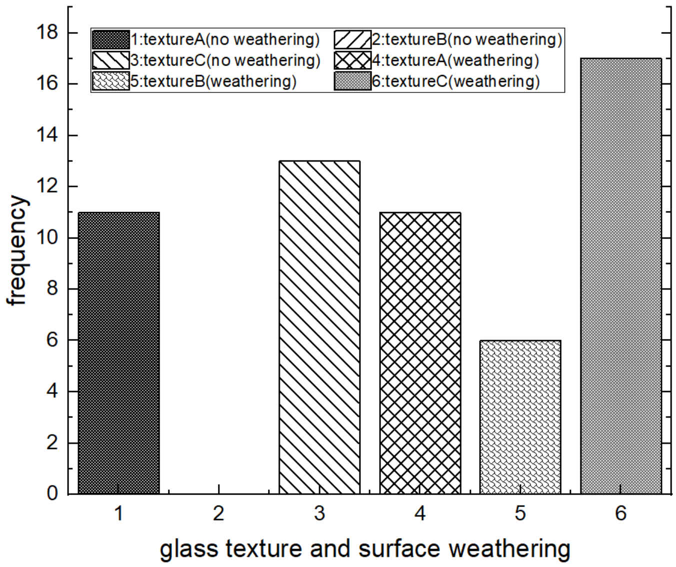

Figure 2: Glass ornamentation and surface weathering

2.1.2. Chi-square test model

According to the frequency analysis chart, it can be preliminarily determined that there is a correlation between colour, but it can not be ruled out that there is also a correlation between ornamentation and glass type and weathering, so further analysis and test are needed, that is, chi-square test. The results of the chi-square test are shown in Table 1. By observing P in the table, it can be seen that the P value of glass type is less than 0.05, while the P value of ornamentation is greater than 0.05. Therefore, it can be concluded that there is a correlation between glass type and weathering degree.as shown in Table 1.

Table 1: Chi-square test indicators of ornamentation and type

Indicators	Glass type	Glass ornamentation
chi(2)	6.88	15.19
P	0.009	0.11

2.1.3. Normality test

According to the divided data table, draw the frequency distribution histogram of the correlation between the surface weathering degree and the chemical composition content index of high-potassium glass and lead-barium glass respectively, as shown in Figure 3 and Figure 4. By observing the data in the figure, [4] it can be seen that the chemical composition of high-potassium glass shows a downward trend after weathering, while that of lead-barium glass shows an upward trend on the contrary.as shown in Figure 3 and Figure 4.

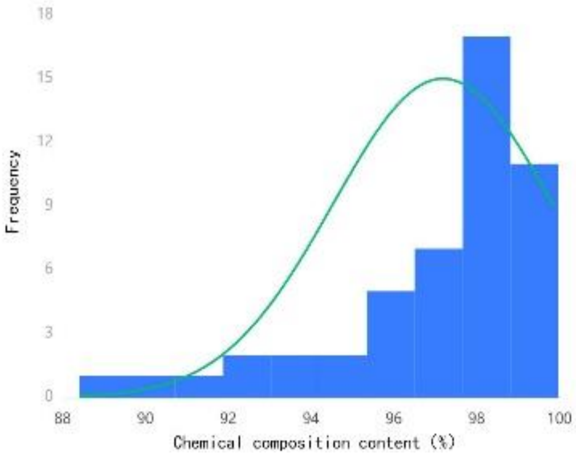

Figure 3: Frequency distribution of lead and barium

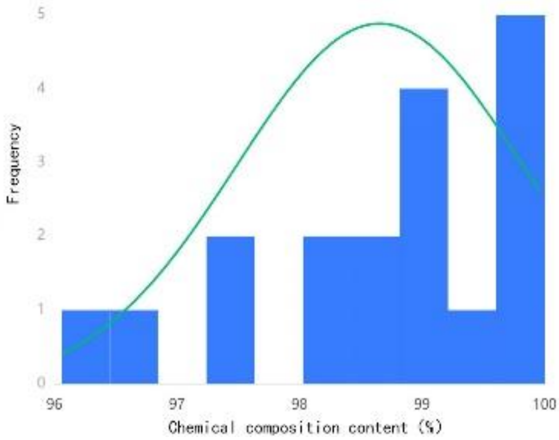

Figure 4: Frequency distribution of high potassium

2.1.4. Surface fitting

Due to the lack of available sample data, the trend of each chemical component's content under different conditions can be obtained intuitively by means of the method of turning data into geometric changes. Here, considering the correlation of data, the content of each chemical component with or without weathering of high potassium is fitted with a three-dimensional curved surface by Matlab, as shown in Figure 5 and Figure 6.

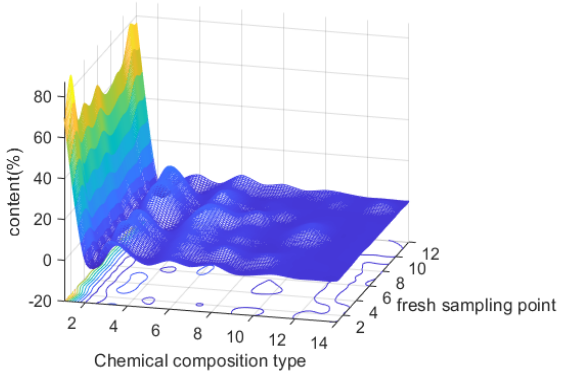

Figure 5: Unweathered components with high potassium

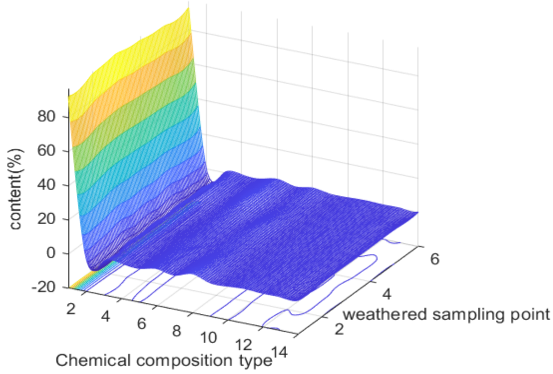

Figure 6: Weathered components with high potassium

Contour processing of the 3D surface plot before and after weathering, as shown in Figure 7.8.

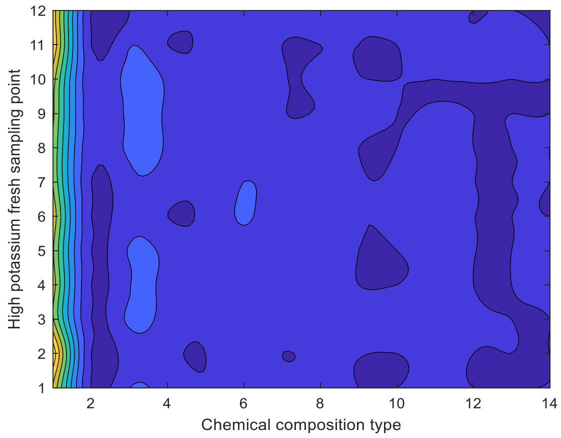

Figure 7: Unweathered components with high potassium

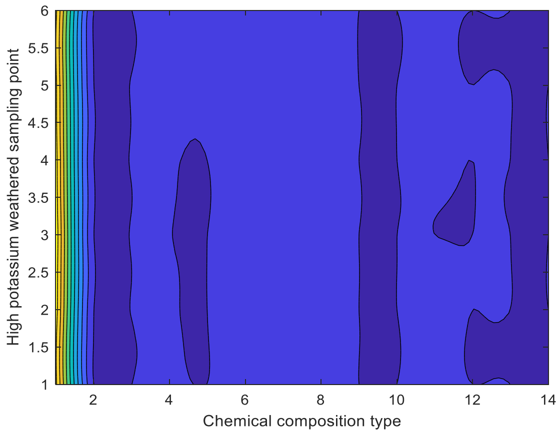

Figure 8: Weathered components with high potassium

2.2. K-means clustering

The method comprises the following steps: carrying out sub-classification on five types of silicon dioxide, potassium oxide, aluminium oxide, lead oxide and phosphorus pentoxide by a preliminary hypothesis, carrying out K-means clustering on the components in the high-potassium glass and the lead-barium glass by dividing glass types, randomly initializing K clustering centres, selecting data to be analyzed, measuring the distance between the data and the K clustering centres, The data is divided into the cluster corresponding to the nearest cluster centre, [6] and the corresponding distance is obtained by calculating the average value of each cluster and taking the average value as a new cluster centre.

$$C_t = \frac{\sum_{X_i \epsilon S_j} X_i}{|S_j|}$$

In the formula, represents the center of the jth cluster, represents the number of objects in the jth cluster, and represents the ith object in the J cluster.$C_j 1 \le j \le k |S_j| X_i 1 \le i \le |S_j|$.

After the K-means algorithm is carried out, the distance D of each point is calculated. The calculation results are shown in Table 2 and Table 3 (see the appendix). After integration and averaging, it is obtained that for high potassium, the average distances of potassium oxide, aluminium oxide and lead oxide are 398, 272 and 411 respectively, which are close and have great influence. Subclassification of these three components in high potassium glasses; For lead-barium, the distance averages of silicon dioxide and phosphorus pentoxide are 733 and 824, respectively, which are small in their total composition, so lead-barium glass subclassifies these two components.

2.3. BP neural network prediction model

For the glass relics of unknown types, the indexes shall be reasonably divided first, and the glass of unknown types shall be predicted by using BP neural network. According to the given data, the proportion of each chemical composition screened out by multiple data and cluster analysis shall be taken as the input value, and the glass type shall be taken as the output value. The prediction result is shown in the figure 9.

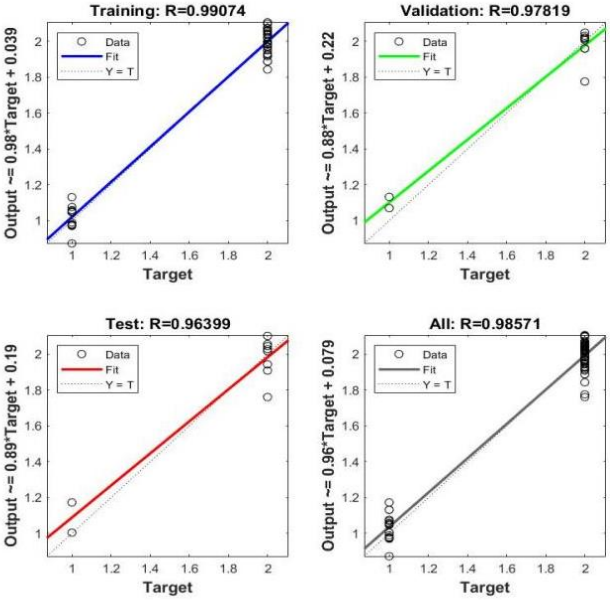

Figure 9: Training results of BP neural network

Before prediction, the sample data is trained. To get the training result with a smaller error, the training range is set to 70%, and the sample verification and sample test are set to 15%. Ten groups of samples are verified and tested, and the Scaled Conjugate Gradient algorithm, which runs slowly but has higher satisfaction, is selected for training. [7]The correlation coefficient for the first time is 0.88, which is lower than 0.9, so the training result is not ideal. Therefore, the sample data are trained repeatedly. In the third training, the training result is as shown in Figure 10, and the correlation coefficient is as high as 0.99. At this time, the output value can be predicted, and the training results are retained. The data of each chemical component of the unknown type of cultural relics are input to predict the results. The predicted values are shown in Table 2, and the predicted results are shown in Table 3.

Table 2: Type values of unknown artifacts

Number	Type (1-high potassium, 2-lead barium)
A1	0.9012
A2	1.6371
A3	1.9945
A4	1.9803
A5	1.9501
A6	1.0236
A7	1.0991
A8	1.9496

Table 3: Glass Types of Unknown Artifacts

Number	Type (1-high potassium, 2-lead barium)
A1	High potassium
A2	Lead and barium
A3	Lead and barium
A4	Lead and barium
A5	Lead and barium
A6	High potassium
A7	High potassium
A8	Lead and barium

2.4. Grey correlation evaluation

In this paper, the grey correlation evaluation is used to analyze the correlation degree between the indexes. The grey correlation evaluation is to determine whether the data are closely related by determining the data column and comparing the similarity of the geometric shapes of several data columns, which can be used to reflect the correlation degree between the curves.

$$\xi_i(k) = \frac{min_i min_k |a_0(k) - a_i(k)| + \rho * max_i max_k |a_0(k) - a_i(k)|}{|a_0(k) - a_i(k)| + \rho * max_i max_k |a_0(k) - a_i(k)|}$$

$a_0(k)$ is the reference column; p is the discrimination coefficient, with a range of (0, 1), which is used to control the discrimination. The smaller the p-value is, the greater the discrimination is. The larger the p-value is, the smaller the discrimination is. Usually taken as 0.5.

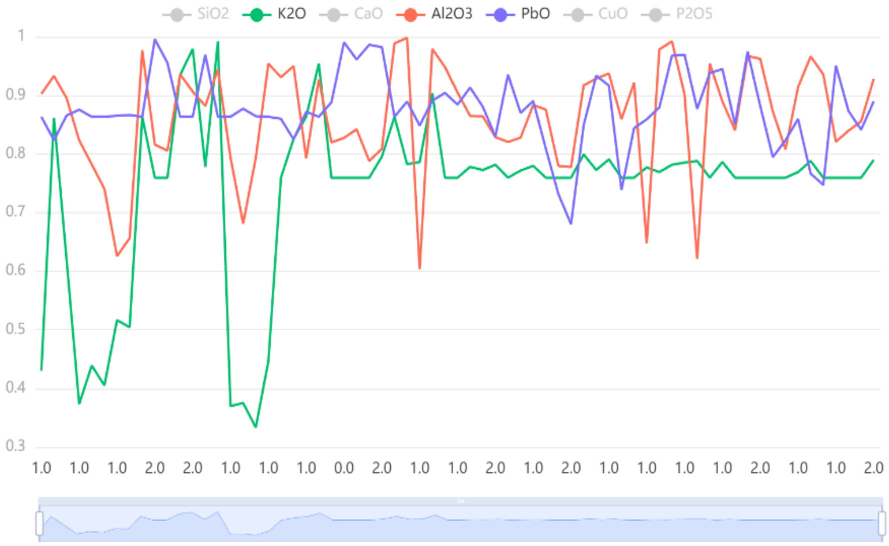

Figure 10: Correlation degree of high potassium subclass components

According to the correlation analysis of each component and glass type, the glass type is taken as the parent sequence, and the correlation analysis diagram of sub-classified components of high-potassium glass is shown in Figure 10. By observing the correlation degree in the diagram, [8] it can be seen that the correlation degree of lead oxide is relatively high, and the specific correlation degree analysis of each component is shown in Table 4. [8] According to the larger the correlation degree, the stronger the correlation with the parent sequence, which also represents the higher the evaluation. From the analysis of Table 4, it was found that the lead oxide (PbO) was evaluated the highest, followed by the phosphorus pentoxide (P_2O_5).

Table 4: Correlation results between glass type and chemical composition

Correlation degree result		
Evaluation item	Correlation degree	Rank
Lead oxide (PbO)	0.873	1
Silica (SiO_2)	0.866	2
Alumina (Al_2O_3)	0.86	3
Calcium oxide (CaO)	0.834	4
Copper oxide (CuO)	0.816	5
Phosphorus pentoxide (P_2O_5)	0.799	6
Potassium oxide (K_2O)	0.728	7

3. Conclusion

This paper mainly studies the analysis of the composition of ancient glass products, in the face of the exchange of complex internal elements and external environmental factors, through the calculation of the composition of unknown types of glass products, analysis and prediction of the external characteristics of the products. The contour map is drawn by three-dimensional surface fitting to show the trend of variables, and different models are used to predict unknown data to improve the accuracy of the analysis process.

The total number of data samples here is relatively small. First, interpolate the missing values of some data samples to improve the data set. If the total number of available sample data is considerable, the proportional relationship of linear average weighting can be used, and the normal distribution function can be used as the weighting function, which can solve the predicted value more simply and effectively. When clustering the content of different chemical elements, it belongs to the unsupervised learning range, so we can consider using K-means algorithm to automatically optimize the best number of clusters, and analyze the ratio of data set tightness and separation, the ratio of intra-cluster and inter-cluster distance and contour coefficient to evaluate the quality of the model.

References

[1] Chen Tiemei, Wang Jianping. Composition determination, data processing and archaeological interpretation of ancient ceramics [J]. Heritage Conservation and Archaeological Science, 2003 (04): 50-56.

[2] Zheng Meiling, Liu Qianjin, Yang Jinchu, Li Ruili, Xu Kejing, Li Yaoguang, Du Jia, Zhang Junsong. Comprehensive evaluation of aroma quality of different sweet potato extracts based on principal component analysis and cluster analysis [J]. China Food Additives, (3): 196-206, 2022.

[3] Peng Xiangang, Lai Jiawen, Chen Yi, Intelligent identification method of customer power consumption mode based on clustering analysis [J]. Power System Protection and Control, 42 (19): 69-73, 2014.

[4] Li Zhenkun, Ren Yuliang, Wang Fashun, Yang Xiu, Mi Yang, Investment benefit evaluation and sensitivity analysis of incremental distribution network based on annual time series simulation [J]. Power Grid Technology, 43 (6): 2164-2170, 2019.

[5] ZHOU Wenhao, ZENG Bo. Review of Grey Relational Model [J]. Statistics and Decision, (15): 29-34, 2020.

[6] CHen Ganghua, HU Cong, ZENG Yali, MA Zhonggao, Logging identification method of fracture-cave fillings in carbonate reservoirs based on BP neural network [J]. Petroleum Geophysical Prospecting, 54 (1): 100-104, 2015.

[7] Ren Jiaguo, Gong Ke, Ma Fujun, Gu Qingbao, Wu Qianqian. Prediction of heavy metals and PAHs contents in soils of contaminated sites based on BP neural network [J]. Research of Environmental Sciences, 34 (9): 2238-2247, 2021.

[8] Cheng Qiyun, Sun Caixin, Zhang Xiaoxing, et al. Short-Term load forecasting model and method for power system based on complementation of neural network and fuzzy logic [J]. Transactions of China Electrotechnical Society, 2017, 19(10): 53-58.