Study on the Change Law of Weathering Composition of Glass Based on Data Statistics and Regression Model

Yihan Zhang*, Chang Feng, Kunwu Liu

Air Defense and Antimissile School of Air Force Engineering University, Air Force Engineering University, Xi'an, China, 710051

*Corresponding author: 15929936830@163.com

Abstract. Because of its chemical composition and technological characteristics, cultural relics glass is prone to weathering under the influence of the natural environment. It is significant to study its weathering mechanism and composition change law for cultural relics protection. In this study, the data correlation changes before and after glass weathering, the classification of glass types, and the degree of correlation among various chemical components were analyzed. According to the chi-square test table, the kind and ornamentation of glass have a great influence on weathering, while color has little impact on weathering. This paper provides a research idea and framework that is applied to the related fields of glass materials analysis. Based on the Fisher discrimination method, principal component analysis method, and grey correlation analysis method, it is concluded that the glass type has a great influence on weathering. The different kinds of decoration also have a significant impact on weathering.

Keywords: Chi-square Test, Fisher Discriminant Method, Principal Component Analysis, Grey Relational Analysis.

1. Introduction:

With the changes of the times and the influence of the natural environment, many cultural relics have gradually undergone weathering and aging. Among them, cultural relics glass is more likely to undergo chemical changes under long-term storage and erosion of the external environment due to its unique chemical composition and technological characteristics[1]. The problem of glass weathering not only affects the appearance and structural integrity of cultural relics but also poses a severe challenge to the preservation and restoration of cultural relics. Therefore, it is of great theoretical and practical significance for cultural relics protection to deeply understand the mechanism, influencing factors, and the changing law of chemical composition before and after the weathering of glass[2-3].

The existing research mostly focuses on analyzing the weathering phenomenon of glass through experimental means and discusses the influence of different environmental factors on the weathering process of glass. However, most often have problems such as relying on expert analysis or difficulty in processing analysis when there is a large amount of data. Therefore, this study started to use intelligent algorithms such as the chi-square test Fisher discriminant method, principal component analysis method, and grey relationship analysis method. Among them, Chi-SquareTest[4-5] is a widely used hypothesis testing method, mainly used for statistical inference of classified data, and is widely used in medicine, product design, engineering, and decision-making of almost all research projects.

Firstly, by analyzing the quantity and distribution proportion of various types of glass before and after weathering, and the changes of different chemical components of high-potassium glass and lead-barium glass before and after weathering, the chi-square test model of fixed variables and the linear regression model of chemical components before and after weathering were established. The data of glass before and after weathering were analyzed by spss software, and the proportion of each component before and after change by Matlab software[6], and the chemical components before weathering at weathered points before and after weathering were obtained. By analyzing the attachment data and counting the chemical composition in high-potassium glass and lead-barium glass, the specific characteristic types of high-potassium glass and lead-barium glass are obtained, and the principal component analysis model is established.

The data is processed by spss software, and then the subclassification method is obtained according to the results of principal component analysis, that is, based on high potassium and lead and barium, and then the Fisher discriminant analysis model is established by analyzing the data characteristics of classified glass. The data is processed by spss software, and then the difference between groups and the difference within groups are obtained for classification. Finally, by analyzing the data distribution among different chemical components, a grey correlation analysis model is established, and the grey correlation degree is obtained by using spss software to process the data, and the correlation degree and difference between different chemical components are obtained.

2. Data source and preprocessing:

The data of this study were derived from www.mcm.edu.cn.

2.1. Data cleaning

Because the data characteristics of the main components at the sampling point of cultural relics are compositional, that is, the cumulative sum of each component proportion should be 100%, but due to detection means and other reasons, the cumulative sum of component proportion between 85% and 105% may be regarded as valid data. According to the proportion of the main components at the corresponding cultural relics sampling points in the database, through the SUM function in Excel, the total proportion of the main components at 15 and 17 cultural relics sampling points is 79.47% and 71.89%, both of which are less than 85%. Accordingly, the data at the artifact sampling points 15 and 17 are deleted.

2.2. Missing value filling

Since the colors of the four samples with artifact numbers 19, 40, 48, and 58 in the database are not given, while other attributes, as well as the content of the chemical composition, are known, the missing color data are supplemented by classifying the samples according to the colors displayed by different metal cations in the compounds. The processing table is shown in the supporting document. Since the proportion of the corresponding main component of the cultural relic sampling point is given in the database, where the blank space indicates that the component is not detected, the corresponding missing value is directly filled with "0".

3. Statistical Analysis and Prediction Model of Chemical Composition Before and After Glass Weathering

3.1. The relationship between weathering and glass characteristics

(1) Chi-square test

According to the variables given in the database, it is found that several types have been determined, which belong to fixed-class variables. It is necessary to analyze the relationship between fixed variables and find out the influence of each variable on weathering, that is, to analyze the difference between the actual observed value and the theoretical inferred value. Therefore, the chi-square test can be used to find the variables whose influence on weathering is close to the actual value. The chi-square test is mainly used to compare the difference analysis between definite variables and definite variables. In this question, the glass type, ornamentation, and color are regarded as the fixed variable x respectively, and whether the surface is weathered is regarded as the fixed variable y. The influence degree of variable x on whether it is weathered is analyzed by the chi-square test.

The steps of the chi-square test are as follows:

① Put forth null hypothesis:

Since this problem is discrete, it is assumed that the distribution law of the population is H0: the distribution law of the population x is x = x i, i = 1, 2, 3... where x i is the glass type, color, and ornamentation. The occurrence probability value of each variable is represented by pi.

② Divide the range of x into intervals that do not interact with each other, that is, different glass types (high potassium, lead barium), colors (blue, light blue, blue-green...), and ornamentations (A, B, C).

③ When H0 is true, record the number of samples in each interval, count the group frequency (true value), and then find the theoretical frequency (theoretical value) according to the assumed overall theoretical distribution.

④ Introduce detection statistics:

$$x^2 = \sum_{i=1}^{k} \frac{\left(f_i - np_i\right)^2}{np_i} \tag{1}$$

The data in the database was processed using spss, and the following is the chi-square test table:

Table 1 Chi-square test table[7] of glass type, color, ornamentation, and surface weathering

Title	Name	Surface weathering		Total	χ2	Correction χ2	P
		No weathering	Fenghua				
Color	Light green	2	1	3	7.011	7.011	0.428
	Light blue	6	12	18			
	Dark green	3	4	7			
	Dark blue	2	0	2			
	Purple	2	2	4			
	Green	1	0	1			
	Blue-green	6	9	15			
	Black	0	2	2			
Type	Lead barium	12	24	36	3.861	2.758	0.049**
	High potassium	10	6	16			
Ornamentation	A	11	9	20	5.72	5.72	0.057**
	B	0	6	6			
	C	11	15	26			

Note: * * *, * *, * represent significance levels of 1%, 5%, and 10%, respectively.

According to Table 1, the following conclusions can be drawn: the influence of glass type and ornamentation on the surface differentiation of glass samples is significant, while the color of glass has little influence on it.

Among different glass types, lead-barium glass is easier to weather; among the ornamental types, Class A and Class C have little influence on weathering, while Class B has almost 100% influence. Different color kinds have no obvious effect on weathering.

Results analysis: The type of glass sample is mainly determined by its internal chemical content, and the material structure determines its properties. For example, if the content of lead and barium is high in this question, it is easier to weather, and it is easier for glass to be affected by weathering than lead and barium. Therefore, the type of glass has a greater influence on weathering; different types of ornamentation mainly determine the surface area, volume, and other factors of the sample. Different surface areas will lead to different contacts between the sample and the external environment, so ornamentation also has a great influence on weathering; the color difference only depends on a small amount of sub-group metal elements. Although some colors have a strong influence on weathering, such as light blue, other colors have no great influence on weathering, so the color will not have a great influence on weathering. Have a great influence.

(2) Statistical law of chemical composition content of glass before and after weathering

Glass types are divided into two types: high-potassium glass and lead-barium glass. According to the weathering of the glass surface, these two types of glass can be further divided into weathered high-potassium glass, unweathered high-potassium glass, weathered lead-barium glass, and unweathered lead-barium glass. Firstly, the VLOOKUP function[8-10] in Excel is used to match and merge the corresponding data in the database, and all the characteristics of each cultural relic are obtained. Secondly, the distribution of chemical composition content corresponding to each glass type is classified and counted. Finally, the change law of chemical composition content of high potassium glass and lead-barium glass before and after weathering is counted, and the statistical law is obtained. The operation steps are as follows:

① Combine the data of Form 1 and Form 2.

② Weathered high-potassium glass, unweathered high-potassium glass, weathered lead-barium glass, and unweathered lead-barium glass were obtained by screening, and descriptive statistical analysis was carried out by spas, and the histogram distribution of corresponding chemical components was drawn.

③ According to the chemical composition content before and after weathering, draw the change law of the chemical composition content of high potassium glass and lead barium glass before and after weathering. After descriptive statistical analysis of the chemical composition content of weathered high-potassium glass, unweathered high-potassium glass, weathered lead-barium glass, and unweathered lead-barium glass by spss, the histogram of the corresponding chemical composition content distribution is shown in the following figure:

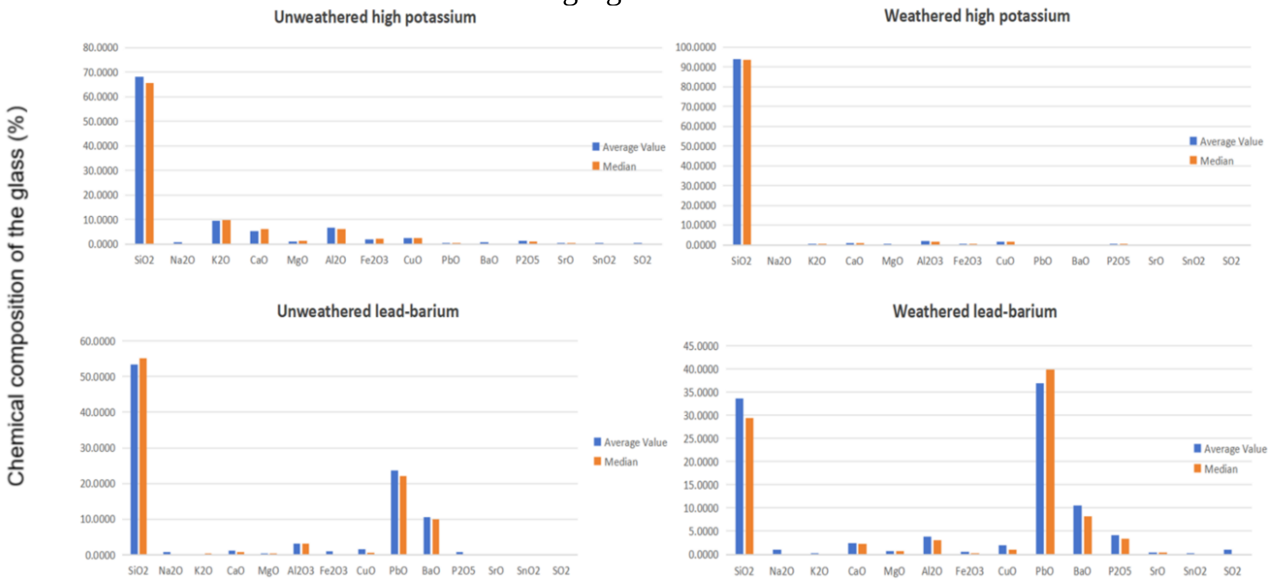

Fig. 1 The distribution of chemical composition content of high potassium glass and lead-barium glass before and after weathering

According to the distribution of chemical components of weathered high-potassium glass from Fig. 1, unweathered high-potassium glass, weathered lead-barium glass and unweathered lead-barium glass, it can be seen that the seven chemical components of silica, potassium oxide, calcium oxide, alumina, iron oxide, copper oxide and phosphorus pentoxide are relatively important before and after the differentiation of high-potassium glass, and the seven chemical components of silica, calcium oxide, alumina, copper oxide, lead oxide and phosphorus pentoxide are relatively important during the differentiation of lead-barium glass.

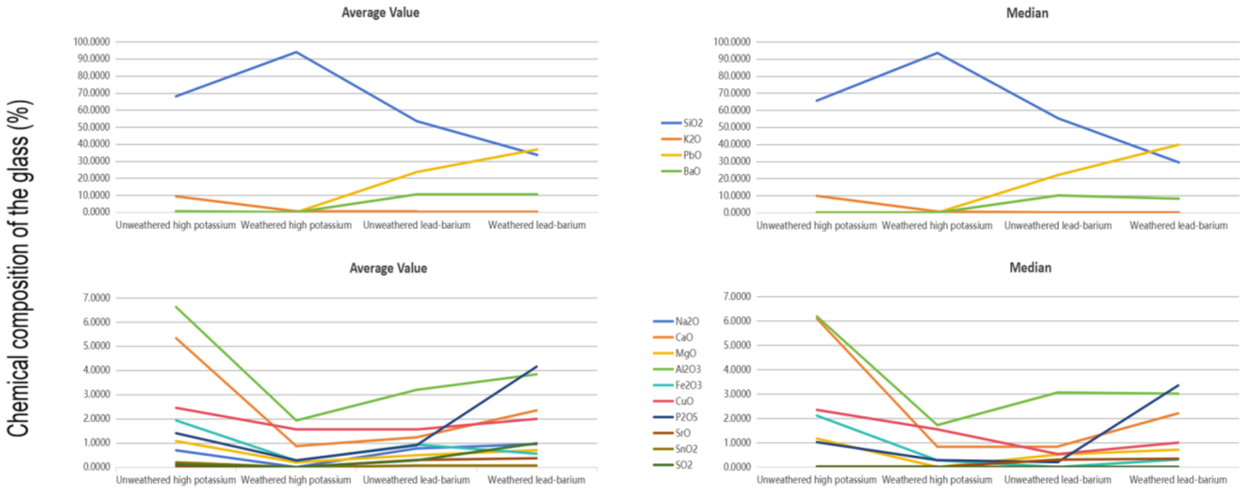

Fig. 2 Trend chart of different chemical components before and after glass differentiation

It can be seen from Fig. 2 that before and after weathering, the content of silica in high potassium glass shows an upward trend, while the contents of six chemical components, namely potassium oxide, calcium oxide, alumina, iron oxide, copper oxide, and phosphorus pentoxide, show a downward trend. Before and after weathering, the content of silica in the lead-barium glass showed a decreasing trend, while the contents of calcium oxide, alumina, copper oxide, lead oxide, phosphorus pentoxide, and sulfur dioxide showed an upward trend.

3.2. Predictive type based on linear regression relationship

(1) Constructing regression relations

From the data in the form, it can be seen that: the distribution and specific content of chemical composition content of each sample before weathering and the distribution and content of each chemical component after weathering, so the change quantity of each commodity can be obtained, and the regression coefficient can be used to solve the quantity before weathering.

However, it is observed that the content of the same substance (such as sodium oxide) in each glass is quite different, so the normal distribution curve is used to weigh and average the data to obtain universal data.

The steps are as follows: the contents of the 14 chemical components before weathering are expressed as q1, q2, q3, …, q14, respectively, and the weights after weathering are expressed as q1, q2, q3, …, q14, and the weights after weathering are expressed as k1, k2, …, k14, respectively, and after weathering, k1, k2, …, k14 are weighted according to the following normal distribution curve:

$$f(q)=\frac{1}{\sqrt{2\pi}}e^{\frac{q^2}{2}} \tag{2}$$

Then the weighted average value of each chemical component is obtained as follows:

$$\begin{cases} \bar{q}=\dfrac{q_1k_1+q_2k_2+\ldots+q_{14}k_{14}}{k_1+k_2+\ldots+k_{14}} \\ \bar{q}'=\dfrac{q_1'k_1'+q_2'k_2'+\ldots+q_{14}'k_{14}'}{k_1'+k_2'+\ldots+k_{14}'} \end{cases} \tag{3}$$

Then the proportion before and after weathering is:

$$\eta=\frac{\overline{d_i}}{\overline{d_i'}} \tag{4}$$

Then, the content of each chemical component(4) at the weathering point is determined by the formula (4), and the content of the chemical component before weathering is determined.

(2) Outcome prediction

The prediction process is implemented by Matlab, and the results are as follows:

Table 2 Prediction data before weathering of some weathering points

Ornamentation	Type	Color	Surface weathering	Cultural relics sampling point	Silica	Potassium oxide	Lead oxide	Barium oxide
B	High potassium	Blue-green	Weathering	7	75.58805	0	0	0
B	High potassium	Blue-green	Weathering	27	73.81449	0	0	0
B	High potassium	Blue-green	Weathering	9	73.34773	11.64304	0	0
B	High potassium	Blue-green	Weathering	22	59.92126	12.27491	0	0
B	High potassium	Blue-green	Weathering	12	67.4564	18.47232	0	0
C	Lead barium	Dark green	Weathering	34	68.9343	0.42148	19.93275	6.877643
C	Lead barium	Dark green	Weathering	36	60.65529	0.18779	14.17597	5.926191
C	Lead barium	Light blue	Weathering	43 Site 1	37.24268	0	39.91956	7.809824
C	Lead barium	Blue-green	Weathering	57	58.19698	0	22.94848	14.1389

First, glass samples of the same type are selected and grouped. For each group, the weighted average of each chemical component is calculated. The mean values are then used to determine the content of each chemical component before and after weathering for each sample type. Subsequently, a linear regression model is constructed to analyze the relationship between the chemical compositions before and after weathering. Using the ratios of different chemical components before and after weathering, the content before weathering is predicted, resulting in the data presented in Table 2.

4. Conclusions

In this paper, a chi-square test model was used to analyze the relationship between fixed variables and surface weathering, while a linear regression model was established to explore the changes in chemical composition before and after weathering. The results indicate that glass type has a significant impact on weathering, with lead-barium glass being more prone to degradation. Similarly, ornamentation plays a crucial role, as different types of surface designs influence the exposure of the glass to environmental factors. However, color differences, despite being influenced by certain subgroup metals like light blue, exhibit a minimal overall effect on weathering. This confirms that structural and compositional characteristics outweigh aesthetic attributes in determining the glass’s susceptibility to weathering.

This study provides a methodological framework for analyzing glass materials, offering insights into the weathering mechanisms and compositional changes of cultural relic glass under natural environmental conditions. Future research will focus on expanding the dataset to include more diverse glass samples and environmental conditions, refining predictive models for chemical composition changes, and exploring the microstructural effects of weathering. These efforts aim to deepen the understanding of glass weathering processes and further enhance conservation strategies for ancient glass artifacts.

References

[1] Qu Song Research on the Disease Mechanism and Weathering Degree Evaluation System of Marble Cultural Relics in Beijing Area [D] Beijing University of Chemical Technology, 2018

[2] Zhao Fangqin Application of Material Characterization Technology in the Study of Weathering of Sandstone Cultural Relics [D] Harbin Institute of Technology, 2010

[3] Lu Chunhai Research on Material Design and Applicability for Preventing Cultural Relics from Weathering [D] Chengdu University of Technology, 2009

[4] Fang Xiangzhong Chi square distribution and chi-square test [J] China Statistics, 2022, (05): 29-31

[5] Chai Yaping, He Ying, Zhou Yanxiang, etc Analysis of the Implementation Effect of Techniques and Tactics in Excellent Table Tennis Athletes Based on chi-square Test [J] Sports Boutique, 2024, 43 (04): 89-92

[6] Li Changxiao Matlab Program Implementation of Ray Tracing Algorithm and Its Application in Aberration Analysis of Optical Systems [D] Huazhong University of Science and Technology, 2023

[7] Li Shiqi, Chen Zhengwei, Wu Weican Research on the home advantage of the 2018-2019 CBA league based on chi-square test [J] Sports Science and Technology Literature Bulletin, 2021, 29 (08): 103-106

[8] Gao Jianbin Application of VLOOKUP Function in Paperless Metrology Verification [J] Measurement and Testing Technology, 2023, 50 (12): 62-64

[9] Huang Jia Application Analysis of Single column Search and Multi-column Search Methods for VLOOKUP Function in Excel Data Tables[J] Computer Knowledge and Technology,2023,19 (24):34-36

[10] Huang Jia The Basic Usage of VLOOKUP Function [J] Computer Age, 2022, (10): 121-122+126