A Study of Weathering of Glass Artifacts Based on Spearman's Correlation Coefficient and Chi-Square Test

Xinyu Du *, Xinying Li

Guangdong University of Technology, Guangzhou, China

* Corresponding Author Email: 2782275192@qq.com

Abstract. Ancient glass is very susceptible to weathering by the influence of the burial environment, in order to determine whether the glass artifacts weathered with the type of artifacts, ornamentation and color and other characteristics of the relationship, this paper first identifies the correlation between the artifacts, the use of Spearman correlation coefficient of the artifacts of the ornamentation, the type and the color of the artifacts with the surface of the correlation between the weathering, and then the use of the fixed type of data of the difference between the analysis of the chi-square test. It was finally concluded that whether the surface of the artifacts was weathered or not was correlated with their decoration, type and color, while whether the surface of the artifacts was weathered or not was significantly different from their type. Finally, type was used as a fixed variable to determine the relationship between weathering of artifact surfaces and their decoration, color and total chemical content range. Weathered artifacts are then mapped one by one to unweathered artifacts through Euclidean distances, and the artifact with the smallest vector distance is used as a predictor of its pre-weathered chemical composition content.

Keywords: Spearman's correlation coefficient, chi-square test, Weathering of glass artifacts.

1. Introduction

Glass, valuable physical evidence of early trade between the West and China in ancient times, is made from quartz sand, whose main chemical component is silicon dioxide (SiO2) [1-2]. However, due to the high melting point of pure quartz sand, it is necessary to add a flux when using it to make glass, as well as limestone as a stabilizer. Different fluxes are added, and their chemical compositions are also different - lead-barium glass adds lead ore as a flux during the firing process, and its lead oxide (PbO) and barium oxide (BaO) content is higher, and it is usually regarded as a glass variety of China's own invention, and the glass of the Chu culture is dominated by lead-barium glass; potash glass uses substances with a high content of potassium, such as wood ash, as a stabilizing agent; and potassium glass uses substances with a high potassium content, such as wood ash, as a stabilizing agent. Potassium glass is made of potassium-containing substances such as grass ash as a flux fired, mainly popular in Lingnan, China, as well as Southeast Asia and India and other regions. Ancient glass is very susceptible to weathering under the influence of the buried environment, during which its internal elements are exchanged with the external environment in a large number of elements. For the surface weathering of these glass relics, the relationship between their glass type, decoration and color is analyzed; based on the type of glass, the statistical law of the chemical composition content on the surface of the relic samples with or without weathering is analyzed and based on the data of the weathering point detection, the chemical composition content of the relics before weathering is predicted [3].

2. Modeling and analysis of relationship analysis

In this paper, we analyze the correlation and variability of surface weathering with its ornamentation, type, and color by using Spearman's correlation coefficient and chi-square test to obtain the relationships [4-5].

2.1. Spearman's correlation coefficient analysis

The correlation coefficients for each feature of the glass artifacts are shown in Table 1.

Table 1 Correlation coefficients

	color	typology	figure decorative motif	Surface weathering
Color	1.000(0.000***)	0.105(0.435)	-0.303(0.021**)	0.309(0.018**)
Type	0.105(0.435)	1.000(0.000***)	0.119(0.374)	0.344(0.008***)
Texture	-0.303(0.021**)	0.119(0.374)	1.000(0.000***)	0.037(0.781)
Surface weathering	0.309(0.018**)	0.344(0.008***)	0.037(0.781)	1.000(0.000***)
Note: ***, **, * represent 1%, 5%, and 10% significance levels, respectively.				

The method of analysis is as follows:

1. first test whether there is a statistically significant relationship between color, type, grain and surface weathering to determine whether the P value presents significance ($P<0.05$);

2. if it presents significance, it means that there is a correlation between the two variables, and vice versa, there is no correlation between the two variables.

3. analyze the positive and negative direction of the correlation coefficient and the degree of correlation.

2.2. Chi-square test

The results of the chi-square test analysis are shown in Table 2, and the chi-square test equation:

$$D_f = (C - 1)(R - 1) \tag{1}$$

Table 2 Results of chi-square test analysis

title	Name	Surface weathering		total	X^2	Calibrate X^2	P
		weathering-free	public morals				
Color	Blue-green	6	9	15	6.287	6.287	0.507
	Light blue	8	12	20			
	Purple	2	2	4			
	Dark green	3	4	7			
	dark blue	2	0	2			
	Light Green	2	1	3			
	Black	0	2	2			
	Green	1	0	1			
Total		Total	30	54			
Type	High Potassium	12	6	18	5.400	4.134	0.020**
	Lead Barium	12	24	36			
Total		24	30	54			
Pattern	C	13	15	28	5.747	5.747	0.056*
	A	11	9	20			
	B	0	6	6			
Total		24	30	54			
Note: ***, **, * represent 1%, 5%, and 10% significance levels, respectively.							

The methods of analysis were as follows:

1. analyze whether the model presents significance ($P<0.05$).

2. if it presents significance and the original hypothesis is rejected, it means that there is a significant difference between the samples, which is described according to the percentage of difference in the categories; conversely the data do not present significant differences [6].

The results of the analysis are as follows:

1. the results of the chi-square test analysis show that based on the surface weathering and color, the significance p-value is 0.507, the level does not present significance, the original hypothesis is accepted, so for the surface weathering and color data do not have significant differences.

2.The result of chi-square test analysis shows that based on surface weathering and type, the significance p-value is 0.020**, which presents significance at the level and the original hypothesis is rejected, therefore there is a significant difference for surface weathering and type data.

3. The results of the chi-square test analysis show that based on surface weathering and ornamentation, the significance p-value is 0.056*, which does not present significance at the level and the original hypothesis is accepted, therefore there is no significant difference for surface weathering and ornamentation data.

Table 3 Quantitative analysis of effects Table 3 Quantitative analysis of effects

Field Name/Analysis Item	Phi	Crammer's V	Number of columnar links	lambda
Color	0.341	0.341	0.323	0.000
Type	0.316	0.316	0.302	0.000
Pattern	0.326	0.326	0.310	0.000

The quantitative analysis of effects is shown in Table 3 above and analyzed as follows:

1. when a significant difference is presented (prerequisite), the difference is quantified in conjunction with the analysis of the quantitative indicators of effect [7].

2. the quantitative indicators of effects reflect the degree of correlation between variables.

3. depending on the type of crossover, different effect size indicators can be selected (crossover type is expressed as: the number of horizontal grids of the crossover table x the number of vertical grids).

4. phi coefficient: the size of the phi correlation coefficient indicates the degree of correlation between the two samples. When the phi coefficient is less than 0.3, it indicates a weak correlation; when the phi coefficient is greater than 0.6, it indicates a strong correlation (used in the 2 × 2 cross type table).

5. Cramer's V: works similarly to the phi coefficient, but the Cramer's V coefficient has a wider range of effects. When two variables are independent of each other, V=0. When there are only 2 dichotomous variables in the data, Cramer's V coefficient has the same result as phi (if $m \neq n$, it is recommended to use Cramer's V);

6. Column linkage number: C coefficient for short is used for 3×3 or 4×4 crosstabs, but it is affected by the number of rows and columns, which increases with R and C. Therefore, depending on the number of rows and columns and the calculated coefficients, the C coefficient is the same as phi. Therefore, column linkage numbers calculated based on different rows and columns are not easy to compare, unless the number of rows and columns in the two linked tables are the same;

7. lambda: used to react to the prediction effect of the independent variable on the dependent variable, in general, its value of 1 indicates that the independent variable predicts the dependent variable better, and 0 indicates that the independent variable predicts the dependent variable poorly (X or Y have fixed-order data, it is recommended to use lambda).

The results of the analysis are as follows:

1. the results of the quantitative analysis of effects show that the analytical term: color Cramer's V-value is 0.341, so the degree of difference between color and surface weathering is a moderate difference.

2. the results of the quantitative analysis of effects show that the analytical term: type Cramer's V value is 0.316, so the degree of difference between type and surface weathering is a moderate difference.

3. the results of the quantitative effects analysis show that the analytical term: ornamentation has a Cramer's V value of 0.326, so that the degree of difference between ornamentation and surface weathering is medium.

3. Analysis of statistical patterns

From the above analysis of variability, we analyzed the relationship between weathering and ornamentation, and color by using glass type as a fixed variable and obtained Table 4.

Table 4 Descriptive statistics league table

Glass Type	Surface weathering	decorative motif	Color	Percentage of each color type	Total content range
High Potassium	Weathered	B	Blue green	100%	[99.81%,100%]
	Unweathered	A、C	blue green	61.53%	[97.25%,100%]
			Light Blue	15.38%	
			dark green	7.69%	
			Light Blue	7.69%	
			Dark Blue	7.69%	
Lead Barium	Weathered	A、C	Light Blue	42.86%	[90.17%,99.89%]
			Purple	7.14%	
			Gray	14.29%	
			blue green	10.71%	
			dark green	14.29%	
			Light green	3.57%	
			black	7.14%	
	Not weathered	A、C	light blue	30.77%	[88.41%,99.98%]
			Purple	15.38%	
			dark blue	15.38%	
			Light Green	7.69%	
			Dark Green	7.69%	
			green	7.69%	

As can be seen from the above table, in the high potassium glass, there are only B decorations, the color of blue-green weathering type of glass, indicating that the high potassium glass with B decorations is the easiest to weather, and the color after weathering is blue-green, and the range of its chemical composition content is [99.81%,100%]. In the lead-barium glass, there are only two types of decorations A and C, and their chemical composition ranges from [90.17%,99.89%], and the highest weathering color is light blue, and gray-yellow, black is the unique color of the weathering type, and we can speculate that the light blue color is the color of the beginning of the weathering soon after the weathering, and the gray-yellow and the black color is the color of the weathering after the severe weathering.

4. Prediction of chemical composition content

In this paper, the type and ornamentation of the weathered artifacts are quantified, the unweathered artifacts identical to them are screened out, the distance is calculated through the Euclidean distance metric, and the unweathered artifacts that are most similar to the weathered artifacts are derived from multiple data comparisons, so as to predict the pre-weathering chemical composition content.

At the same time, take the artifact number 2 as an example, and the remaining weathered artifacts' pre-weathering chemical composition content by analogy [8-9].

Let the chemical composition content of the artifacts be vector Y= (x1, x2, x3,x4,......,x14), where x1,x2,x3,x4,......,x14 denote the silica, sodium oxide and potassium oxide of the known types of artifacts, respectively, Calcium oxide ...... sulfur dioxide content.

Euclidean distance formula: D= sqrt ((Yi-Yj) ^2)

Specific calculations are shown in Table 5.

Table 5 Euclidean distance calculation results

Artifact number	decorative motif	typology	Color	Surface weathering	Euclidean distance D
02	A	lead and barium	Light blue	Weathering	0
20			Light blue	No weathering	54.9311977
23 Unweathered			blue green		51.82318429
28 Unweathered point			light blue		51.07815367
29 Unweathered point			Light Blue		55.28549976
30 Part 1			dark blue		26.17267507
30 Part 2			dark blue		19.24988255
42 Unweathered Point 1			Light Blue		44.68934835
42 Unweathered Point 2			Light Blue		46.83525144
44 Unweathered point			Light Blue		55.82096002
45			Light Blue		55.07529869
46			Light Blue		42.73206521
47			light blue		41.85588261
49 Unweathered point			Black		31.82979354
50 Unweathered point			black		28.91615469
53 Unweathered point			light blue		57.93128995

From Table 5 above, it can be seen that the artifact number 02 is most similar to the artifact number 30 part 2, so it is predicted that its artifact chemical composition content is the same as that of the artifact number 30-part 2 sampling site [10].

5. Conclusions

Ancient glass is very susceptible to weathering under the influence of the buried environment, during which its internal elements are exchanged in large quantities with those of the external environment.

This paper analyzes the correlation and difference between the glass artifacts, the glass artifacts are more accurately classified into several types for comparison, and these types of indicators are further analyzed using different methods. First of all, to identify the correlation between artifacts, due to the artifact data belongs to non-parametric relationship, so this paper adopts the Spearman correlation coefficient on the artifacts of the decoration, type and color and whether the surface weathering for correlation analysis. And in the difference analysis, because all kinds of data are fixed class data, this paper uses fixed class data to analyze the difference in the way of chi-square test. The final conclusion is that whether the surface of cultural relics is weathered or not has correlation with its decoration, type and color, and whether the surface of cultural relics is weathered or not has significant difference with its type. Based on this conclusion, this paper takes type as a fixed variable as the premise of the study. Through descriptive statistics, this paper determines the percentage of whether the artifacts are weathered or not and obtains the relationship between the surface weathering of the artifacts and its decoration, color and total chemical content range. Then, through the Euclidean distance, the weathered artifacts are mapped one by one with the unweathered artifacts, and the artifact with the smallest vector distance is used as the predictor of its chemical composition content before weathering.

References

[1] YAN Weiwen, REN Hyunyu, SHEN Yadong. Research on surface differentiation of ancient glass artifacts based on clustered K-means algorithm [J]. Chemical Engineering and Equipment,2023(07):37-43+70.

[2] T.Y. Deng,Y.H. Qiu,J.H. Chi. CLR-based compositional analysis and classification model for glass artifacts[J]. Mathematical Modeling and its Applications,2023,12(02):41-51.

[3] He Y. Compositional identification of weathered glass artifacts based on NSGA-II algorithm and entropy value TOPSIS method [J]. Academic Journal of Materials and amp; Chemistry,2023,4(3).

[4] Wang H,Zhai J,Tao W. Prediction of chemical composition content of ancient glassware before weathering [J]. Materials and amp; Chemistry Journal,2023,4(1).

[5] Ying Z. Analysis of molecular composition of glass chemical compositions based on time series and clustering methods [J]. Molecules,2023,28(2).

[6] Li Q,Wang C,Zhai J. Compositional analysis and identification study based on ancient glass artifacts [J]. Academic Journal of Materials and amp; Chemistry,2022,3(2).

[7] YANG Yun,SONG Qingyi,YUN Xinyu et al. Research on knowledge point extraction of glass artifacts based on BiLSTM-CRF [J]. Journal of Shaanxi University of Science and Technology,2022,40(03):179-184.

[8] Brigitte B,Laure D. Specialized potash glass artifacts from Tissamaharama (Sri Lanka) [J]. Journal of Glass Research, 2022, 64.

[9] Carter A,Dussubieux L,Polkinghorne M, et al. Glass artifacts from Angkor: evidence of exchange [J]. Archaeological and Anthropological Sciences,2019,11(3).

[10] Cao Caixia,Guo Hong. Classification of disease types of stone, ceramic and glass artifacts in collections using fuzzy mathematical knowledge [J]. Journal of Beijing Union University (Natural Science Edition),2009,23(04):58-60.