ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

Research on Composition Analysis and Identification Based on Ancient Glass Products

Qing Li, Chuang Wang, Jianwei Zhai

HLJU, Software Faculty, Harbin, Heilongjiang, 150080, China

Abstract: *Ancient glass products are very vulnerable to weathering caused by burial environment. In the process of weathering, a series of changes will often occur in the proportion of chemical components inside. The chemical composition of these glass relics was analyzed and identified.This paper makes chi square test and Spearman correlation analysis on the surface weathering of glass cultural relics and other attribute values respectively. It is found that the surface weathering of cultural relics is closely related to the type of cultural relics and has significant differences. The quantitative relationship between the weathering of cultural relics and other attribute values is fitted through multiple regression in three ways. Finally, pure quadratic multiple regression is selected, and the fitting accuracy reaches 75.9259%. Then, the data are normalized, and the rule of whether there is weathering chemical composition content in cultural relics samples is summarized through the relationship diagram between the content of each chemical component and the type of chemical component in each sample. Then, the influence of each component on the weathering results of the samples is obtained by using the principal component analysis method to verify the rationality of the conclusion. It is found that the content of phosphorus pentoxide and copper oxide has a significant impact on the weathering results of cultural relics. Finally, given the initial values of the variables in the weathering samples that have a low impact on the weathering degree, the chemical composition content of the samples before weathering can be predicted by multiple linear regression.In this paper, firstly, based on the two types of high potassium glass and lead barium glass, the average and variance of the chemical composition content in each sample were counted, and the classification rules of the two types of glass were obtained by analyzing the difference of the chemical composition of the two types of glass. Then, using the Kmeans clustering method, the high potassium glass and lead barium glass were clustered, and the principal component analysis method was used to reduce the dimensions of the data, Combined with each principal component and the clustering process, the sub classification standard was determined. Finally, the sensitivity of the classification results was analyzed by PSI method, and the rationality of the classification results was analyzed by sampling inspection method. The accuracy of the classification results was 89.036%.*

Keywords: *chisquare test, Spearman test, multiple regression fitting, Kmeans clustering*

1. Research background

Glass, as an indispensable material in today's daily life, is an important material in foreign trade. It not only reflects the wisdom of the ancient working people, but also demonstrates the ancient craft level of China [1][2]. However, because the chemical composition is affected by factors such as flux and weathering, it is difficult to identify its type by the content of chemical composition on the surface of cultural relics. Therefore, the analysis of chemical composition and classification information of glass cultural relics has important research value for exploring ancient glass. This paper studies the composition and classification of glass relics.

2. Study on chemical composition of glass

This paper studies the relationship between the weathering of samples and the color, texture and type of samples, and summarizes the statistical law of chemical composition of each sample in combination with the sample type. Finally, it is necessary to predict the chemical composition content of weathered samples before weathering. First, preprocess the sample data, quantize the attribute values in Form 1, and convert them into numerical categorical variables. Secondly, analyze the chemical composition content of each sample in Form II, and eliminate the abnormal samples that do not meet the meaning of

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

the question. At the same time, in combination with the analysis results of chemical composition content, the missing variables are supplemented. Here, "0" is used to complete the missing values. The chi square test[3] and Spearman test[4] were performed on the weathering of the sample and other attributes of the sample in Table 1 to obtain the significant difference and correlation between the categorical variables, and the quantitative relationship between the weathering of the sample and the values of other attributes of the sample was fitted by means of multiple regression[5] and accuracy analysis was performed. According to Form 2, the data involved shall be normalized first, and the relationship between the content of each chemical component and the type of chemical component of each sample shall be drawn to summarize the content laws of chemical components of high potassium and high barium samples before and after weathering. Finally, the influence of various components on the weathering results of samples is obtained by principal component analysis[6] to verify the rationality of the conclusions. According to the above summarized chemical composition content rules of samples, the initial values of variables that affect weathering degree in weathering samples are given, and the chemical composition content of samples before weathering is predicted by multiple linear regression.

Analyze the classification rules of high potassium glass and lead barium glass according to the known information, and sub classify them according to the difference of chemical composition content based on each category. Finally, analyze the rationality and sensitivity of the classification results. The interval of each chemical composition content of high potassium glass and lead barium glass is counted respectively, and based on each category, the average value and variance of the chemical composition content in each sample are counted, and the statistical results are visualized to analyze the difference of the chemical composition of the two types of samples, so as to obtain the classification rules of high potassium glass and lead barium glass. By using the Kmeans clustering method, the cluster analysis is carried out on the samples with high potassium and lead barium respectively by iterating the number of clusters gradually, and the relatively ideal number of clusters is selected. At the same time, in order to make the results visible, the principal component[7] analysis method is used to reduce the dimensions of the data, and each principal component and the clustering process are combined to determine the sub classification standard. In this paper, PSI method is used to analyze the sensitivity of the classification results, and sampling inspection method is used to analyze the rationality of the classification results.

3. Establishment and solution of research model

3.1 Data preprocessing

In order to intuitively obtain the relationship between sample weathering and sample color, type, and texture, each nominal class variable is assigned to quantify, so that it can be transformed into a numerical class variable. For the data in Form 2, fill in the missing values with reference to assumptions. At the same time, calculate the cumulative sum of the proportion of chemical substance content in each sample, and remove the abnormal samples whose results are not in the range of 85%~105%. Finally, in order to facilitate the subsequent calculation and analysis, the data in Form 2 is normalized.

3.2 Chi square test and spearman correlation analysis

In order to analyze the relationship between categorical variable data, according to the topic, SPSS was used to conduct chi square test on surface weathering and glass type, surface weathering and grain, surface weathering and color, and the results are shown in the following table. In order to analyze the relationship between categorical variable data, according to the topic, SPSS was used to conduct chi square test on surface weathering and glass type, surface weathering and grain, surface weathering and color, and the results are shown in the following table 1~3.

Table 1: Chi square test for relationship between type and surface weathering

	value	freedom	Progressive significance (bilateral)
Pearson chi square	6.287[a]	7	0.507
likelihood ratio	8.156	7	0.319
Linear correlation	0.707	1	0.401
Number of valid cases	54		

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

Table 2: Chi square test for relationship between color and surface weathering

	value	freedom	Progressive significance (bilateral)
Pearson chi square	6.287a	7	0.507
likelihood ratio	8.156	7	0.319
Linear correlation	0.707	1	0.401
Number of valid cases	54		

Table 3: Chi square test for relationship between grain and surface weathering

	value	freedom	Progressive significance (bilateral)
Pearson chi square	5.747a	2	0.056
likelihood ratio	7.993	2	0.018
Linear correlation	0.205	1	0.650
Number of valid cases	54		

The chi square test showed that the analysis of surface weathering and sample type passed the significance test, and there was a significant difference between the two. There is also a significant difference between the surface weathering and the sample texture, but the difference does not pass the significance test. The asymptotic significance between surface weathering and sample color is too large, which can be understood that there is no direct relationship between them.

Next, analyze the chi square test of the relationship between sample color, sample type and sample texture, and the test results are shown in the following table 4~5.

Table 4: Chi square test for the relationship between sample color and sample type

	value	freedom	Progressive significance (bilateral)
Pearson chi square	22.693a	7	.002
likelihood ratio	25.201	7	.001
Linear correlation	9.507	1	.002
Number of valid cases	54		

Table 5: Chi square test for the relationship between sample color and sample grain

	value	freedom	Progressive significance (bilateral)
Pearson chi square	38.109a	14	.001
likelihood ratio	43.503	14	.000
Linear correlation	2.001	1	.157
Number of valid cases	54		

The chi square test shows that there are significant differences between sample color and sample type. There are also significant differences between sample color and sample grain.

To sum up, the surface weathering and color of the sample have a significant relationship with the sample texture and type, while the surface weathering of the sample has no significant relationship with the sample color.

Next, SPSS is used to analyze the relationship between sample surface weathering and sample color, sample texture and sample type. The results are shown in the following table 6~8.

Table 6: Chi square test for relationship between sample surface weathering and sample type

		value	Asymptotic standard error a	Approximate Tb	Progressive significance
Interval to interval	Pearson R	.316	.130	2.404	.020c
Order to order	Spearman correlation	.316	.130	2.404	.020c
Number of valid cases		54			

Table 7: Chi square test for relationship between sample surface weathering and sample color

		value	Asymptotic standard error a	Approximate Tb	Progressive significance
Interval to interval	Pearson R	-115	0.136	-838	0.406c
Order to order	Spearman correlation	-112	0.136	-812	0.421c
Number of valid cases		54			

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

Table 8: Chi square test for relationship between sample surface weathering and sample grain

		value	Asymptotic standard error a	Approximate Tb	Progressive significance
Interval to interval	Pearson R	0.062	0.138	0.450	0.655c
Order to order	Spearman correlation	0.048	0.140	0.346	0.731c
Number of valid cases		54			

To sum up, through Spearman correlation analysis, we can see that the Spearman relationship between sample surface weathering and sample type has a strong correlation through the significance test, while the relationship between sample surface weathering and sample color, sample texture has not passed the Spearman significance test, and the correlation is not significant.

It can be concluded that the surface weathering of the sample has a significant difference with the change of sample type and sample texture, while the significant difference is small compared with the sample color; The weathering of the sample surface has a large correlation with the sample type, but has a weak correlation with the sample texture and sample color. The above conditions are in line with the actual conditions. The weathering of cultural relics will lead to the degradation of some elements, and change the content of the main components of the glass cultural relics, which will lead to the change of the category of glass cultural relics. Therefore, the weathering of the sample surface has a strong significant difference and correlation with the type of samples.

3.3 Multiple regression

In order to quantify the relationship between sample surface weathering and sample type, sample texture and sample color, multiple regression[8] method can be used. So we can set the function as follows.

$$s = f\left(V_1, V_2, V_3\right) \tag{1}$$

With the help of multiple linear regression, multiple pure quadratic regression and cross multiple quadratic regression in Matlab, the threshold value of 0.5 is taken as the verification of the results. If the threshold value is exceeded, it means weathering or normal. The following results are obtained through inspection, as shown in Figure 1, Figure 2 and Figure 3.

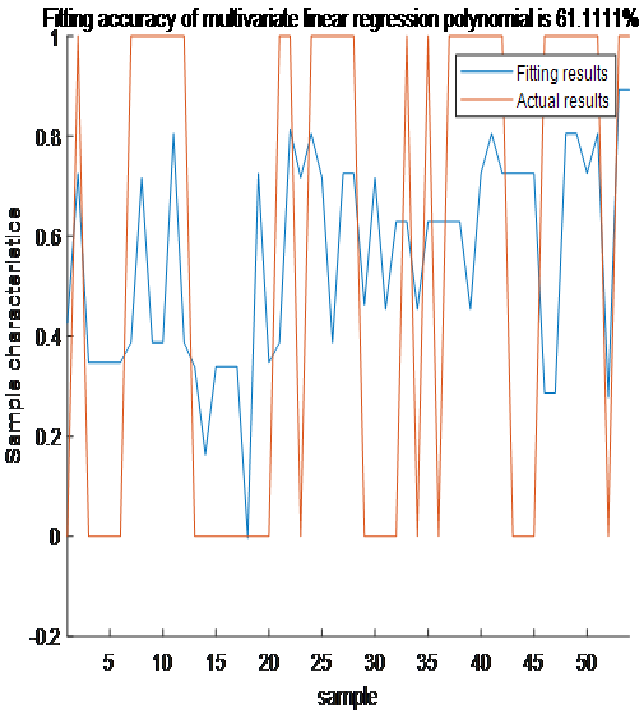

Figure 1: Fitting effect of linear multiple regression

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

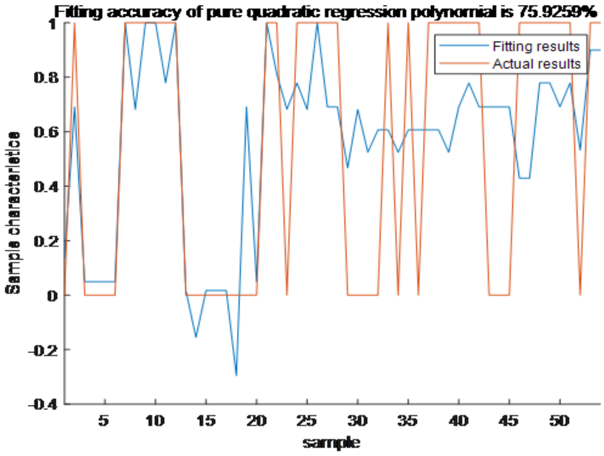

Figure 2: Fitting effect of pure quadratic multiple regression

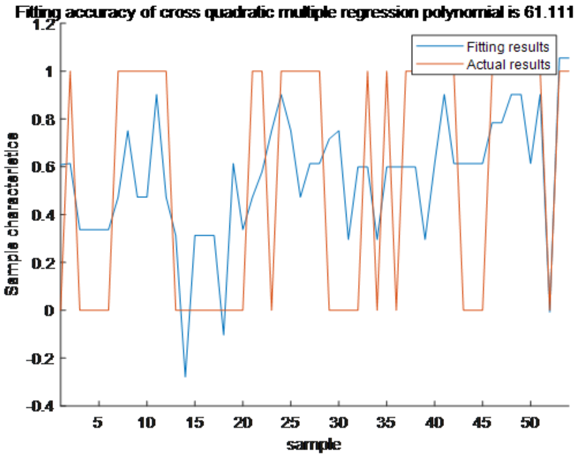

Figure 3: Cross multiple quadratic regression fitting effect

To sum up, when pure quadratic multiple linear regression is selected, the fitting effect is the best, and the accuracy reaches 75.9259%. According to this process, the pure quadratic multiple regression function is obtained as follows.

$$f(V_1, V_2, V_3) = 0.0491 + 3.7156 * V_1 + 0.7622 * V_2 - 0.9185 * V_3 - 3.6277 * V_1^2 + 0.5514 * V_3^2 \quad (2)$$

3.4 Analysis and prediction of statistical laws of chemical components in samples

According to Form 2 after data preprocessing, classify the high potassium samples and lead barium samples, and visualize the chemical species and corresponding content of each sample. The visualization results of high potassium samples are as Figure 4.

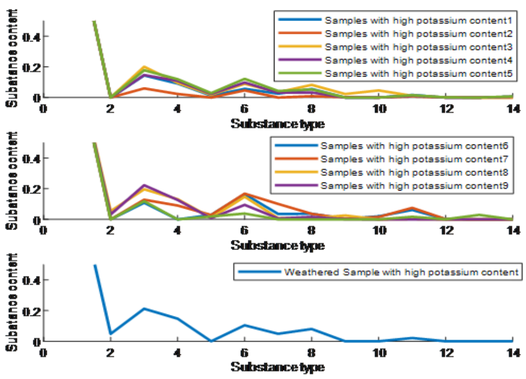

Figure 4: Distribution of substance content in each high potassium sample

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

According to the distribution pattern of chemical composition of high potassium samples, it is concluded that the species and content of high priced samples conform to the following pattern: the unweathered samples are mainly composed of silicon dioxide, the potassium and aluminum content is high, and the sample image is relatively prominent in the content of copper oxide and phosphorus pentoxide; The situation of weathered samples is similar to that of non weathered samples. The difference is that the images of weathered samples have protrusions at the nodes of copper oxide and phosphorus pentoxide. To sum up, the weathering degree of high priced samples may be greatly affected by the content of copper oxide and phosphorus pentoxide.

For the above normalized data, it is impossible to judge whether the content of silicon dioxide is related to weathering samples. The following results are obtained by analyzing the visualization of silicon dioxide content of each high potassium sample.

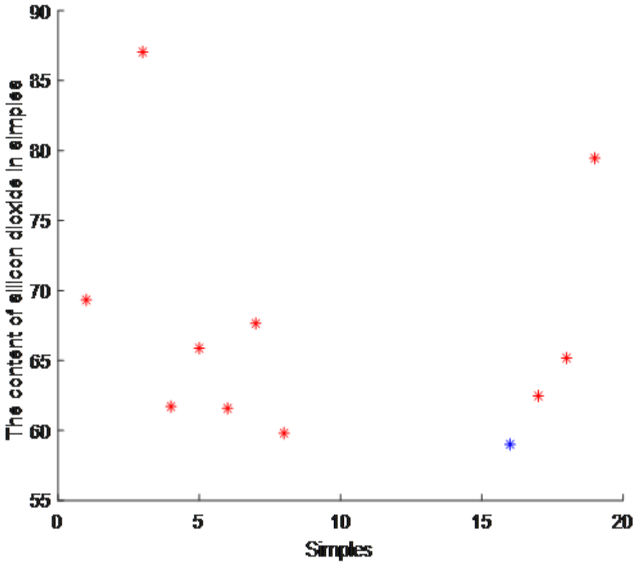

Figure 5: Distribution of Silica Content in Each High Potassium Sample

According to the analysis of the above Figure 5 (the blue speckles are the silica content of the high potassium weathered samples), the weathered high potassium samples have the lowest silica content. Therefore, the content of silica obtained from the analysis is related to weathering of high potassium samples.

Through the above analysis, it is found that the weathering of high priced samples is related to the content of copper oxide, phosphorus pentoxide and silicon dioxide. For this reason, this paper temporarily ignores the actual content of other chemical components. To predict the material content of high potassium samples before weathering, it is only necessary to predict the content of copper oxide, phosphorus pentoxide and silicon dioxide. The content of other chemical substances can be replaced by the average value of the corresponding material content of the fresh samples, and the following multiple linear regression function group can be constructed.

$$\begin{cases} S_1\left(A_1,A_2,A_3,A_4,A_5,A_6,A_7,A_8,A_9,A_{10},A_{11}\right)=W_{11}A \\ S_2\left(A_1,A_2,A_3,A_4,A_5,A_6,A_7,A_8,A_9,A_{10},A_{11}\right)=W_{12}A \\ S_3\left(A_1,A_2,A_3,A_4,A_5,A_6,A_7,A_8,A_9,A_{10},A_{11}\right)=W_{13}A \\ W_1=\left[w_1^1,w_2^1,w_3^1,w_4^1,w_5^1,w_6^1,w_7^1,w_8^1,w_9^1,w_{10}^1,w_{11}^1,w_{12}^1,w_{13}^1\right] \\ W_2=\left[w_1^2,w_2^2,w_3^2,w_4^2,w_5^2,w_6^2,w_7^2,w_8^2,w_9^2,w_{10}^2,w_{11}^2,w_{12}^2,w_{13}^2\right] \\ W_3=\left[w_1^3,w_2^3,w_3^3,w_4^3,w_5^3,w_6^3,w_7^3,w_8^3,w_9^3,w_{10}^3,w_{11}^3,w_{12}^3,w_{13}^3\right] \\ A=\left[A_1,A_2,A_3,A_4,A_5,A_6,A_7,A_8,A_9,A_{10},A_{11}\right]^{\mathrm{T}} \end{cases} \tag{3}$$

Finally, the average value of the corresponding material content of the fresh sample is substituted into the prediction to get the final result as shown in the following Figure 6.

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

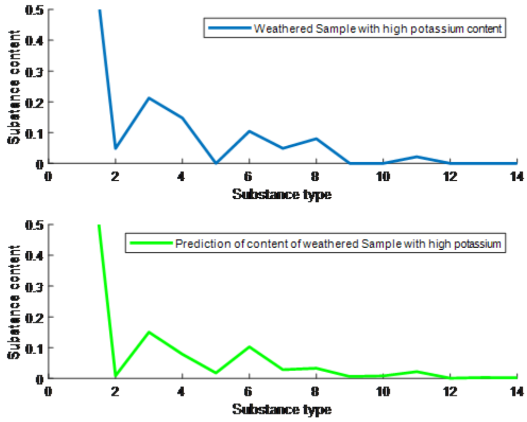

Figure 6: Comparison of Chemical Composition of Weathering Samples and Predicted Content of Samples before Weathering

Through the analysis of the above figure, it is found that the protrusions on the copper oxide nodes of the chemical composition content image of the weathered samples before weathering are significantly reduced, and the phosphorus pentoxide content is also reduced, which also confirms the correctness of our hypothesis, as shown in Figure 7, Figure 8 and Figure 9.

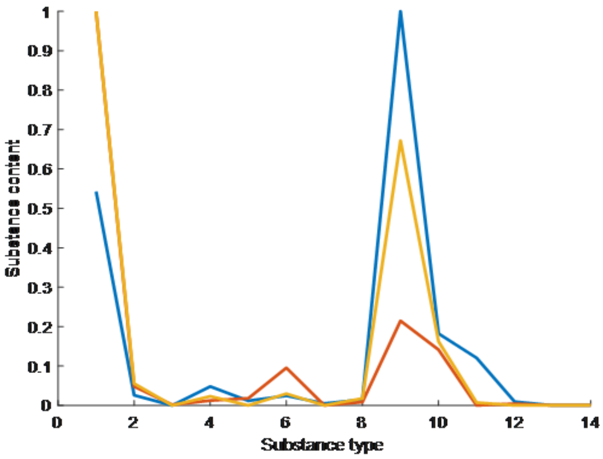

Figure 7: Chemical Content of Some Unweathered Lead Barium Samples

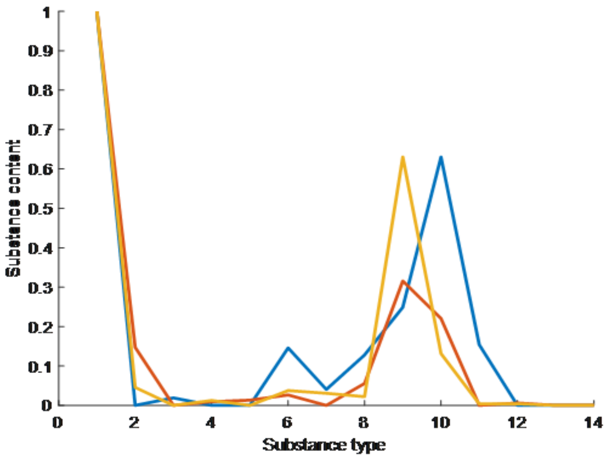

Figure 8: Chemical Content of Some Unweathered Lead Barium Samples

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

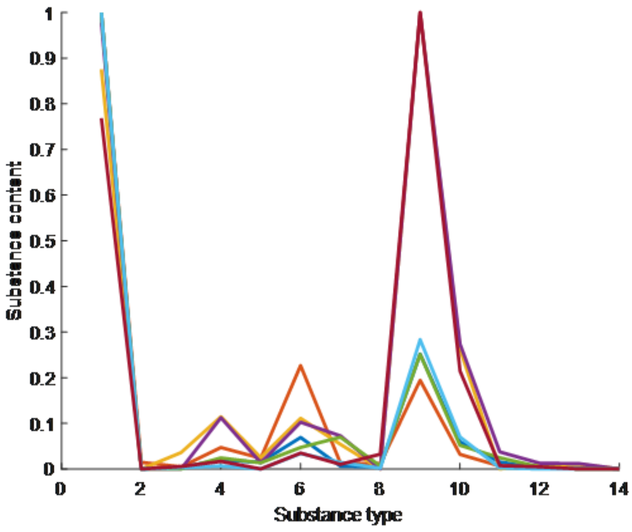

Figure 9: Chemical Content of Some Unweathered Lead Barium Samples

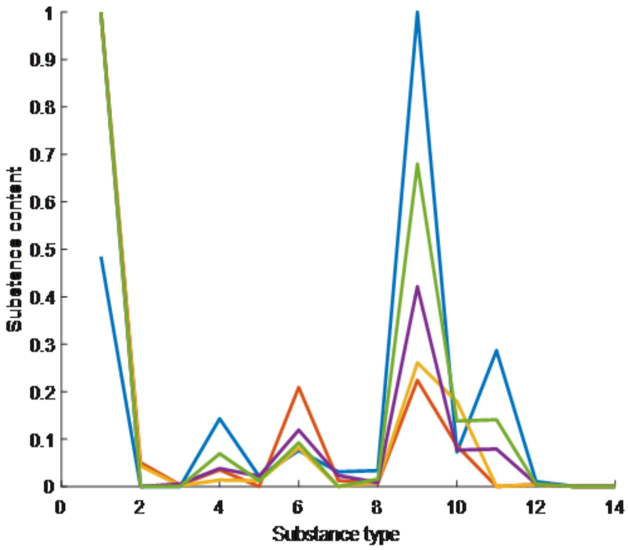

Figure 10: Chemical Content of Some Unweathered Lead Barium Samples

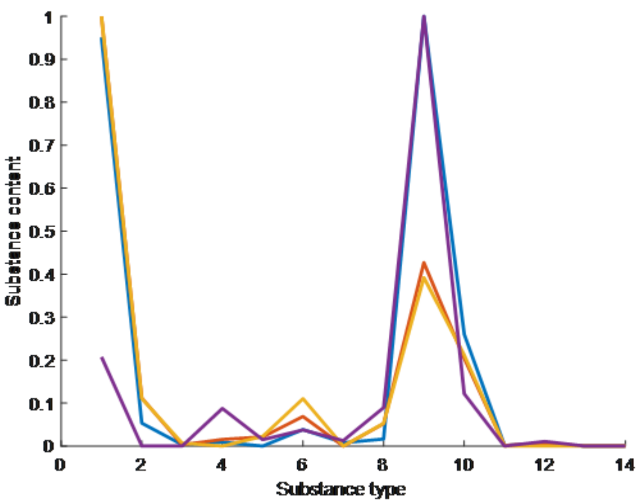

Figure 11: Chemical Content of Some Unweathered Lead Barium Samples

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

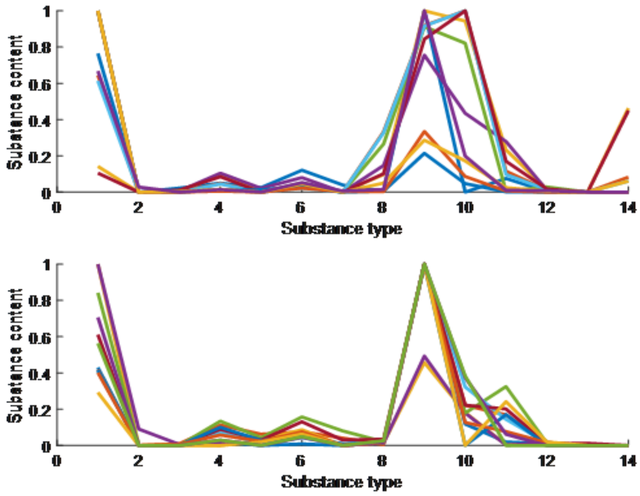

Figure 12: Contents of various chemical substances in all weathered lead barium samples

Through the above graphic analysis, it is known that the chemical content of lead barium samples has the following rule: the sample is mainly composed of silica, followed by the lead oxide and barium oxide content of lead barium samples, and finally the chemical composition diagram of this sample has a bulge at the calcium oxide, aluminum oxide, phosphorus pentoxide nodes, with a high content, as shown in Figure 10, Figure 11 and Figure 12. For the weathered sample, the biggest difference between its image characteristics and that of the fresh sample is the sulfur dioxide content. Therefore, this paper will focus on the impact of sulfur dioxide on lead barium samples. To predict the chemical content of the weathered lead barium samples before weathering, only one prediction of sulfur dioxide content is required. The prediction method still focuses on the content of the target material, and the content of other components can be replaced by the average of the corresponding material content of the fresh sample. Therefore, we only need to establish a multiple linear regression function for sulfur dioxide content, namely:

$$\begin{cases} K(B_1,B_2,B_3,B_4,B_5,B_6,B_7,B_8,B_9,B_{10},B_{11},B_{12},B_{13}) = MB \\ M = [m_1,m_2,m_3,m_4,m_5,m_6,m_7,m_8,m_9,m_{10},m_{11},m_{12},m_{13}] \\ B = [B_1,B_2,B_3,B_4,B_5,B_6,B_7,B_8,B_9,B_{10},B_{11},B_{12},B_{13}]^{\mathrm{T}} \end{cases} \tag{4}$$

4. Conclusion

In this paper, chi square test is carried out based on the weathering condition of the sample and the color, pattern and type of the sample. It is found that there is a significant difference between the style of the sample and the type of the sample. At the same time, Spearman correlation analysis was conducted on the above relationships, and it was found that the weathering of samples was closely related to the type of samples. Secondly, based on the data given in the appendix, this paper conducts multivariate fitting to quantify the relationship between the weathering of samples and other factors; At the same time, draw the content relationship diagram of the corresponding substances of each sample to analyze and summarize the chemical substance content characteristics of the samples with high potassium and lead barium. Then, in order to further explore the statistical characteristics of chemical substances in the sample, the influence of each element on the weathering results of the sample is found through principal component analysis. Finally, the content of some substances in the weathered sample data is predicted by referring to the fresh sample data, and the material content obtained is the solution of the model.

References

[1] 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): 108-116+151+0

[2] Wang Chengyu, Jiang Yanyan, Pang Shihong, Tang Huajuan, Tao Ying. Development of glass composition design[J]. Journal of the Chinese Ceramic Society, 2007, 35:143-148

ISSN 2616-5880 Vol. 3, Issue 2: 63-72, DOI: 10.25236/AJMC.2022.030210

[3] Fang Xiangzhong. Chi Square Distribution and Chi Square Test [J]. China Statistics, 2022 (05): 29-31

[4] Wang Xiaoyan, Li Meimei. A Brief Talk on Grade Correlation Coefficient and Spearman Grade Correlation Coefficient [J]. Journal of Guangdong Light Industry Vocational and Technical College, 2006, 5 (4): 2

[5] Qu Junhua, Ni Jiaming. Analysis and design of multiple regression model [J]. China Electric Power Education, 2007

[6] Jolliffe I T. Principal Component Analysis Springer Verlag [J]. Proceedings of wseas int.conf.on signal speech & image processing corfu, 2002.

[7] Liu Hailin, Xiao Junrong. An evolutionary super multi-objective optimization dimension reduction algorithm based on decomposition and hyperplane fitting [J]. Journal of Electronics and Information, 2022,44 (09): 3289-3298

[8] Wang Songgui. Linear statistical model: linear regression and analysis of variance [M]. Higher Education Press, 1999