Prediction of chemical composition of cultural relics glass based on moving average algorithm of least square method

Yuanjin Wen *

Department of physics, School of Physics and Electronic Engineering, Chongqing Normal University Chongqing, China, 401331

* Corresponding author: chuichuier1022@163.com

Abstract. In order to explore the chemical composition of ancient cultural relics glass and understand the structure of ancient cultural relics glass, according to the basic information of glass cultural relics, the characteristics of glass are quantified, and the chi-square independence test is carried out on whether weathering and the properties of glass. Based on the cross table and chi-square test, it can be concluded that high potassium glass is generally not easy to weathering, while lead barium glass is more prone to weathering. A total of four categories were established to effectively reflect the statistical law of the chemical composition content of glass with different properties, and the principal component factors of each category were explored to obtain the chemical composition of each category. The scatter line diagram of the chemical composition of the unweathered sampling points was drawn. Based on the least square method, the ideal trend value of each chemical composition was calculated by using the moving average algorithm, and the prediction of various chemical compositions was completed. The chemical composition of the glass can be successfully predicted by the model constructed in this study. This method has certain technical support and theoretical support for the study of ancient glass, and provides some reference.

Keywords: Glass relics, Chi-square independence test, Chemical composition prediction, the least squares method, Moving average algorithm.

1. Introduction

The production process of ancient glass has developed well in China, and has also achieved good development in the world. The two river basins in the 20th century BC may be the first place to find ancient glass. From the 10th to the 12th century, the method of making glass changed greatly and began to develop in different directions. Due to the different technical routes and social influences of Chinese and foreign glass making, glass products with different characteristics and functional attributes are now produced [1]. Understanding the production techniques of ancient cultural relics glass can reflect the cultural self-confidence of a country, and can also provide some theoretical basis for the world to study ancient cultural relics glass, so that we can more deeply understand the unique charm of five thousand years of Chinese culture.

The main raw material of glass is quartz sand, and the main chemical composition is silicon dioxide. At present, most of the methods of refining glass are to reduce the melting point of pure quartz sand by adding cosolvents. Different cosolvents will lead to different chemical composition of glass. Lead barium glass and potassium glass are distinguished because of the different cosolvents. At the same time, there are many factors that affect the judgment of ancient glass types, one of which is the weathering of glass. Ancient glass is easily affected by the environment and weathering. In the whole process of weathering, the internal elements of the glass will exchange a lot with the environmental elements, resulting in the change of the composition ratio of the glass, which will affect the judgment. Unweathered glass relics can clearly see its color and decoration, but there will also be local weathering, and the surface color of the weathered relics will change.

This study will analyze the relationship between the surface weathering of these glass relics and their glass types, patterns and colors. Combined with the type of glass, the statistical law of the content of weathering chemical components on the surface of cultural relics samples is analyzed. According to the detection data of weathering points, the content of chemical components before weathering is predicted, so as to analyze the chemical composition of glass cultural relics more accurately.

2. Data Acquisition and Preprocessing

2.1. Data acquisition and preprocessing

All the data in this study come from the C problem of the national college students ' mathematical modeling competition in 2022. (http://www.mcm.edu.cn/html_cn/node/5267fe3e6a512bec793d71f2b2061497.html).

Because there are some invalid data in the data given by Problem C of the National College Students ' Mathematical Contest in Modeling in 2022, the data need to be preprocessed before the model is established. By comparing the obtained data with the data needed in this study, it is found that there are two sampling points in the data of chemical composition content of each sampling point provided in Form 2 of the given data. The proportion of components is accumulated and less than 85 %, and this kind of obvious invalid data is eliminated and not included in subsequent research and analysis.

2.2. methods to introduce

2.2.1. chi-square test

As a widely used hypothesis testing method for frequency data, chi-square independence test is often used to analyze the correlation degree of two categorical variables. The basic idea is to compare the coincidence degree between theoretical frequency and actual frequency. The principle of the chi-square test model is specifically, assuming that the two categorical variables$x$and$y$, respectively, $x = 1,2,3,\ldots,m, y = 1,2,3,\ldots,n$ from which x and y can be obtained to obtain a contingency table [2], as shown in Table 1.

Table 1. Chi-square independence test and contingency table

categorical variables	$y = 1$	...	$y = s$	footing
$x = 1$	r_{11}	...	r_{1n}	a_1
...	...	...	...	...
$x = m$	r_{m1}	...	r_{mn}	a_m
footing	b_1	...	b_n	r

Where r_{ij} is the actual frequency of the survey results of choices $x = i$ and $y = j$, and each frequency 4 corresponds to an expected frequency e_{ij} defined as follows:

$$e_{ij} = \frac{a_i b_j}{r} \quad i = 1,2,\ldots,m, j = 1,2,\ldots,n \tag{1}$$

The expected frequency can be understood as strictly according to the survey object attribute frequency. The degree of agreement between the actual frequency and the expected frequency can be determined by formula (2).

$$\chi^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(r_{ij}-e_{ij})^2}{e_{ij}} \tag{2}$$

Statistics χ^2 is proposed by British statistician Pearson, so χ^2 is usually called Pearson χ^2. And it is easy to observe, if χ^2 is very small, it means that the actual frequency is basically consistent with the attribute frequency of the survey object, reflecting that the interaction between variables x and y is relatively small; on the contrary, if χ^2 is large, it indicates that the two have a great influence on each other, which can prove that the two have a great correlation.

2.2.2. least-square method

The least squares method is a mathematical tool widely used in many disciplines such as error estimation, uncertainty, system identification and prediction, forecasting and other data processing.

In this study, the chemical composition content of the weathered cultural relic glass before weathering is predicted. Considering that the cultural relic glass, whether it is weathered or not, is in the same environment in its construction age, the same era, the same production time, so the chemical composition content before weathering should be the same as that of the unweathered chemical composition content. When considering the chemical composition content before weathering, the different types of glass will also make the chemical composition content different. Therefore, the least square method can be used to predict the two different types of high potassium and lead barium glass.

2.2.3. moving average algorithm

This study improves on the basis of the original least squares method and adds a moving average algorithm. The moving average method is a common method to predict the demand of the company's products and the company's production capacity in the next period or several periods with a set of recent actual data values. The moving average method is suitable for immediate prediction. When the product demand is neither rapid growth nor rapid decline, and there is no seasonal factor, the moving average method can effectively eliminate the random fluctuations in the forecast, which is very useful. The moving average method can be divided into simple moving average and weighted moving average according to the weight of each element used in the prediction.

3. Establishment and Solution of Model

3.1. Statistical analysis of the content of weathering chemical components on the surface of cultural relics samples based on glass types

A total of 69 samples were collected from 58 heritage glass samples. It is necessary to analyze the statistical law of the chemical composition content of whether the surface of the cultural relics is weathered based on the type of glass. The basis for the classification of the types of cultural relics glass is based on the content of chemical elements. For the definition of weathering of cultural relics, there is still a vague definition of weathering after eliminating invalid sampling points.

Therefore, in the preliminary analysis of the statistical law of the type of glass on the chemical composition content of whether the surface of the cultural relics is weathered, the unweathered point of the weathered glass is removed first, so there are only 59 effective sampling points, among which the weathered glass will only take the weathered sample point, so the sample point can be used as an effective sampling point to reflect the weathering of the glass ; among them, only unweathered samples are taken for unweathered glass, so this sample can be used as an effective sampling point to reflect unweathered glass [3].According to the type of cultural relics glass and whether the glass is weathered or not, it is divided into four categories : high potassium weathering, high potassium unweathered, lead barium weathering, and lead barium unweathered for factor analysis.

After the preliminary analysis of the four types of factors, the chemical composition content of the weathered and unweathered parts on the surface of the weathered glass is further considered. Weathered and unweathered samples were taken on the differentiated glass. Based on the fact that the glass has been weathered, there are 3 weathering sampling points and 10 unweathered sampling points.

The counts of weathered and unweathered cultural relics glass in different types, different colors and different patterns are counted respectively. The specific statistical results are shown in Table 2.

Table 2. Comparison of power load forecasting of 403 line

Type attribute	Attribute details	Weathering	Unweathered
Type	High potassium	6	12
	Lead barium	28	12
Decoration	A	11	11
	B	6	0
	C	17	13
Color	Blue green	9	6
	Light blue	12	8
	Purple	2	2
	Deep green	4	3
	Black	2	0
	Green	0	1
	Deep blue	0	2
	Light green	1	2
	Unknown	4	0

From the data in the counting of weathered and unweathered cultural relics glass under different types of attributes in Table 2, it can be seen that the frequency law of weathering is not the same under different types of different attributes. However, only from the number of frequencies, it is not possible to directly obtain the specific correlation analysis between the surface weathering of glass and glass type, decoration and color. From a statistical point of view, the difference in quantity cannot reflect the size of the correlation degree. For example, at the level of glass decorated as A, there are 11 weathered glass and 11 unweathered glasses. From the perspective of the glass level decorated with B, a total of 6 cultural relics glass are all weathered. Obviously, it can be seen that the latter decorated with B is related to the research problem. But it is difficult to be convincing only from more and less levels. Therefore, this paper needs to further analyze the method of independence test. Independence test is a statistical method to test whether the two categorical variables are related (or independent). The most commonly used method is chi-square independence test [4].

3.2. Preliminary analysis of the statistical law of the chemical composition content of whether the glass surface is weathered or not

3.2.1. Factor analysis of the chemical composition of high potassium weathering glass was carried out by SPSS.

A. **Data preprocessing**

Due to the different dimensions of different indicators, there is no comparability. In this paper, before the formal factor analysis of the chemical composition content of high potassium weathering glass, it is necessary to standardize the original data of the chemical composition content of the weathering sampling points on the surface of the high potassium weathering glass, as shown in Table 3, so as to eliminate the influence of dimension [5].

The standardized formula is as follows:

$$Z_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{3}$$

Where X_i is the original data and Z_i is the standard data.

Table 3. Comparison of power load forecasting of 403 line

	N	Minimum value	Maximum value	Mean value	Standard deviation
SiO2	6	92.35	96.77	93.9633	1.73362
K2O	6	0	1	.54	.445
CaO	6	.21	1.66	.8700	.4877
MgO	6	0	1	.20	.306
Al2O3	6	.81	3.50	1.8217	1.05965
Fe2O3	6	.17	.35	.2650	.0695
CuO	6	.55	3.24	1.5617	.934820
P2O5	6	.00	.61	.2800	.20995
Effective N	6				

B. **preprocessing**

Using SPSS, the eigenvalue and variance contribution rate of each factor are obtained. The SPSS results show that the first two common factors explain 88.466 % of the total variance, indicating that the extracted two common factors can represent 88.466 % of the original eight chemical composition content indicators that measure whether weathering, indicating that the data information loss is less, can better explain the initial data, so the extracted two common factors are X_1, X_2.

C. **factor loading**

The maximum variance method is used for factor rotation. The common factor 1 has a large load on Al2O3, CaO and MgO, which can be classified into one category and defined as a factor of high potassium weathering. The common factor 2 has a large load on potassium oxide (K2O) and iron oxide (Fe2O3), which can be classified as a class of high-potassium weathering b molecules.

(1) Statistical analysis of chemical composition content of high potassium weathering glass

According to the factor analysis results of the statistical law of chemical composition content of high-potassium weathering glass, it can be concluded that for high-potassium glass, the weathering state of the glass can finally be mainly determined by two common factors. The contribution rate of factor 1 to the weathering state of high-potassium glass is as high as 51.645 %, while the contribution rate of factor 2 to the weathering state of high-potassium glass is 36.821 %. Therefore, the contribution rate of factor 1 and factor 2 to the weathering of high-potassium glass is 88.466 %.

For the composition of factor 1 and factor 2, factor 1 is mainly composed of Al2O3, CaO and MgO, while factor 2 is mainly composed of K2O and Fe2O3. In the contribution analysis of the two factors, it can be seen that the contribution of factor 1 is larger than that of factor 2, so the contribution of chemical composition can be further analyzed roughly. The contribution to Al2O3 is 0.982 × 0.51645 = 0.5071539; the contribution of CaO is 0.942 × 0.51645 = 0.4864959; the contribution of MgO is 0.906 × 0.51645 = 0.4679037; the contribution of K2O is 0.857 × 0.36821 = 0.31555597; the contribution of Fe2O3 is 0.850 × 0.36821 = 0.3129785.

Therefore, it can be roughly estimated that the content of Al2O3, CaO, MgO, K2O and Fe2O3 from large to small affects the weathering state of high potassium glass.

3.2.2. Factor analysis of the chemical composition of high potassium weathering glass was carried out by SPSS.

Similarly, the appeal method was used to analyze the chemical composition content of high-potassium unweathered glass. The first four common factors explained 81.606 % of the total variance, indicating that the extracted four common factors can represent the original 12 chemical components. The content index of 81.606 % indicates that the data information is less lost and can better explain the initial data, so the four common factors are extracted as $\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3, \mathrm{Y}_4$.

The common factor 1 has a large load on alumina (Al2O3) and phosphorus pentoxide (P2O5), which can be classified as a class, defined as high potassium unweathered a factor; common factor 2 has a large load on calcium oxide (CaO), which can be classified as a class, defined as high potassium unweathered b molecule ; the common factor 3 has a large load on lead oxide (PbO) and can be classified as a class, which is defined as a high potassium unweathered c molecule ; common factor 4 had no significant load on each chemical component.

According to the factor analysis results of the statistical law of chemical composition content of high potassium unweathered glass, it can be concluded that for high potassium glass, the unweathered state of glass can be mainly determined by four common factors. The contribution rate of factor 1 to the weathering state of high potassium glass is 33.256 %, while the contribution rate of factor 2 to the unweathered state of high potassium glass is 21.453 %. The contribution rate of factor 3 to the unweathered state of high potassium glass is 16.016 %, and the contribution rate of factor 4 to the unweathered state of high potassium glass is 10.881 %. Therefore, the four components of factor 1, factor 2, factor 3 and factor 4 contribute 81.606 % to the unweathered determination of high-potassium glass.

For the composition of factor 1, factor 2, factor 3 and factor 4, factor 1 is mainly composed of Al2O3 and P2O5, while factor 2 is mainly composed of CaO and factor 3 is mainly composed of PbO.

Therefore, it can be roughly estimated that the content of Al2O3, P2O5, CaO and PbO from large to small has an effect on the unweathered state of high potassium glass.

3.2.3. Factor analysis of the chemical composition of lead-barium weathered glass by SPSS.

Similarly, it can be roughly estimated that the content of BaO, BaO, SnO2, Al2O3, SrO and K2O from large to small affects the weathering state of lead-barium glass.

3.2.4. Factor analysis of the chemical composition of lead-barium weathered glass by SPSS.

According to the factor analysis table of the chemical composition of lead-barium unweathered glass, the first five common factors explain 83.153 % of the total variance, indicating that the extracted five common factors can represent 83.153 % of the original 14 chemical composition content indicators that measure whether the glass is weathered, indicating that the loss of data information is less, which can better explain the initial data, so the five common factors are extracted as S_1, S_2, S_3, S_4, S_5 [6].

Common factor 1 has a large load on CaO, SnO2 and PbO, which can be classified as a class, defined as lead-barium unweathered factor a ; common factor 2 has a large load on BaO and CuO, which can be classified as a class, defined as lead-barium unweathered b molecule ; the common factor 3 has a large load on Al2O3 and can be classified as a class, which is defined as lead barium weathering c molecule ; common factor 4 had no significant load on each chemical component ; the common factor 5 has a large load on SO2 and can be classified as a class, defined as lead-barium weathering e molecule.

It can be roughly estimated that the content of CaO, SnO2, PbO, BaO, CuO, Al2O3 and SO2 has an effect on the unweathered state of lead-barium glass.

3.2.5. Factor analysis of the chemical composition of the weathered part of the weathered glass was carried out by SPSS.

Similarly, according to the relevant table of factor analysis of the chemical composition content law of the weathered part of the weathered glass, the previous common factor can explain 92.874 % of the total variance, indicating that the extracted common factor can represent the original 11. The chemical composition content index of weathering is 92.874 %, indicating that the loss of data information is less, which can better explain the initial data, so the common factor is extracted.

The common factor has a large load on lead oxide (PbO), silicon dioxide (SiO2), phosphorus pentoxide (P2O5), magnesium oxide (MgO), alumina (Al2O3) and strontium oxide (SrO), which can be classified into one category and defined as the weathering part a factor.

According to the factor analysis results of the statistical law of the chemical composition content of the weathered part of the glass, it can be concluded that the weathering state of the glass can finally be determined mainly according to a common factor. From the descriptive statistical scale, it can be seen that the contribution rate of the factor to the weathering state of the lead glass is 92.874 %, so the contribution rate of the factor component to the weathering of the glass is 92.874 %.

For the composition of the factor, factor 1 is mainly composed of lead oxide (PbO), silica (SiO2), phosphorus pentoxide (P2O5), magnesium oxide (MgO), alumina (Al2O3), strontium oxide (SrO).

Therefore, it can be roughly estimated that the content of lead oxide, silicon dioxide, phosphorus pentoxide, magnesium oxide, alumina and strontium oxide from large to small affects the weathering state of glass.

3.2.6. Factor analysis of the chemical composition of the unweathered part of the weathered glass was carried out by SPSS.

According to the relevant tables of factor analysis of the chemical composition content of the unweathered part of the weathered glass, the first five common factors explain 91.176 % of the total variance, indicating that the extracted five common factors can represent the original 14. The chemical composition content index of weathering is 91.176 %, indicating that the loss of data information is less, which can better explain the initial data, so the five common factors are extracted as 1.

Common factor 1 had no significant load on each chemical component; common factor 2 had no significant load on each chemical component; common factor 3 had no significant load on each chemical component; common factor 4 had no significant load on each chemical component; common factor 5 had no significant load on each chemical component [7].

Therefore, it can be concluded that there is no specific law for the chemical composition content of the unweathered state of the weathered glass.

3.3. Prediction of moving average algorithm based on least square method

For the weathered cultural relic glass, the chemical composition content before weathering is predicted. Considering that the cultural relic glass is in the same environment in the same age, the same age and the same production time, the chemical composition content before weathering should be the same as that before weathering. When considering the chemical composition content before weathering, the different types of glass will also make the chemical composition content different. Therefore, two different types of glass, high potassium and lead barium, should be predicted separately [8].

3.3.1. Prediction of chemical composition content of high potassium glass before weathering

Using Python program software, the broken line scatter plot of the chemical composition content of each unweathered sample is drawn. As shown in Figure 1, it can be clearly observed that the chemical composition content of unweathered cultural relic glass has certain rules and trends [9,10].

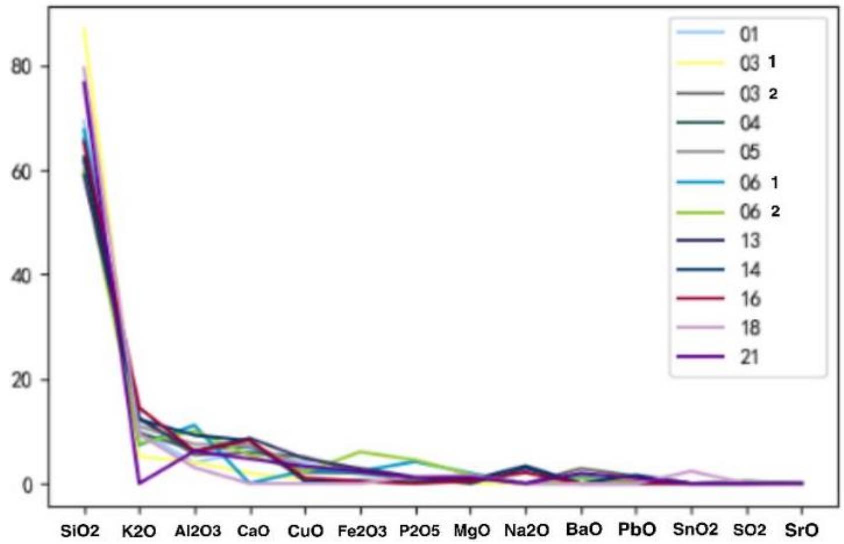

Figure 1. The broken line scatter plot of chemical composition content of unweathered samples

The least square method was used to predict the trend of chemical composition content of high potassium glass without weathering. Firstly, 12 groups of case data of chemical composition content of unweathered high-potassium glass were sorted out. For the chemical composition content group of each sampling point case, it was composed of 14 chemical composition contents. The chemical composition content of each sampling point was drawn. The scatter line chart of each chemical composition content is shown in Figure 1 above.

For any chemical composition content $i(i=1,2,3,\ldots,13)$, the actual content of the sample is set to y_i, and the ideal trend content is x.

If satisfied:

$$\min\left(\sum_{i=1}^{n}(x-y_i)^2\right) \tag{4}$$

The trend content of the chemical composition is x.

The ideal chemical composition content of 14 chemical components of 13 groups of case data in the case data of the chemical composition content of high-potassium glass without weathering was predicted. The chemical composition content of high-potassium glass before weathering was predicted as follows Table 4.

Table 4. Comparison of power load forecasting of 403 line

Unweathered high potassium glass	SiO_2	K_2O	Al_2O_3	CaO	CuO	Fe_2O_3	P_2O_5
	67.9842	9.33083	6.62	5.3325	2.4525	1.93167	1.4025
	MgO	Na_2O	BaO	PbO	SnO_2	SO_2	SrO
	1.07917	0.695	0.59833	0.41167	0.19667	0.10167	0.04167

3.3.2. Prediction of chemical composition content of lead-barium glass before weathering

The trend of chemical composition content of unweathered lead-barium glass was predicted by least square method. Firstly, 13 groups of case data of chemical composition content of lead-barium glass without weathering were sorted out. For the chemical composition content group of each sampling point case, it was composed of the content of 14 chemical components. The chemical composition content of each sampling point was drawn. The scatter line diagram of the chemical composition content is shown in Figure 2 below.

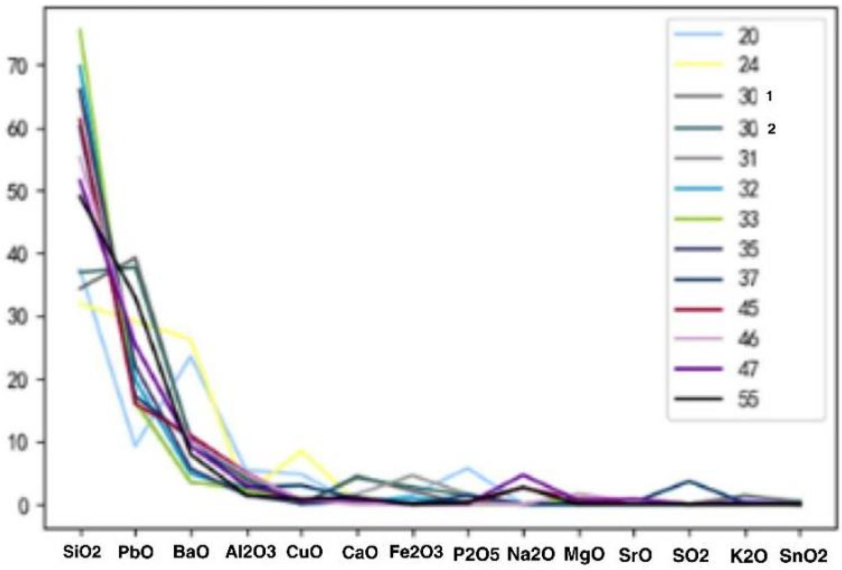

Figure 2. The chemical composition scatter plot of unweathered lead-barium glass

The ideal chemical composition content of 14 chemical components in 13 groups of case data of chemical composition content of unweathered lead-barium glass was predicted. The prediction of chemical composition content of lead-barium glass before weathering is shown in Table 5.

Table 5. Comparison of power load forecasting of 403 line

Unweathered lead barium glass	SiO_2	K_2O	Al_2O_3	CaO	CuO	Fe_2O_3	P_2O_5
	53.4439	23.5939	10.4992	3.19462	1.55692	1.23154	0.93308
	MgO	Na_2O	BaO	PbO	SnO_2	SO_2	SrO
	0.90385	0.77154	0.49231	0.29692	0.28154	0.25846	0.06462

4. Conclusion

In this paper, the basic information and chemical composition ratio of two kinds of glass samples, as well as the chemical composition ratio of classified glass relics are classified by using the relevant data of a batch of ancient glass samples, and the correlation between the influence indexes of classification is studied. In summary, the moving average algorithm based on the least squares method is a more accurate method for predicting cultural relics glass. It is predicted that the chemical composition of high potassium glass before weathering is mainly SiO_2, which is as high as 67.9842 %, followed by K_2O accounting for 9.33083, while Na_2O, BaO, PbO, SnO_2, SO_2 and SrO are all less than 1 %. Prediction of the trend of chemical composition content of lead-barium glass without weathering the chemical composition content is mainly SiO_2, P_2O_5, P_2O_5, Na_2O, BaO, PbO, SnO_2, SO_2 and SrO are also less than 1 %. Different from high-potassium glass, the content of K_2O in lead-barium glass is as high as 23.5939 %, and the proportion of Al_2O_3 is 10.4992.

References

[1] Li Yingli. Research on glassware unearthed in China from the 10 th to the 12 th century [D]. Dalian: Liaoning Normal University, 2021.

[2] Jin Danxing, the application of chi-square test in the analysis of college students ' English final scores and learning styles, Overseas English, 2019.

[3] Jia Chunyan, the application of chi-square test in the analysis of second-child fertility willingness-Taking the faculty of Beijing Polytechnic Institute as an example, Journal of Beijing Polytechnic Institute, 2016.

[4] Li Zipeng, Wang Decai, Zhang Yu, Tan Guohui, Sun Hua, Shu Shu, Mei Shuqi, Peng Xianwen. Principal Component Analysis and Cluster Analysis of Slaughter and Body Size Traits in Selenium Black Pigs [J]. Chin J Animal Husbandry, 2022, 58 (08): 56 - 60.DOI: 10.19556 / j.0258 - 7033.20220615 - 17.

[5] Li Yang, SPSS, Journal of Chifeng University (Natural Science Edition), 2011.

[6] Duan Baona; feng Sanying, difference analysis of organic acid content in jujube based on principal component analysis, [J]. China condiment, 2021, 149 - 151, 149 - 151.

[7] Zhang Shuaifan, Sun Ruijuan, Ren Hanzhuo. Chemical composition analysis and correlation study of ancient glass products.

[8] Lu Jiajia. Ancient glass classification model based on integrated feature selection and random forest [J]. Silicate Journal, 2023,51 (04): 1060 - 1065.DOI: 10.14062 / j. issn.0454-5648.20220790.

[9] Yin Yulong. Composition analysis of ancient glass products by correlation prediction [J]. Contemporary chemical research, 2023 (01): 122 - 126.

[10] Li Lingling. Analysis and identification of ancient glass products based on statistical analysis methods [J]. Chinese new technology and new products 2023 (02): 99 - 101.DOI: 10.13612 / j. cnki. cntp.2023.02.014.