ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

Compositional analysis and identification of ancient glass products using clustering algorithm

Xinyao Zhang, Weijun He

Ocean College, Jiangsu University of Science and Technology, Zhenjiang, China, 212114

Abstract: *Ancient glass is extremely susceptible to weathering due to the environment in which it was buried. In the process of weathering, there is a massive exchange of interior elements with those of the environment, resulting in the compositional proportions of the glass being affected, thereby interfering with the accurate judgment of its category. In this paper, aiming at the problem of ancient glass products, using contingency table chi-square test, linear binary classification support vector machine, K-means++ clustering algorithm and other methods, the composition analysis and identification model of glass is established, and the problem of classifying unknown samples is solved, and did a model test and sensitivity analysis.*

Keywords: *Contingency table chi-square test, support vector machine, K-means++*

1. Introduction

The Silk Road was an important channel for cultural exchanges between China and the West in ancient times, and the precious evidence of early trade exchanges is glass. In the early days, glass was often made into bead-shaped ornaments from West Asia and Egypt and introduced to my country. After absorbing its technology in ancient my country, glass was made from local materials. Therefore, although it looks similar to foreign glass products, its chemical composition is different.

Ancient glass is extremely susceptible to weathering due to the environment in which it was buried. In the process of weathering, there is a massive exchange of interior elements with those of the environment, resulting in the compositional proportions of the glass being affected, thereby interfering with the accurate judgment of its category. For example, some cultural relics are marked with no weathering on the surface, and the decoration and color of the cultural relics can be clearly seen on the surface, but it is possible to have relatively light weathering locally; while some cultural relics are marked as surface weathered, with a large area of gray on the surface. The obviously weathered areas in yellow are the regolith, and the purple parts are the general weathered surface. In partially weathered artifacts, there are also unweathered areas on the surface.

In this paper, aiming at the problem of ancient glass products, using contingency table chi-square test, linear binary classification support vector machine, K-means++ clustering algorithm and Fisher discriminant and other methods, the composition analysis and identification model of glass is established to solve the problem of unknown samples. Classification and other issues, and did a model test and sensitivity analysis.

2. Model assumptions

Assumption 1: The blank data value of the proportion of chemical composition in the attachment is 0, that is, the sample does not have this chemical composition;

Assumption 2: The sample with blank color in the attachment thinks that the weathering is too serious and the color of the glass cannot be seen;

Assumption 3: The valid data provided are true and reliable, and the testing methods are standard;

Assumption 4: The chemical composition of the weathering process will not be artificially affected;

Assumption 5: The color part of the attachment is the color of the glass rather than the color of the weathered surface;

Assumption 6: The effects of weathering are irreversible.

ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

3. Model building and solution

3.1 Relational Analysis Model Based on Contingency Table Chi-Square Test

The contingency table can be used to study the independence between two variables [1]. This paper uses a two-dimensional contingency table to study the relationship between the surface weathering of glass cultural relics and the glass type, decoration, and color. For example, the surface weathering and decoration make the frequency contingency table as follows:

Table 1: Frequency contingency table of surface weathering and ornamentation

Weather or not / Ornamentation	Weathering	No weathering	Total
A	n_{11}	n_{12}	$n_{\cdot 1}$
B	n_{21}	n_{22}	$n_{\cdot 2}$
C	n_{31}	n_{32}	$n_{\cdot 3}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	n

In Table 1 n_{ij} is the frequency value that satisfies the i-th row variable and the j-th column variable at the same time.

The chi-square test was performed on the basis of the three contingency tables established above, and the above-mentioned analysis of the relationship between surface weathering and ornamentation was taken as an example.

First put forward the null hypothesis H_0: the independence between ornamentation and surface differentiation;

Then H_1: Decoration and surface differentiation are not independent.

The statistics for the test are calculated as follows:

$$\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{2}\frac{(n_{ij} - n_{i\cdot} \cdot n_{\cdot j} / n)^2}{n_{i\cdot} \cdot n_{\cdot j} / n} \quad (1)$$

The degree of freedom is the product of the number of row variables and column variables minus 1 respectively. The degree of freedom here is 2. For a given significance level:

$$\alpha = 0.05 \quad (2)$$

Then the critical value is obtained by looking up the table:

$$\chi^2_{0.05} = 5.992 \quad (3)$$

If $\chi^2 \geq \chi^2_{0.05}$, reject H_0. If there is no independent relationship between ornamentation and surface differentiation, then H_0 is accepted, they are considered independent [2].

If the two are related, this paper can use the Pearson series connection number to study the degree of correlation between the two categorical variables:

$$r = \sqrt{\frac{\chi^2}{n+\chi^2}} \quad (4)$$

It is worth noting that the contingency coefficient is between 0 and 1. If it is 0, there is no correlation between them. The larger the correlation coefficient is, the stronger the correlation is.

3.2 Solution of chi-square test relation analysis model

Taking the surface weathering and glass type of glass cultural relics as an example to make two-dimensional contingency Table 2.

ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

Table 2: Two dimensional contingency table of surface weathering and glass type

Type \ Weather or not	Weathering	No weathering	Total
High potassium glass	6	12	18
Lead barium glass	28	12	40
Total	34	24	58

It can be seen that the surface of most high potassium glass relics is weathered, and the surface of most lead barium glass relics is not weathered.

The other two two-dimensional contingency tables of surface weathering and glass decoration and color are shown in Appendix 1, from which it can be concluded that the number of samples with and without weathering on the surface of the glass with A and C decoration is not significantly different, but all the six glass cultural relics samples with B decoration are completely weathered, which does not rule out the contingency of the small sample size, so it can be inferred that there is a certain relationship between glass decoration and surface weathering, but it is not very obvious, Among them, the glass surface with pattern B is likely to be weathered.

However, according to the two-dimensional contingency table of the glass color and surface weathering, except that the four samples with no color can be seen are all glass relics with weathering on the surface, the other colors are not obviously related to the surface weathering. Therefore, it is believed that the glass color has little influence on surface weathering.

Using the descriptive statistics function of the software SPSS, select the cross table to get the following results:

Table 3: Data Results of Relationship between Surface Weathering and Texture, Type and Color

Variable	Test statistics	Freedom	Boundary value	Progressive significance	Contingency coefficient
Ornamentation	4.957	2	5.992	0.084	0.281
Type	6.880	1	3.843	0.009	0.326
Colour	9.432	8	15.507	0.307	0.374

According to the results in Table 3, the gradual significance of glass color is large, so it can be verified that it has little relationship with surface weathering, while the glass type has the least gradual significance, so it has a great relationship with surface weathering. From this, it can be concluded that the glass type has the largest relationship with the surface weathering and the color has the smallest relationship with the three variables.

3.3 Establishment and solution of statistical law analysis model based on run test

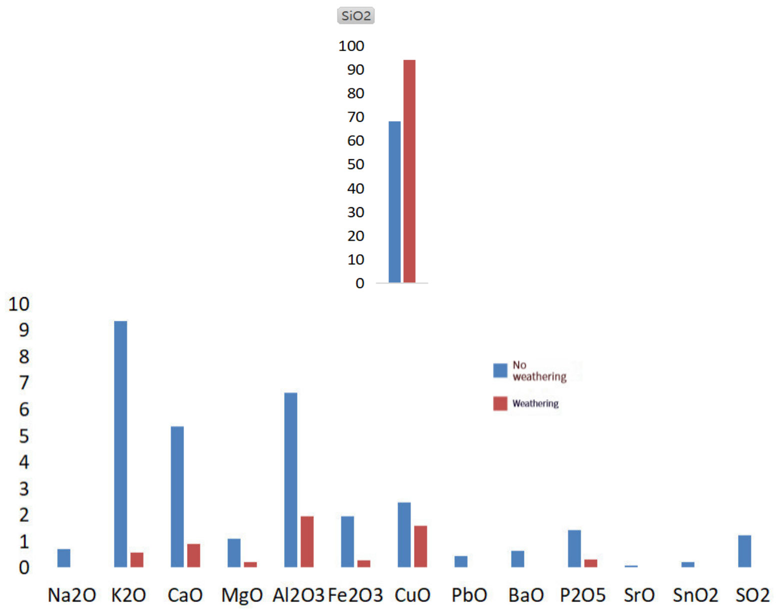

Figure 1: Statistics of average proportion of each component of high potassium glass

ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

Considering the composition of different types of composite glass, the following two figures are made according to the data statistics. They are respectively the proportions of different components of the surface weathered and non-weathered cultural relics of high potassium glass and lead barium glass. Figure 1 is the data statistics chart of high potassium glass, Figure 2 is the data statistics chart of lead barium glass. The abscissa is the main component, and the ordinate is the average proportion of each component of the surface weathered and non-weathered cultural relics. Among them, the data of cultural relics sampling points that are seriously weathered are regarded as severely weathered, and the data of non-weathered points are regarded as non-weathered no matter whether the cultural relics are weathered or not.

In the high potassium glass cultural relics, it can be seen that no matter whether there is weathering on the surface of the cultural relics, the main component is silicon dioxide (SiO_2), but SiO_2 itself is not easy to be oxidized, and the weathering process is also an oxidation process to some extent, so the proportion of SiO_2 in the surface weathered glass cultural relics is much higher than that in the non weathered glass cultural relics. This paper speculates that it is due to the exchange of elements between the internal and external environment. It can be understood that the weathered cultural relics are in the state before the weathering cultural relics. The components of weathered glass cultural relics are potassium oxide (K_2O), calcium oxide (CaO), aluminum oxide (Al_2O_3) and other components that are greatly reduced. It can be speculated that the proportion of SiO_2 in the weathered glass cultural relics is greatly increased due to some chemical reaction or exchange with some elements in the external environment.

Through the search of data, it is known that the content of some elements on the surface of the glass will increase and other elements will lose during weathering. The weathering surface of the high potassium glass on the relevant experimental surface is relatively non weathered. The element with higher weathering is Si, etc. The loss of K is very obvious, which can verify the accuracy of the above conjecture [3].

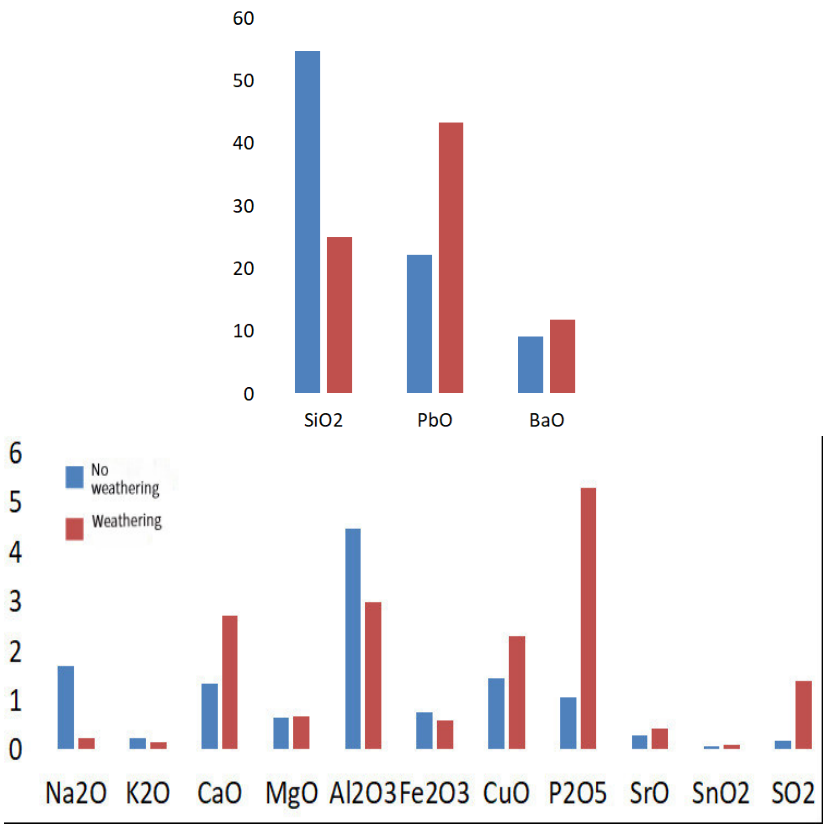

Figure 2: Statistical Chart of Average Proportion of Each Component of Lead Barium Glass

From the above statistical chart, it can be seen that in the lead barium glass relics, the more serious the weathering, the lower the proportion of SiO_2 components, and the higher the proportion of PbO and BaO components. It can be confirmed from the experimental detection results in the literature searched that the Si element is constantly lost in the weathering process of lead barium glass, and the Ba element accumulates in the outer layer of weathering to form a large amount of insoluble barium salts, so this process is increased, the PB element is also gradually increased in the weathering process, and many other chemical components also have certain changes, but most of the changes are in the dimension of a few

ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

percent, Therefore, the rule of its change cannot be determined. For example, the statistical data of CaO in the sample is that the proportion of weathered glass is lower than that of the sample of fresh glass relics. However, through the reference literature, it is found that the proportion of its content should increase with the weathering process. However, the difference of statistical data can be ignored in this dimension, so this paper does not think that these components have a rule on the surface weathering of lead barium glass [4].

Therefore, based on the statistical results of the two glass types, the following results can be obtained in Table 4.

Table 4: Qualitative Analysis on Statistical Law of Chemical Composition of Weathering on the Surface of Cultural Relics

Glass type	Regular chemical composition	Relative weathering free content change
High potassium glass	K_2O, MgO, Al_2O_3, Fe_2O_3, CaO	Reduce
	SiO_2	Increase
Lead barium glass	PbO, BaO	Increase
	SiO_2	Reduce

The run test is a test method to judge the number of runs formed by the arrangement and performance of sample marks, which is used to compare two independent samples and determine whether the observed results are random [5].

Take weathering and chemical composition SiO2 on the surface of high potassium glass relics as two variables to study whether they have randomness.

First put forward the original hypothesis H_0: two variables are random; H_1: The two variables have no randomness.

Then, the number of runs r, approximates to normal distribution. The proportion of SiO2 components in each high potassium glass cultural relic sample is sorted from small to large. With weathering as the test variable, this paper changes it to a numerical variable, that is, no weathering is recorded as 0, and weathering is recorded as 1. The number of runs is 2.

Then there are test statistics:

$$Z = \frac{r-E(r)}{\sigma_r} \quad (5)$$

The mean value is:

$$E(r) = \frac{2mn}{m+n} + 1 \quad (6)$$

m and n are the number of samples with and without weathering, respectively, and the other variance is:

$$\sigma_r = \sqrt{\frac{2mn(2mn-m-n)}{(m+n)^2(m+n-1)}} \quad (7)$$

If the sample is random, according to the classical probability, when R=2k is even, the probability distribution is:

$$P\{R = 2k\} = \frac{2C_{m-1}^{k-1}C_{n-1}^{k-1}}{C_{m+n}^{m}}, k = 1,2,\ldots,m; \quad (8)$$

When R=2k+1 is an odd number, its probability distribution is:

$$P\{R = 2k+1\} = \frac{C_{m-1}^{k-1}C_{n-1}^{k} + C_{m-1}^{k}C_{n-1}^{k-1}}{C_{m+n}^{m}}, k = 1,2,\ldots,m; \quad (9)$$

Finally, given the significance level α The critical value can be obtained to judge whether the original hypothesis is accepted.

ISSN 2616-5775 Vol. 6, Issue 3: 69-74, DOI: 10.25236/AJCIS.2023.060309

4. Conclusions

In this paper, qualitative and quantitative comprehensive analysis is used for discussion. Qualitative analysis is to visualize the data of the attachments and analyze the differences and changes of different chemical components between weathering and non-weathering based on the reference documents, title and other information. Then in quantitative analysis to make the above more accurate, this paper chooses run test to judge whether the observed values are random. It is found that the chemical components regularly weathered with high potassium glass mainly include SiO2, K2O, MgO, Al2O3, CaO, among which only SiO2 accounts for a higher proportion of surface weathering than non weathering; The chemical components regularly weathered with lead barium glass mainly include SiO2, PbO and BaO, among which only SiO2 accounts for a lower proportion of surface weathering than non-weathering.

References

[1] Lu Qingyun, Liu Hongxia. Analysis on the use of chi square test of contingency table in mathematical education research [J]. Statistics and Decision Making, 2008 (02): 156-158

[2] Sheng Sud, Xie Shiqian, Pan Chengyi. Probability Theory and Mathematical Statistics Beijing Higher Education Press 2019.12.

[3] Siqin Biliger, Li Qinghui, Gan Fuxi. Analysis of components of ancient Chinese potassium glasses by laser ablation inductively coupled plasma atomic emission spectrometry/mass spectrometry. Shanghai Institute of Optics and Precision Machinery, Chinese Academy of Sciences, Fudan University. Analytical Chemistry. 2013, 41 (09), 1328-1333

[4] Wang Jie, Li Mo, Ma Qinglin, Zhang Zhiguo, Zhang Meifang, Wang Julin. Study on the weathering of an octagonal lead barium glassware in the Warring States Period [J]. Glass and Enamel, 2014,42 (02): 6-13. DOI: 10.13588/j.cnki.g.e.1000-2871.2014.02.002

[5] Mathematical Cihai Committee. Mathematical Cihai (1-6). Volume 3 [M] Nanjing, 2002