Classification and Prediction of Ancient Glass Artifacts based on Chi-square Test

Yifan Chen *, Jiayi Nie, Ke Xu

Sichuan University Pittsburgh Institute, Sichuan University, Chengdu, China

* Corresponding Author Email: chenyifan2526@stu.scu.edu.cn

Abstract. This paper focuses on the composition analysis and type identification of ancient glass objects. The correlation between the weathering of the surface of glass artifacts and its type, color and decoration is obtained by establishing a chi-square test model based on the characteristics of fixed class variables. After that, the prediction model was established based on the weighted average method according to the characteristics of the data, and the predicted value of each chemical composition before weathering was solved.

Keywords: Chi-square Test; glass artifacts; predictive model.

1. Introduction

The Silk Road was an important channel for trade and commerce between the East and West in ancient times, and an important link for cultural exchanges between the East and West [1]. With the prosperity of the ancient Silk Road, the material and cultural exchanges between China and Central Asia, West Asia and even the Egyptian region continued to be close, promoting the continuous development of commerce, transportation and handicrafts along the road, and playing an important role in the development of ancient civilization and history [2].

There is a batch of data related to ancient glass products in China, and archaeologists have conducted chemical composition analysis of the samples and classified them into two types [3]: high-potassium glass and lead-barium glass. In addition, basic data have been collected to record the basic information of glass artifacts, the proportion of chemical composition of classified glass artifacts and the proportion of chemical composition of unclassified glass artifacts, respectively [4]. A mathematical model was developed to analyze the relationship between the weathering of the surface of glass artifacts and their glass type, decoration and color; the statistical pattern of the chemical composition content with and without weathering on the surface of the artifact samples was analyzed, and on this basis, the chemical composition content of the relevant glass artifacts before weathering was predicted.

2. Model assumptions and notation

2.1. Assumptions [5]

1. Assume that each variable in the sample conforms to a normal distribution
2. Assume that all data measurements are true and valid, without considering the effects of human error
3. It is assumed that the small amount of data is also representative
4. the significance level in this paper is 0.05.

2.2. Notations

Important notations used in this paper are listed in Table 1.

Table 1. Notations

Symbols	Meaning
x_{i1}	The decoration of the i-th sample
x_{i2}	Type of sample i
x_{i3}	Color of sample i
x_{i4}	Surface weathering of sample i
$y_{ij}\ j = 1,2,\ldots,14$	Chemical composition of sample i
w_{ij}	Weighting weights of sample i
m	Number of clusters in the clustering model

3. Model construction and solving

3.1. Analysis of the relationship between factors on the surface of glass artifacts

3.1.1 Chi-square Test -based assessment modeling

In order to further analyze the relationship between surface weathering and glass type, decoration and color, a chi-square model was developed to assess the significance of the relationship between surface weathering and other factors [6-8].

3.1.2 Solution of chi-square value

According to the basic definition of the chi-square distribution [9], it can obtain the formula for calculating the chi-square value as.

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} \tag{1}$$

In support of the null hypothesis, we consider that all variables xi satisfy independent distribution and do not have significant correlation with the variable x_4.

On this basis, the probability of all values of xi is theoretically equal. Under the condition of independent distribution, the probability of occurrence of all three values is 1/3, so the expectation value of the number of occurrences of A, B and C is 58/3, and under this condition, there is no correlation between the variables x_1 and x_4.

Therefore, the expected value of e1 in Eq. is 29/3. By continuing with the observed values of the number of occurrences of each value of the variables x_1 and x_4, it can obtain

The chi-square values of the ornamental variables are obtained.

The results of the chi-square test for each variable are shown in Table 2.

Table 2. Chi-square test on ornamentation, type and color of artifact samples

	X^2	Numerical value
Ornament	5.72	0.057
Type	3.861	0.049
Color	7.011	0.428

From the above table, it can be seen that there is a significant correlation between the surface weathering of artifact samples and artifact decoration and color, and there is no significant correlation with glass type.

3.2. Statistical law analysis on chemical composition

On the premise that the correlation between the weathering on the surface of the samples and the decoration, glass type and color of the artifacts has been obtained by the chi-square test, in order to further obtain the statistical pattern of the chemical composition content of the samples with or without weathering for a certain type of glass, the 67 sets of data were divided into 4 categories, namely g1 high potassium weathered glass, g2 high potassium unweathered glass, g3 lead-barium weathered glass, g4 lead-barium unweathered glass, and the characteristics of the chemical composition content within each category were obtained by means of averaging, and the results are shown in Table 3 below.

Table 3. Mean values of chemical composition for each type of glass sample

	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
G1	93.96	0.00	0.54	0.87	0.20	1.93	0.27	1.56	0.00	0.00	0.28	0.00	0.00	0.00
G2	67.77	0.83	9.57	5.73	1.05	6.41	1.79	2.42	0.38	0.46	1.18	0.03	0.24	0.12
G3	27.73	0.27	0.15	2.56	0.65	3.07	0.63	.22	43.20	11.01	4.51	0.37	0.07	0.22
G4	54.93	0.84	0.22	0.97	0.47	3.12	0.81	1.69	22.35	10.51	0.92	0.29	0.04	0.31

The highest content of silica (SiO2) in the high potassium weathered glass was 93.96%, while the lowest content of silica (SiO2) in the lead-barium weathered glass was only 27.73%. At the same time, the four types of glass samples also showed significant characteristics in terms of various chemical compositions such as potassium oxide (K2O) and calcium oxide (CaO).

In order to further analyze the patterns existing behind the data, the variability of specific variables in each type of glass samples was visualized and statistically summarized.

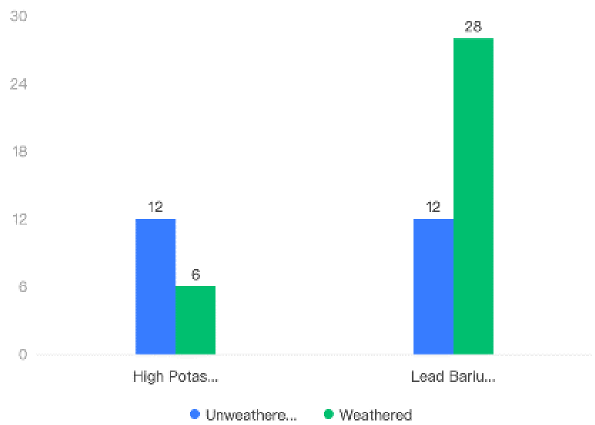

Figure 1. The relationship between glass type and weathering degree

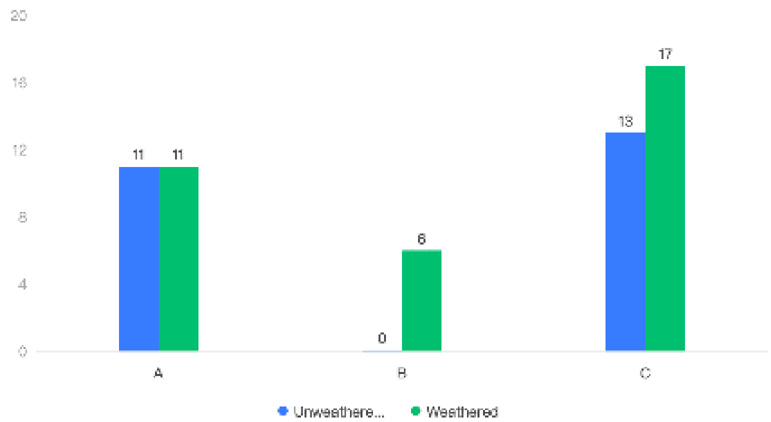

Figure 2. The relationship between ornamentation andthe degree of weathering

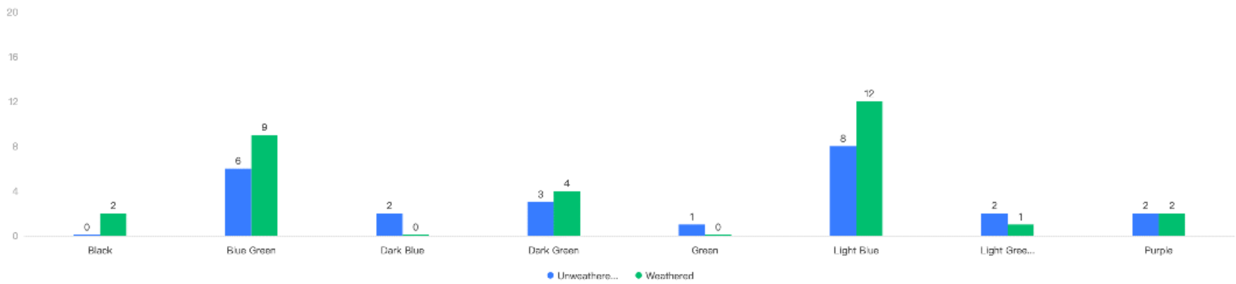

Figure 3. The relationship between color and weathering degree

As shown in Figure 1 above, the number of weathered lead-barium glass samples far exceeds the number of unweathered lead-barium glass samples, while high potassium glass shows the opposite characteristics. Therefore, lead-barium glass has a higher probability of weathering compared to high-potassium glass.

As shown in Figure 2 above, all of the glass samples with motif B have been surface weathered, while the number of glass samples with motif A or C is about half of the overall number. Therefore, we can conclude that the samples with motif B are more susceptible to weathering, while motifs A or C do not have a significant effect on surface weathering.

As shown in Figure 3 above, comparing the number of samples with surface weathering within each sample color, it can be concluded that dark blue and green glass samples may be less susceptible to weathering.

3.3. Prediction of chemical composition content before weathering

Among the samples of artifacts, there were large differences in chemical composition content before and after weathering, as shown in Figure 4 below.

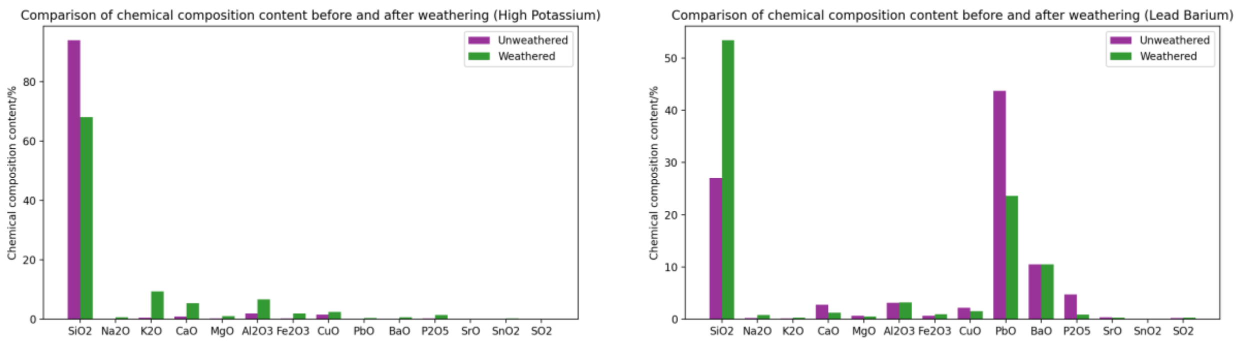

Figure 4. Comparison of chemical composition content of glass samples before and after weathering

Therefore, to improve the prediction accuracy of the chemical content of the samples before weathering, we add a weight to each chemical composition variable and derive its total weighted average value on this basis [10].

It is assumed that the weighted average of each item of the sample after weathering is $v_i{}'$ and before weathering is v_i.

$$v_i = \frac{\sum_{i=1}^{14} w_i y_i}{\sum_{i=1}^{14} w_i} \tag{2}$$

$$v_i = \frac{\sum_{i=1}^{14} w_i y_i}{\sum_{i=1}^{14} w_i} \tag{3}$$

The ratio r of the two is

$$r_i = \frac{v_i}{v_i{}'} \tag{4}$$

On this basis, the predicted values of the chemical composition of different artifact samples before weathering were obtained by programming in Python, and the results are shown in Table 4 below.

Table 4. On the prediction of chemical content of artifact samples before weathering

Glass Type	Artifact Number	Silicon Dioxide	Sodium oxide	Potassium oxide	Calcium oxide	Magnesium oxide	. . .	Sulfur Dioxide
High Potassium	7	66.4367	0.8834	9.02317	5.9345	0.85733		0.122
High Potassium	7	68.8267	0.834	9.61317	5.4845	0.85733		0.122

4. Conclusion

In this paper, based on the mechanism of the occurrence of surface weathering of glass samples, the type and appearance of glass artifacts may directly determine their surface weathering. Meanwhile, based on the available data, the analysis of the statistical pattern of the chemical content of artifact samples with and without weathering can help predict the chemical content of the samples before weathering. Combining the known data, the surface weathering of glass artifacts may be related to glass type, decoration, and color. Among them, the types of glass artifacts are high potassium as well as lead-barium, the decorations are A, B, and C, the colors are mainly composed of blue-green, light blue, light green, dark blue, dark green, and purple, and the surface weathering is weathered or unweathered. Considering that all the above four variables belong to definite class variables, a model based on chi-square distribution can be established to assess the relationship between the surface weathering of glass artifact samples and their glass type, decoration and color. On this basis, combining the data, different glass artifacts differ in their chemical composition and proportions, and based on the artifact number, a link can be established between the data and thus the glass type can be added to the study of the pattern of chemical composition content. Considering the variability of chemical composition data of different types of artifact samples, we chose to use the average method to summarize the overall characteristics of the samples under the same classification, and through the form of tables and histograms to visualize the statistical laws. At the same time, based on the study of the statistical law, the chemical composition content of each cultural relic sample was weighted and averaged, and the ratio of the weighted average to the general average was obtained by the proportion method, based on which the prediction of the chemical composition content before weathering was established by Python.

References

[1] Takuya Iwanaga, William Usher, and Jonathan Herman. Toward SALib 2.0: Ad- vancing the accessibility and interpretability of global sensitivity analyses. Socio- Environmental Systems Modelling, 4:18155, May 2022.

[2] Jon Herman and Will Usher. SALib: An open-source python library for sensitivity analysis. The Journal of Open Source Software, 2(9), jan 2017.

[3] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blon- del, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

[4] Organizing Committee of Gaojiao Society. 2022 National University Student Mathematical Modeling Competition. The National Mathematical Modeling Competition for University Students, 2022.

[5] Wang Jie; Li Mo; Ma Qinglin; Zhang Zhiguo; Zhang Meifang; Wang Julin;. Weathering of an octagonal lead-barium glass vessel from the Warring States period. Institute of Metallurgy and Materials History, University of Science and Technology Beijing; Beijing Key Laboratory of Electrochemical Processes and Technologies, Beijing University of Chemical Technology; Chinese Academy of Cultural Heritage, 2014.

[6] McHugh, Mary L. "The chi-square test of independence." Biochemia medica 23.2 (2013): 143-149.

[7] Lewis, Don, and Cletus J. Burke. "The use and misuse of the chi-square test." Psychological bulletin 46.6 (1949): 433.

[8] Franke, Todd Michael, Timothy Ho, and Christina A. Christie. "The chi-square test: Often used and more often misinterpreted." American journal of evaluation 33.3 (2012): 448-458.

[9] (2008). Chi-square Distribution. In: The Concise Encyclopedia of Statistics. Springer, New York, NY. https://doi.org/10.1007/978-0-387-32833-1_54

[10] Mahmoud, Essam. "Accuracy in forecasting: A survey." Journal of forecasting 3.2 (1984): 139-159.