Analysis and Prediction of Chemical Composition of Ancient Glass Relics

Tao Wan[1, *], Wen Qi[2], Kangqi Cheng[2]

[1]School of Management Science and Engineering, Anhui University of Finance and Economics, BengBu, China

[2]School of economics, Anhui University of Finance and Economics, BengBu, China

*Corresponding author: wtttt1009@163.com

Abstract. Glass was introduced into China long ago via the Silk Road. After the ancient glass was buried, its internal elements would change under the influence of the environment. It is of great significance for archaeological work to study the changes of these chemical components. In this paper, the chemical composition of ancient glass products is analyzed and predicted by using R-type clustering, BP neural network, multiple linear regression and other models.First, after using SPSS statistical data, Pearson chi square test was conducted to draw a conclusion that the category, decoration and color were not related to whether the surface of the cultural relics was weathered. Then, descriptive statistics are made on the data. Finally, the multiple linear regression model is used to predict the unknown type of cultural relics. The results show that after the weathering of cultural relics, the content of potassium oxide, calcium oxide, aluminum oxide and magnesium oxide decreases compared with that before weathering, while the content of iron oxide increases.The main chemical components of high potassium glass relics and lead barium glass relics are calculated by entropy method. Compared with silicon dioxide and aluminum oxide, they have less influence on other chemical components. Therefore, these two chemical components are eliminated to classify cultural relics. Then, the R-type clustering model is established to classify the cultural relics, and the high potassium weathering type is obtained: 22 cultural relics sampling points form a category, and the other cultural relics form a category.

Key words: Composition of glass products, BP neural network, Multiple linear regression model, Python.

1. Introduction

Weathering refers to the process in which materials near the surface change physically or chemically in contact with the atmosphere and water to form loose deposits, which is a form of energy conversion. The weathering of glass is due to the loss of crystal water in the natural state of the glass, and the appearance of foggy film, or pointy, fine line containing weathering substances on the surface of the glass, which reduces the light transmittance of the glass[1]. It is worth noting that the causes, processes and results of weathering of glasses with different chemical compositions are different. The composition proportion of ancient glass buried underground will change when it is weathered. Based on the existing basis, our team will analyze the composition of weathering of ancient glass through the data provided by archaeologists, and identify its type before weathering.

China is one of the countries that produced and manufactured glass earlier in the world. The glass relics unearthed can be traced back to the Shang and Zhou Dynasties. However, the production system of lead barium glass in Chu culture is mainly PbO-SiO_2 and PbO-BaO-SiO_2, which are brittle and fragile, so they can not be used as daily appliances and can not be well integrated into folk life. However, thanks to the influence of the Silk Road, soda lime silicate glass from West Asia, India and other regions was introduced into China, which improved the lead barium glass made in China since the Spring and Autumn Period and the Warring States Period[2]. Since the 18th National Congress of the Communist Party of China, China has placed more emphasis on cultural confidence and a community of shared future for mankind. The identification and analysis of the composition of ancient glass products can provide sufficient evidence for the ancient civilization of China, and also help the exchanges between China and other countries in the world. On this basis, it is also conducive to the development and improvement of glass products, promoting the development of China's industry, and also providing strong help for the archaeological community.

2. Model establishment and solution

First, the data is visualized and chi square test is performed to verify its relevance; Secondly, through the classification and descriptive statistics of the data, the statistical rules are obtained; Finally, a multivariate linear regression model is established to predict the chemical composition of the weathered part when it is not weathered[3].

Since the above variables are nominal variables, it is difficult to obtain the correlation between them through observation, and the amount of data that can be calculated is small. The sample size of the types, patterns and colors of glass relics is small, which is less than 40. Therefore, chi square test is used to verify their correlation. The chi square test hypothesis and alternative hypothesis are:

H_0: There is a significant positive correlation between these variables;

H_1: There is no significant positive correlation between these variables

After hypothesis testing is set, the expected frequency of type, texture and color is calculated, and then the chi square value and degree of freedom are calculated by the formula. Finally, the hypothesis is verified by checking the P value.

Use SPSS software to perform chi square test on the above three correlations, and the results are shown in Table 1and Table 2:

Table 1. chi square test of three correlations

name	Pearson	p
High potassium lead barium	2.000	0.157
A	/	/
B	6.000	0.199
C	/	/

Table 2 chi square test of three correlations

<table>
<tr><th>name</th><th>Pearson</th><th>p</th></tr>
<tr><td>black</td><td>/</td><td>/</td></tr>
<tr><td>Blue-green</td><td>/</td><td>/</td></tr>
<tr><td>Light blue</td><td>/</td><td>/</td></tr>
<tr><td>Light green</td><td rowspan="2">35.00</td><td rowspan="2">0.243</td></tr>
<tr><td>Dark Blue</td></tr>
<tr><td>Dark green</td><td>/</td><td>/</td></tr>
<tr><td>purple</td><td>/</td><td>/</td></tr>
<tr><td>No color record</td><td>/</td><td>/</td></tr>
</table>

According to the SPSS analysis, in the correlation test between the glass cultural relic category and whether the cultural relic surface is weathered, the Pearson value is 2.000, and the significance value is 0.157 greater than 0.05, there is no significant difference, that is, the glass type is not related to whether the cultural relic surface is weathered; In the correlation test between the glass cultural relics' ornamentation and whether the cultural relics' surface is weathered, the Pearson value is 6.000, and the significance value is 0.199 greater than 0.05, there is no significant difference, that is, the glass cultural relics' ornamentation is not related to whether the cultural relics' surface is weathered; In the correlation test between the color of glass relics and whether the surface of cultural relics is weathered, the Pearson value is 35.000, and the significance value is 0.243, which is greater than 0.05. There is no significant difference, that is, the color of glass is not related to whether the surface of cultural relics is weathered[4].

In this paper, the data are first described and statistically analyzed, and then the chemical composition data of various categories are visualized. Through descriptive statistics, an analysis model is established to obtain statistical rules.

This paper uses EXCEL software to draw a scatter plot, as shown in Fig 1:

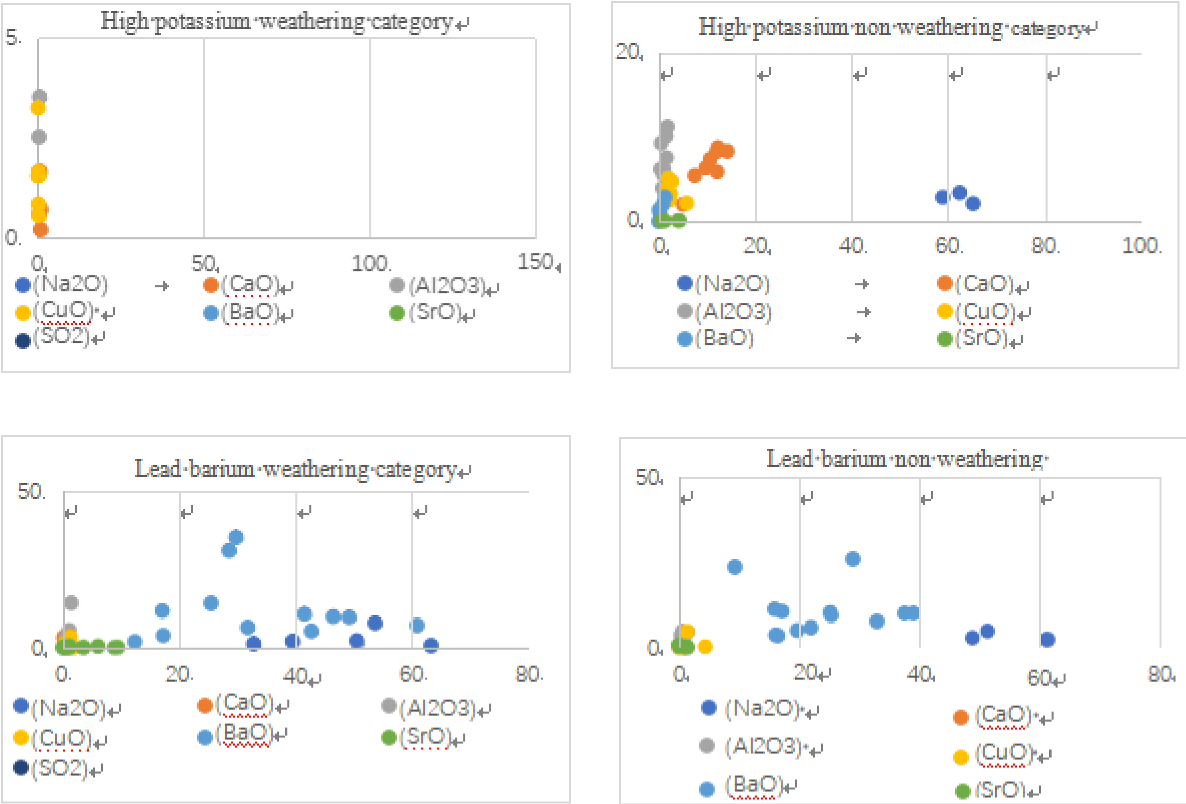


Fig. 1 Content of each component

The descriptive statistics by SPSS software are shown in Table 3. The text only shows the descriptive statistics results of high potassium weathering categories, and the rest groups are shown in the appendix:

Table 3. High potassium weathering

name	minimum value	Maximum	mean value	standard deviation	variance
silicon dioxide	92.35	96.77	93.9633	1.73362	3.005
sulfur dioxide	0	0	0	0	0
Stannic oxide	0	0	0	0	0
Copper oxide	0.55	3.24	1.5617	0.93482	0.874
ferric oxide	0.17	0.35	0.265	0.0695	0.005
Sodium oxide	0	0	0	0	0
Potassium oxide	0	1.01	0.5433	0.44518	0.198
calcium oxide	0.21	1.66	0.87	0.48777	0.238
magnesium oxide	0	0.64	0.1967	0.30631	0.094
alumina	0.81	3.50	1.93	0.96449	0.93
Lead oxide	0	0	0	0	0
Barium oxide	0	0	0	0	0
Phosphorus pentoxide	0	0.61	0.28	0.20995	0.044
Strontium oxide	0	0	0	0	0

In order to make descriptive statistics more comprehensive and intuitive, for four types of glass, namely, high potassium weathered glass, high potassium non weathered glass, lead barium weathered glass and lead barium non weathered glass, this paper removes the statistical description of the chemical composition of silicon dioxide, so as to avoid the neglect of other statistics due to the excessive content of this chemical composition, and fine tune each type of glass. According to the above figure and the above table, we can get the statistical laws of chemical components of each category. To make the results more illustrative, this paper removes the silica data in all categories in advance, and the results are shown in Table 4:

Table 4. Sequence of Main Chemical Components of Cultural Relics of Each Category

Category of glass cultural relics	Ranking of main chemical components
High potassium weathering	Copper oxide、calcium oxide、alumina
High potassium no weathering	calcium oxide、Copper oxide、alumina、Barium oxide、Sodium oxide
Lead barium weathering	Barium oxide、Sodium oxide、Strontium oxide
Lead barium not weathered	calcium oxide、Copper oxide、Barium oxide、Sodium oxide、Copper oxide、Strontium oxide

Therefore, it can be concluded that the statistical law of weathering of each category is: for high potassium type glass relics, whether the surface is weathered or not causes a large change in the proportion of chemical elements; For the lead barium type glass relics, whether the surface is weathered or not causes little change in the proportion of chemical elements[5].

2.1. Solution of model

This paper uses MATLAB software to process the data, and the results are shown in Table 5. The text only shows the predicted value of the high potassium weathering category, and the remaining data are shown in the appendix:

Table 5. Predicted Values of High Potassium Weathering Categories

number	Na2O	K2O	CaO	MgO	AL2O3	Fe2O3	CuO	PbO	BaO	P2O5	SrO	SnO2	SO2
07		8.76	6.21		5.7	1.3	0.98		1.98	1.56	0.08		
09		7.43	7.45	1.21	8.9	2.4	3.67	0.42	0.66	0.35	0.04		0.45
10	3.65	9.21	8.61	0.98	4.5	0.9	2.79	0.22		2.71	0.11		
12		6.56	4.43	1.45	6.8	1.45	0.87			0.56	0.05		0.37
22	2.31	7.43	3.45	1.34	7.9	2.02	4.56	1.34	2.21	1.54		2.45	
27		8.96	6.57	0.78	4.3	0.98	4.31			0.88	0.09		

The predicted data visualization is shown in Fig 2 below:

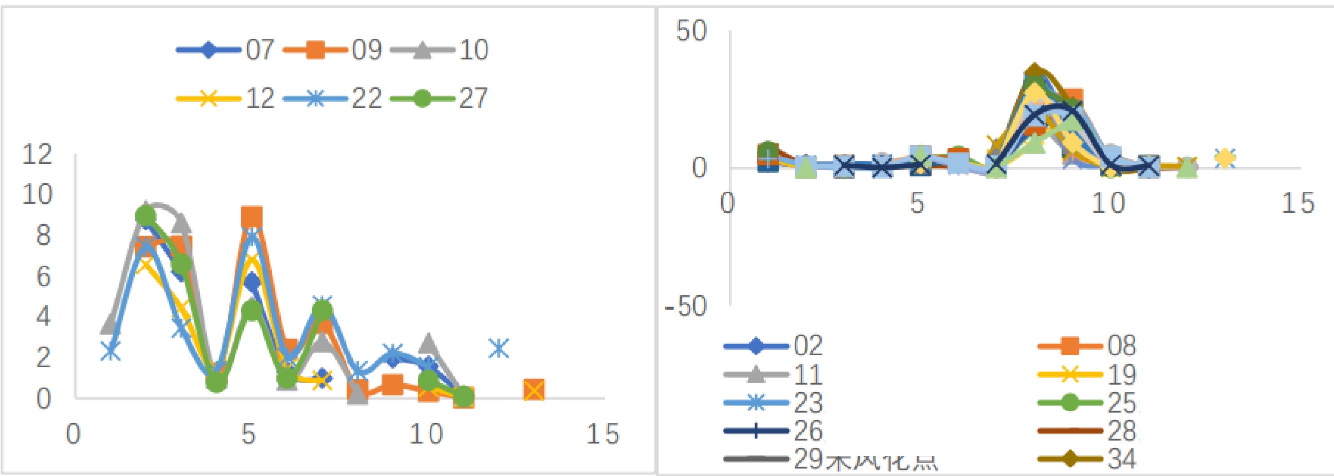


Fig. 2 Predicted data visualization diagram

Next, the parameter estimates of the model are tested, and the model is statistically tested. In this paper, the goodness of fit test and the significance test of the regression equation are used to calculate the R^2 value and the F value of the model. In this paper, through SPSS software, the R^2 value is 0.812, and the F value is 0.863, and the results with a high degree of fit are obtained.

3. Entropy data preprocessing

First, we preprocess the data, use entropy method to screen the chemical components of the cultural relics sampling points, remove the chemical components that occupy a small weight, and then calculate the correlation coefficient matrix between each chemical component, use the R-type clustering model to give the classification results [6-7].

In this paper, the appropriate chemical composition is selected for each category of cultural relics sampling points for analysis. Therefore, the entropy method is selected to screen the chemical composition of different categories of cultural relics sampling points. Under the condition of ensuring a high degree of interpretation of whether the explained variables are weathered, the chemical composition with a small weight is removed[8].

3.1. Select sample data

Select *i* indicators of chemical composition, a total of *j* cultural relics sampling points, then X_{mn} is the percentage of the nth chemical composition in the nth cultural relics sampling sample, $m = 1,2,3\cdots i$, $n = 1,2,3\cdots j$.

When the measurement direction and unit of each chemical component are inconsistent, it is necessary to standardize the data of cultural relics sampling points. The processing of positive and negative data is as follows:

$$X = \frac{X_{mn} - Min(X_{mn})}{Max(X_{mn}) - Min(X_{mn})} \tag{1}$$

$$X = \frac{Max(X_{mn}) - X_{mn}}{Max(X_{mn}) - Min(X_{mn})} \tag{2}$$

Calculate the proportion of the m-th cultural relic sampling point in the chemical composition under the n-th chemical composition value. The formula for calculating the sample weight is as follows:

$$P_{mn} = \frac{X_{mn}}{\sum_{n=1}^{j} X_{mn}} \tag{3}$$

Calculate the entropy value of the nth chemical component. The formula for calculating the entropy value is as follows:

$e_n = -K * \sum_{m=1}^{i} \left(P_{mn} * \ln(P_{mn})\right), K = \frac{1}{\ln(i)}$,Where i is the number of samples.

The information utility value of a cultural relic's chemical composition value depends on the difference between the value's information entropy e_n and 1, and its value directly affects the weight of the chemical composition. The greater the value of information utility, the greater the weight of this chemical component. The formula is as follows:

$$dif_n = 1 - e_n \tag{4}$$

The entropy method is used to calculate the weight of each chemical index[9]. The essence is to calculate it through the difference coefficient of chemical composition, which is positively correlated. The weight formula of the nth index is as follows:

$$wht_i = \frac{dif_j}{\sum_{n=1}^{j} dif_n} \tag{5}$$

3.2. Solution of model

Cultural relics are divided into four categories: high potassium weathered glass, high potassium fresh glass, lead barium weathered glass and lead barium fresh glass. SPSS is used to calculate the chemical composition of each category of glass by entropy method. The weight of each category of chemical composition is shown in Table 6 and Table 7:

Table 6. Chemical composition weight of high potassium weathered and non weathered categories

Name of chemical composition	Information entropy	Information utility value	weight coefficient	Information entropy	Information utility value	weight coefficient
silicon dioxide	0.994	0.006	0.15%	0.9967	0.0033	0.09%
Sodium oxide	0.42	0.58	14.15%	0.4769	0.5231	14.54%
Potassium oxide	0.8395	0.1605	3.92%	0.9451	0.0549	1.53%
calcium oxide	0.8438	0.1562	3.81%	0.893	0.107	2.97%
magnesium oxide	0.8205	0.1795	4.38%	0.8881	0.1119	3.11%
alumina	0.9317	0.0683	1.67%	0.9707	0.0293	0.81%
ferric oxide	0.7895	0.2105	5.14%	0.8317	0.1683	4.68%
Copper oxide	0.909	0.091	2.22%	0.8876	0.1124	3.12%
Lead oxide	0.6073	0.3927	9.58%	0.6902	0.3098	8.61%
Barium oxide	0.5015	0.4985	12.16%	0.5714	0.4286	11.91%
Phosphorus pentoxide	0.7962	0.2038	4.97%	0.8258	0.1742	4.84%
Strontium oxide	0.8106	0.1894	4.62%	0.8403	0.1597	4.44%
Stannic oxide	0.1451	0.8549	20.86%	0.1096	0.8904	24.75%
sulfur dioxide	0.4935	0.5065	12.36%	0.4745	0.5255	14.60%

Table 7. Chemical Composition Weights of Pb Ba Weathering and Non Weathering Categories

Name of chemical composition	Information entropy	Information utility value	weight coefficient	Information entropy	Information utility value	weight coefficient
silicon dioxide	0.9654	0.0346	1.14%	0.9861	0.0139	0.33%
Sodium oxide	0.6078	0.3922	12.90%	0.4403	0.5597	13.42%
Potassium oxide	0.7684	0.2316	7.62%	0.6999	0.3001	7.19%
calcium oxide	0.9185	0.0815	2.68%	0.7935	0.2065	4.95%
magnesium oxide	0.8705	0.1295	4.26%	0.7526	0.2474	5.93%
alumina	0.917	0.083	2.73%	0.9655	0.0345	0.83%
ferric oxide	0.7992	0.2008	6.61%	0.6137	0.3863	9.26%
Copper oxide	0.8595	0.1405	4.62%	0.6649	0.3351	8.03%
Lead oxide	0.9759	0.0241	0.79%	0.9732	0.0268	0.64%
Barium oxide	0.9178	0.0822	2.70%	0.9307	0.0693	1.66%
Phosphorus pentoxide	0.8304	0.1696	5.58%	0.6471	0.3529	8.46%
Strontium oxide	0.9275	0.0725	2.39%	0.7921	0.2079	4.98%
Stannic oxide	0.4481	0.5519	18.16%	0.4831	0.5169	12.39%
sulfur dioxide	0.1542	0.8458	27.83%	0.0854	0.9146	21.92%

The visualization diagram of the proportion of chemical components in each category is shown in Fig 3:

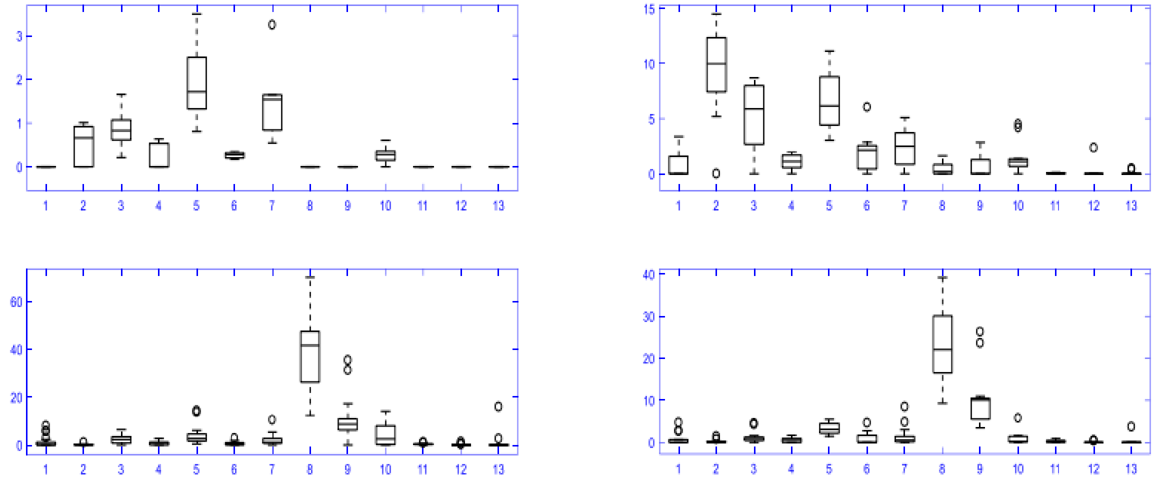


Fig. 3 Visualization Chart of Chemical Composition Proportion of Each Category

According to the specific gravity of each chemical component calculated by the entropy value method, under the condition that 95% of the component weights are accumulated in sequence, silica and alumina are excluded from the four categories of cultural relics glass, and each category is classified according to the rest of the chemical components.

Through the above process, appropriate chemical components are screened, and R-type clustering model is established, so as to realize the subclass division of each category

Standardization of chemical composition data by category

q_{mn} is used to represent the nth chemical composition value of the m cultural relics sampling point. First, the chemical composition value q_{mn} is standardized. The conversion formula is as follows:

$$t_{mn} = \frac{q_{mn} - \mu_n}{s_n}, m = 1, 2, \cdots, i; n = 1, 2, \cdots, j \tag{6}$$

The standardized indicator variables are:

$$u_n = \frac{z_n - \mu_n}{s_n}, n = 1, 2, 3, \cdots, j \tag{7}$$

Calculate the distance between two sample points and construct Euclid distance matrix:

$$dist_{mk} = \sqrt{\sum_{n=1}^{j} (t_{mn} - t_{kn})^2}, m, k = 1, 2 \cdots i \tag{8}$$

The shortest distance method is used to measure the distance between sampling points of different cultural relics, namely:

$$D(G_p, G_q) = \min_{m \in G_p, k \in G_q} \{dist_{mk}\} \tag{9}$$

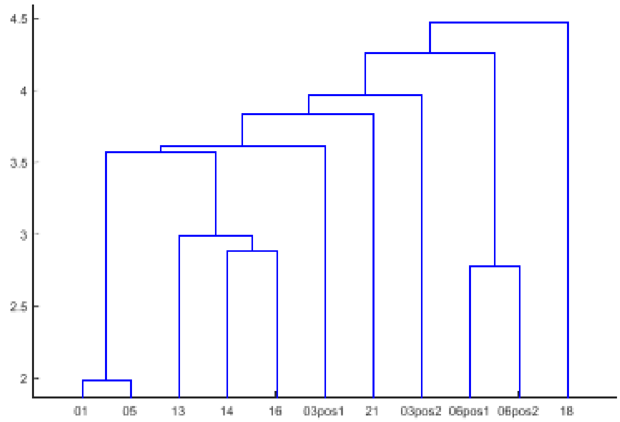
Construct *i* classes, each class contains only one sample point and its platform height is zero

The two classes with the least distance are merged into a new class, and the distance between the new classes is used as the platform height of the cluster graph.

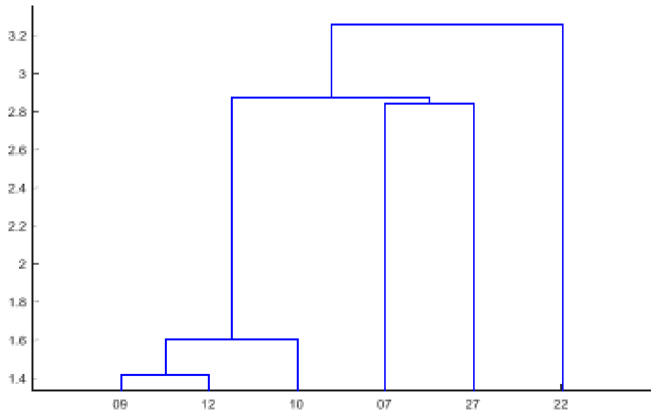
If the number of classes is 1, the cluster graph will be drawn; otherwise, the distance will be recalculated.

3.3. Solution of model

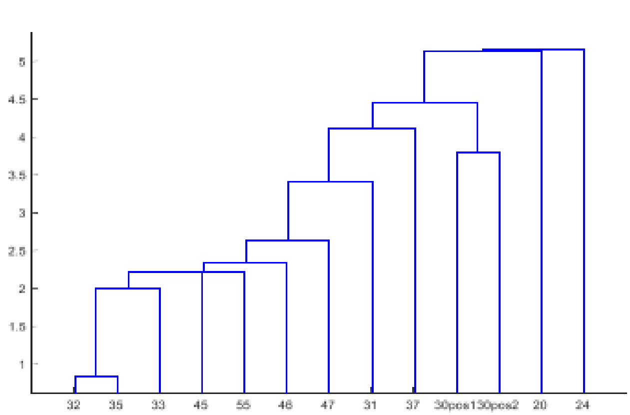
After data processing, use MATLAB software to perform R-type clustering, and the clustering results of each category are shown in Fig 4:



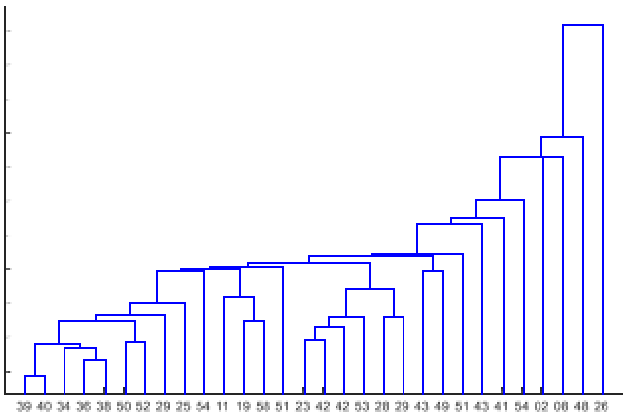
Cluster diagram of high potassium non weathering



Cluster Diagram of High Potassium Weathering



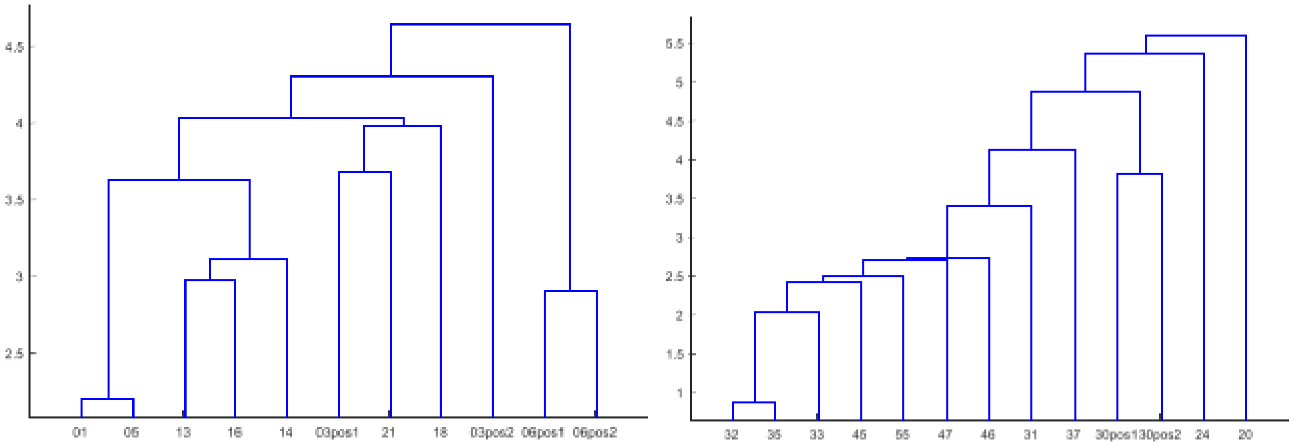
Lead barium non weathering cluster diagram



Lead barium weathering cluster diagram

Fig. 4 R-type cluster diagram of each category

The sensitivity test and rationality analysis are carried out in this paper. According to literature query, the higher the aluminum content is, the deeper the ion exchange layer of the glass is, the greater the surface stress is, and the better the reinforcement performance is. Therefore, the weathering resistance of the glass is also higher [10]. The topic requires us to analyze the rationality and sensitivity of the classification results. Based on the above literature, this paper conducts sensitivity analysis on the glass of the fresh category by adjusting whether the aluminous chemical composition exists. Data processing shall be carried out for the high potassium weathering group and the lead barium weathering group to remove the chemical composition of aluminum in these two groups,Then, the R-type clustering model is used to solve the problem. The operation results of MATLAB software are shown in Fig 5:



High potassium weathering group Lead barium weathering group

Fig. 5 R-type clustering solution diagram of unweathered group

Compared with the above clustering diagram of high potassium and lead barium weathering groups, it is found that the classification results of high potassium weathering groups have greatly changed, while the clustering results of lead barium weathering groups are not significantly different, which indicates that aluminum plays a greater role in weathering resistance of high potassium glass cultural relics, while it plays a smaller role in weathering resistance of lead barium glass cultural relics.

4. Conclusion

In order to verify the influence of the chemical composition aluminum on the classification results, this paper conducts sensitivity analysis by adjusting the existence of aluminum containing chemical composition for the glass of the fresh category. The data of high potassium and lead barium weathering groups were processed, and the chemical composition of aluminum in these two groups was removed. Then the R-type clustering model was used to solve the problem and the conclusion was drawn:It is found that the classification results of the high potassium weathering groups have greatly changed, while the clustering results of the lead barium weathering groups have no significant difference, which indicates that aluminum plays a greater role in weathering resistance in the high potassium type glass relics, while it plays a smaller role in weathering in the lead barium type glass relics.

Reference

[1] Diao Yunchao, Jiang Yanyan, Zhang Qingyu, Wang Chengyu Study on weathering property of coated float glass [C] Summary of the 2011 Global Annual Conference on Glass Science and Technology

[2] Hao Yunhong, Wu Rigen, Zhao Chengguang, Guo Xin, Ya Ruhan. Study on the impact of weathering on the erosion resistance of tempered glass based on the Taguchi method [J]. Silicate Bulletin, 2020, 39 (09)

[3] Gan Fuxi, Zhao Hongxia, Li Qinghui, Li Ling, Cheng Huansheng. Scientific and technological analysis and research on glass products unearthed in the Warring States Period in Hubei Province [J]. Jianghan Archaeology, 2010 (02)

[4] Zou Shifeng, Jiang Hong, Zhao Huifeng, Li Chunhua, Xia Wenbao Effect of Al2O3 content in glass on ion exchange reinforcement [J] Journal of Materials Science and Engineering, 2014.01.008

[5] Qin Qiusheng Construction of prediction model for the number of college students based on multiple linear regression analysis [J] Wind of Science and Technology, 2022 (22)

[6] Yang Guiyuan. Mathematical Modeling [M]. Shanghai: Shanghai University of Finance and Economics, 2015.157-181

[7] Yang Guiyuan, Zhu Jiaming Evaluation and Analysis of Excellent Papers in Mathematical Modeling Contest [M] Anhui: China University of Science and Technology Press, 2013.1-12

[8] Wang Junxin, Li Ping, Zhang Xun, Peng Zicheng, Chen Shuyu, Huang Yunlan, Jiang Tingyu, Qiu Zhonglun. Study on the lead isotope tracing of Xihan ancient glass in Heputang Pai, Guangxi [J]. Nuclear Technology, 1994 (08)

[9] Gong Lixiong, Liu Shixiong, Wang Canlin Research on prediction model of coal consumption based on BP neural network [J]. Journal of Hubei University of Engineering, 2017

[10] Liu Yong, Yang Jun, Chen Kunlong, Chen Study on the Manufacturing Technology and Related Problems of M5 Glass Mat Unearthed in Haihun Marquis Cemetery of the Western Han Dynasty in Nanchang [J]. Southern Cultural Relics, 2021 (06)