

Compositional Analysis and Classification Study of Ancient Glassware Based on K-Means Cluster Analysis

Yunxiang Tan, Mengyu Xu, Yuanjun Zeng *

School of ecology and environment, Hainan University, Hainan, China, 570228

* Corresponding Author Email: illusorybutterfly28@gmail.com

Abstract. The aim of this study is to gain a deeper understanding of the compositional characteristics of ancient glass artefacts and their relationship with the degree of weathering as well as the rules for subclassification of ancient glass based on compositional variations. Through detailed compositional analyses of high-potassium glass and lead-barium glass, the study reveals their significant differences in terms of SiO_2 content, etc., and successfully divides them into subclasses by K-means clustering method, and the clustering results are verified by chi-square test, which shows its effectiveness in distinguishing the degree of weathering. The study not only provides a scientific basis for the classification of cultural relics, but also provides a methodology and preliminary conclusions for further excavation of information on ancient glass relics. This study is of practical significance for the conservation of cultural relics and in-depth understanding of ancient glass craftsmanship.

Keywords: Chi-square test, K-means cluster analysis, Antique glass.

1. Introduction

Under the influence of the Silk Road ^[1], China's ancient craftsmen drew on the glassmaking techniques of South Asia and Egypt to create glass products that were similar in appearance to foreign products but unique in chemical composition. The main raw material, quartz sand, contains silicon dioxide (SiO_2). In order to lower the melting point, they used fluxes such as grass ash, natural soda ash, saltpeter and lead ore, while adding limestone as a stabiliser, which was converted to calcium oxide (CaO) after calcination. Different fluxes lead to different chemical compositions of the glass ^[2,3]. For example, lead-barium glass, which uses lead ore as a flux and contains high levels of lead oxide (PbO) and barium oxide (BaO), is considered to be representative of China's indigenous glass, and was especially prevalent in the Chu culture ^[4]. Potassium glass, on the other hand, uses grass ash as a flux and is mainly found in Lingnan, Southeast Asia and India. The development of this process demonstrates that the ancient craftsmen of China created geographically unique glass products by using different fluxes according to local conditions. Therefore, the study of the chemical composition and identification and classification of glass products is of great practical significance for the protection of ancient glass artefacts.

Currently, scholars mainly use physical and chemical methods to identify the composition of ancient glass and study its evolution in depth ^[5-8], and only a few studies use statistical methods ^[9,10]. Based on this, this paper uses data related to ancient glass compositions, performs chi-square tests and correlation analyses, explores the relationship between surface weathering and its glass type, ornamentation and colour, and divides the subclasses according to the glass compositions by applying K-means clustering analysis. Through these studies, we aim to gain an in-depth understanding of the mechanism of glass weathering, provide a scientific basis for the conservation and restoration of ancient glass artefacts, and further expand our knowledge of ancient glass objects (data accessed at <http://www.mcm.edu.cn>).

2. Surface weathering of glass artefacts in relation to their glass type, decoration and colour

2.1. Pearson chi-square test

Pearson's chi-square test is widely used in the test of independence of categorical variables, which is a method of using the degree of deviation of the actual value of the sample data from the expected value to determine whether the original hypothesis is accepted. The degree of surface weathering, glass type, grain and colour in the data of this paper are all fixed class variables, and the sample size is greater than 40, so it is suitable for the chi-square test in the analysis of variance.

The Pearson chi-square test formula is:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

$$p = 1 - P(\chi^2 > \chi^2(0)) \quad (2)$$

Where, O_i is the actual frequency, E_i is the expected frequency, χ^2 is the chi-square value, $\chi^2(0)$ is the theoretical value, $P(\chi^2 > \chi^2(0))$ is the probability density function of chi-square distribution.

In the chi-square test, the larger the chi-square value, the larger the difference between the actual value and the expected value, i.e., the stronger the correlation between the two variables; the smaller the p-value, the more significant the test result, i.e., the stronger the correlation between the two variables.

This paper establishes the hypothesis:

H_0 : There is no correlation between the indicators studied and the degree of weathering of the sample (original hypothesis)

H_a : There is a correlation between the studied indicators and the degree of weathering of the samples (alternative hypothesis).

The data were subjected to Pearson chi-square test using spss and the results show that for the given level of significance $\alpha = 0.05$, colour and grain do not have a significant effect on the degree of weathering of the glass at the level of significance ($p=0.056$), which is in line with intuitive perceptions. However, the test result of glass type and degree of weathering showed a significant difference ($p=0.020$), rejecting the original hypothesis and supporting the alternative hypothesis that there is a correlation between glass type and weathering of glass. This suggests that the type of glass artefacts may be a key factor influencing the weathering of their surfaces, whereas colour and decoration may have a lesser influence on the degree of weathering.

2.2. List analysis

In order to deeply study the correlation between the surface weathering of ancient glass artifacts and their ornament types and colours, this study conducted a columnar analysis based on SPSS, and the results, as shown in Fig.1, show that when comparing the cases of no weathering and weathering, the correlation between the surface weathering and the ornament types and colours is weaker, with a smaller difference in the frequency counts, and therefore, it can be concluded that the correlation of the effect of these two on the surface weathering is less significant. However, in the same two cases, different glass types present significant differences. The relatively high percentage of lead-barium glass in surface weathering and the relatively low percentage of high-potassium glass suggest a correlation between surface weathering and glass type.

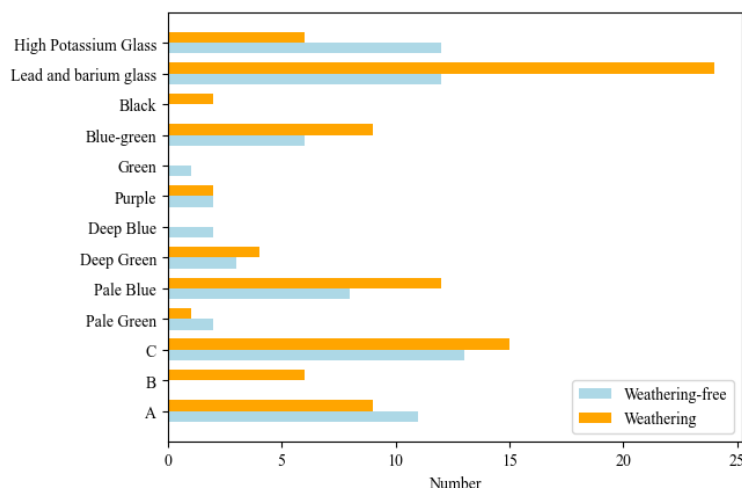


Fig. 1 Cross-plot of glass type, grain and colour versus degree of weathering

2.3. Statistical patterns of chemical composition of various types of glass before and after weathering

2.3.1 Descriptive statistics

In order to ensure the accuracy of the data, this paper processed all the data, converted the data into its proportion of the cumulative sum of all the components at the detection site, and calculated the mean, standard deviation, kurtosis, skewness, and coefficient of variation for the descriptive statistics of the four types of glass, from which we can speculate the changes of the proportions of the various chemical compositions in the glass before weathering and after weathering. The detailed statistical analyses contribute to a more comprehensive understanding of the dynamics of the chemical compositions of different types of glass during the weathering process, and provide in-depth insights into the weathering mechanisms of ancient glass.

Below are some of the descriptive statistics involved in the calculations:

(1) Kurtosis

Used to measure the steepness of the sample probability distribution, usually used to identify outliers (extreme values) in a given data set, the formula for kurtosis is shown in equation (3), this statistic needs to be compared with the normal distribution, kurtosis of 0 indicates that the overall data distribution is the same degree of steepness as that of normal distribution; kurtosis is greater than 0 indicates that the overall data distribution is steeper than that of the normal distribution, for the sharp peak; kurtosis is less than 0 indicates that Kurtosis is less than 0, which means that the overall data distribution is flatter than the normal distribution, with a flat peak. The greater the absolute value of kurtosis, the greater the difference between the steepness of the distribution pattern and the normal distribution.

$$K = \frac{1}{n} \sum_{i=1}^n \left[\left(\frac{X_i - \mu}{\sigma} \right)^4 \right] \quad (3)$$

Where μ is the mean and σ is the standard deviation.

(2) Skewness

Used to measure the asymmetry of the sample probability distribution, the same need to compare with the normal distribution, the formula shown in equation (4). Skewness of 0 indicates that its data distribution pattern has the same degree of skewness as the normal distribution; skewness greater than 0 indicates that its data distribution pattern is positively skewed or right-skewed compared to the normal distribution, i.e., there is a long tail dragging on the right, and the right end of the data has more extreme values; skewness less than 0 indicates that its data distribution pattern is negatively skewed or left-skewed compared to the normal distribution, i.e., there is a long tail dragging on the

left, and the left end of the data has more extreme values. The larger the absolute value of skewness, the more skewed the distribution is.

$$S = \frac{1}{n} \sum_{i=1}^n \left[\left(\frac{X_i - \mu}{\sigma} \right)^3 \right] \quad (4)$$

Where μ is the mean and σ is the standard deviation.

(3) Coefficient of variation

The coefficient of variation (CV) is the ratio of the standard deviation of the original data to the mean of the original data. The CV has no scale, which allows for objective comparisons of the degree of dispersion of different sets of data. The coefficient of variation eliminates the effect of differences in units and/or means on comparisons of the degree of variability of two or more sources.

$$CV = \frac{\sigma}{\mu} \times 100\% \quad (5)$$

Where μ is the mean and σ is the standard deviation.

Through the statistical analysis of the chemical composition contents of weathered high-potassium glass, non-weathered high-potassium glass, weathered lead-barium glass and non-weathered lead-barium glass, the following summaries can be drawn: in high-potassium glass, the content of SiO_2 is relatively high, but it shows a significant difference with a large coefficient of variation during the weathering process, which may reflect the inhomogeneous distribution of the minerals during the weathering process. The Na_2O and K_2O are relatively low, but K_2O shows significant differences in the weathering process. PbO and BaO are high in high potassium glasses, especially PbO has a large coefficient of variation, which may be related to its special formulation and weathering process. In lead-barium glasses, the contents of Al_2O_3 , Fe_2O_3 and CuO are relatively high, especially the coefficient of variation of CuO is larger, which may reflect the significant differences between different samples.

In the case of unweathered lead-barium glass, for example, there are four sets of peaks less than 0, indicating that the distribution is flatter at the top or coarser in the tails compared to the normal distribution, and there are 10 sets of data with peaks greater than 0, indicating that the distribution is sharper at the top or finer in the tails compared to the normal distribution; the skewnesses of SiO_2 , Al_2O_3 and P_2O_5 are close to 0, indicating that the distributions of the data in these sets are almost symmetrical, and the skewness of the rest of the data is greater than 0, indicating that the data fall to the right of the mean on the side of the skewness.

2.3.2 Box diagram

In order to pursue a more intuitive display, the use of matlab on the data grouped in a box-and-line diagram plotted in Fig.2, through the box-and-line diagram can be intuitively seen in the data distribution of the scattering range, the centre position and the data in the anomalies (outliers), to understand the state of the data distribution. In order to make the results more intuitive, in the process of drawing, the mean and standard deviation were selected, the results are more clear, and the comparison effect is more obvious; at the same time, taking into account that the weathering of cultural relics is a process of thousands of years, in which the weathering changes are very small, and the content before and after the weathering is less than 0.5% is counted as an impurity before and after the next step in the precise analysis, therefore, we take into account that only the chemical composition is less than 0.5% to be Therefore, we consider that only chemical compositions below 0.5% have a reference value, and some of the chemical compositions with very small proportions are not analysed in the box plots.

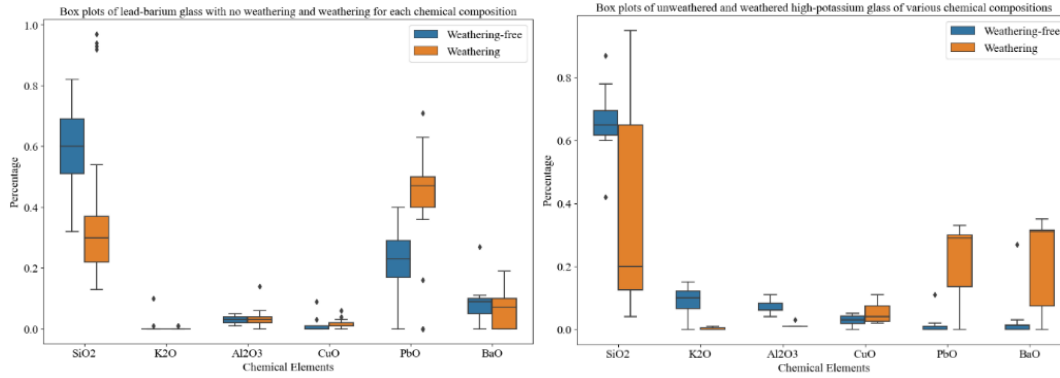


Fig. 2 Two types of glass box line diagram

As can be seen from Fig.2, for lead-barium glass, the percentage of weathering is significantly reduced, and the percentage is significantly higher, except for the other components in which the percentage is reduced more significantly, and the percentage is increased more significantly, and Na₂O, K₂O, Fe₂O₃, SnO₂ and SO₂ and SO₂ are all decreased.

For high potassium glass, in the weathering process, the same is reduced significantly, the difference between the lead-barium glass is the reduction of the magnitude of the second, Na₂O. CaO. MgO. Al₂O₃. Fe₂O₃ components accounted for more or less have decreased.

From this, it can be presumed that SiO₂, Na₂O, K₂O, Fe₂O₃ will be lost in the weathering process.

3. Classification of ancient glass subclasses

3.1. Selection and construction of clustering models

If the subclass division is performed manually, there will be a large deviation due to the subjective difference of each person, and 14 different components If the interval is divided manually, it will be time-consuming and labour-intensive, and the final result will be less than ideal due to the subjective reasons of each person. K-means clustering as a kind of unsupervised clustering, with supervised clustering random forests can be more effective in classifying the categories [11]. Therefore, this paper adopts K-means clustering analysis, through the input of hyperparameters k kinds of cluster classes for the division of subclasses.

$$d(B_p, B_q) = \sqrt[2]{\sum_{i \in e1} (J_{pi} - J_{qi})^2} \quad (6)$$

Based on the above obtained main chemical composition of high potassium glass and lead-barium glass using K-means clustering (Euclidean distance) clustering of high potassium glass and lead-barium glass clustering as well as glass with or without weathering, respectively.

$$d(J_p, J_q) = \sqrt[2]{\sum_{i \in e1} (J_{pi} - J_{qi})^2} \quad (7)$$

The steps involved in classifying high-potassium glass (lead-barium glass) are described.

a) Define the Euclidean distance in terms of the number of main chemical components of high-potassium glass (lead-barium glass) - a multidimensional space;

b) Calculate the distance between two samples of high-potassium glass (lead-barium glass):

where e_1 is the set of main chemical composition numbers of high-potassium glass, $p, q \in r_1$, r_1 is the set of high-potassium glass labelling numbers and $p \neq q$, r_2 is the set of lead-barium glass labelling numbers and $p \neq q$.

The basic steps of K-means clustering are shown as follows: 1) randomly select a sample as the first clustering centre; 2) calculate the distance from each sample to the selected clustering centre, $D(X)$, the larger $D(X)$, the greater the probability that it will be selected as a clustering centre; 3) select the next clustering centre by means of roulette wheel method (the larger the $D(X)$ the higher

the probability of it being selected as a clustering centre); 4) repeat step 2 until k clustering centres have been selected; 5) after the selection of the k clustering centres, that is to say, the initial points can be selected and then clustering can be carried out by using the standard K-means algorithm.

After preliminary analysis and screening, finally high potassium glass was divided into two subclasses by SPSS using K-means clustering method; while lead-barium glass was divided into three subclasses^[12]. Because this cluster analysis involves more variables, each component can be regarded as a variable, so in this cluster analysis we use principal component analysis to reduce the dimensionality, to get two principal components and then clustering; in which the classification of the results of the lead-barium glass cluster analysis is visualised in Fig.3.

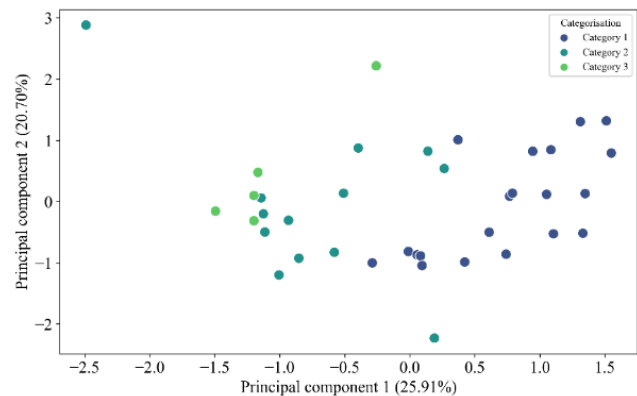


Fig. 3 Distribution of lead and barium glass scattering points.

It can be seen that in the classification of visualisation, this cluster analysis can distinguish the differences between the subclasses more significantly, the analytical items listed in Table 1 and Table 2 below are the analytical items that have a significant effect on this clustering, of which there are four components of high-potassium glass and lead-barium glass shared by the significant analytical items, respectively, SiO₂, PbO, BaO, P₂O₅, which are good indicators of the effect of the next step in the identification of the category of unknown points.

Table 1 Lead and barium glass subclassification interval table

	Clustering categories (mean ± standard deviation)			F	P
	Category 1(n=23)	Category 3(n=10)	Category 2(n=5)		
SiO ₂	0.277±0.074	0.618±0.084	0.916±0.058	183.012	0.000***
CaO	0.029±0.019	0.01±0.008	0.008±0.007	6.915	0.003***
PbO	0.483±0.097	0.211±0.06	0.0±0.0	88.788	0.000***
BaO	0.092±0.061	0.075±0.029	0.0±0.0	6.685	0.003***
P ₂ O ₅	0.049±0.047	0.006±0.006	0.005±0.005	6.042	0.006***
SrO	0.005±0.003	0.002±0.003	0.0±0.0	5.861	0.006***

Table 2 High-potassium glass subclassification interval table

	Clustering categories (mean ± standard deviation)		F	P
	Category 1(n=13)	Category 2(n=6)		
SiO ₂	0.722±0.125	0.21±0.156	59.025	0.000***
K ₂ O	0.082±0.05	0.002±0.003	14.384	0.001***
CuO	0.026±0.015	0.064±0.033	12.708	0.002***
PbO	0.004±0.006	0.264±0.081	143.864	0.000***
BaO	0.006±0.01	0.287±0.071	207.404	0.000***
P ₂ O ₅	0.012±0.015	0.061±0.025	29.177	0.000***

Note: ***, **, * represent 1 per cent, 5 per cent and 10 per cent significance levels, respectively.

Category 1 and 2 is classified as weathering mapping, for high-potassium glass: clustering to get two classes; for lead-barium glass: clustering to get three classes, and the actual weathering degree of agreement is shown in Fig.4.

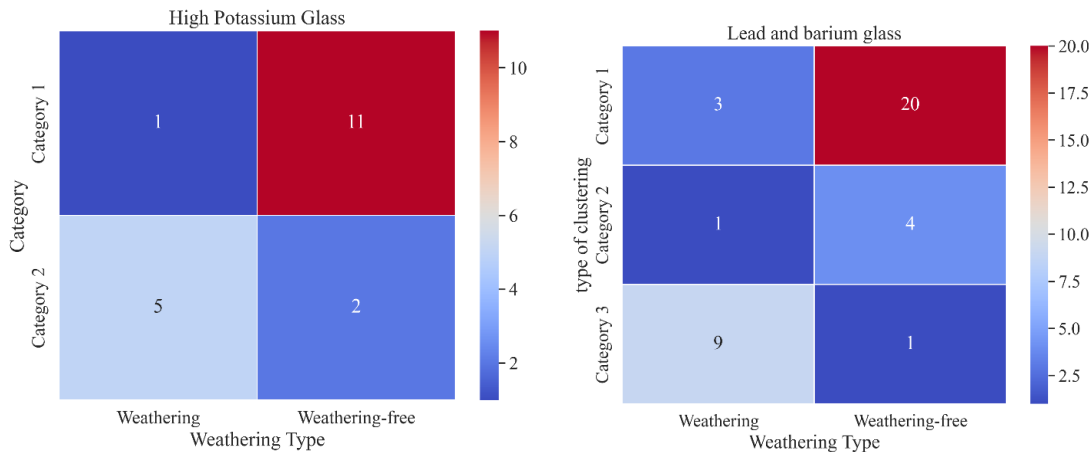


Fig. 4 Heat map of the chi-square test for high-potassium and lead-barium glasses

Using the Q-clustering results, the subclasses and whether weathering chi-square test, to get the test heat map shown in Fig. 4, as presented in the heat map, statistical analysis of the results can be obtained lead-barium glass category 1 weathering accounted for 87%, no weathering accounted for 13%, will be defined as weathering; lead-barium glass category 2 weathering accounted for 80%, no weathering accounted for 20%, will be defined as weathering; lead-barium glass category 3 weathering accounted for 10% and no weathering accounted for 90%, defining it as no weathering; high potassium glass category 1 accounted for 7.7% and no weathering accounted for 92.3%, defining it as no weathering; high potassium glass category 2 accounted for 83% and no weathering accounted for 17% as weathering. The clustering results are not significantly different from the actual degree of weathering.

The classification results can be obtained by using the above equations (6) and (7), where J_{qi} and B_{qi} are the values of the clustering centre, and J_{pi} and B_{pi} are the values to be predicted for the glass samples to be sampled, based on the distance of the samples from the classification centre in the six-dimensional space, and the results are shown in Table 3:

Table. 3 K-means clustering subclassification results.

Glass type	Subclassification	Clustering results	Actual result
High Potassium Glass	Category 1 (unweathered)	1. 3. 4. 5. 6. 7. 9. 13. 15. 16. 17. 18. 22	1. 3. 4. 5. 6. 7. 9. 16. 17. 18. 21. 22
	Category 2 (weathering)	10. 12. 14. 21. 27	10. 12. 13. 14. 15. 27
Lead and barium glass.	Category 1 (weathering)	2. 20. 24. 26. 28. 32. 34. 36. 37. 38. 39. 40. 41. 46. 47. 48. 49. 50. 51. 52. 54. 55. 56	2. 8. 11. 20. 23. 25. 32. 34. 36. 37. 38. 39. 40. 41. 45. 46. 47. 48. 49. 50. 51. 52. 54. 55. 56
	Category 2 (weathering)	8. 11. 19. 23. 25	
	Category 3 (unweathered)	29. 30. 31. 33. 35. 42. 43. 44. 45. 53	19. 24. 26. 28. 29. 30. 31. 33. 35. 42. 43. 44. 53

From Table 3, the accuracy of lead-barium glass without weathering is 69.23%, and the accuracy of weathering is 84%, and the accuracy of high potassium glass without weathering is 91.67%, and the accuracy of weathering is 71.92%. This cluster analysis has a high accuracy rate, the existence of the error may be related to the sample size is too small, whether the weathering eigenvalue is less clear.

3.2. Rationalisation

Through the F-test results and significance P-value test, the results of this clustering and division show quite a high level of significance, and the F-test results and significance P-value test results of the analysed items shared by the two types of glass (SiO, PbO, BaO, P₂O₅) are also highly significant.

This rationality evaluation also introduces the contour coefficient, DBI, and CH index within spss to evaluate the division results, and the results are shown in Table 4. The specific descriptions are as follows:

Contour coefficient: for a collection of samples, its contour coefficient is the average of the contour coefficients of all samples. The range of values of the contour coefficient is [-1,1], the closer the samples of the same category are to each other the further away the samples of different categories are from each other, the higher the score and the better the clustering effect.

DBI (Davies-bouldin): this indicator is used to measure the ratio of the intra-cluster distance followed by the inter-cluster distance of any two clusters. The smaller this metric is indicates the better the clustering effect.

CH (Calinski-Harbasz Score): the closeness of the class is measured by calculating the sum of the squares of the distances between the points within the class and the centre of the class (denominator), and the separation of the dataset is measured by calculating the sum of the squares of the distances between the centroids of the class and the centroids of the dataset (numerator), and CH metrics are obtained from the ratio of the separation and closeness, and the larger the CH, the better the clustering effect.

Table. 4 Evaluation of cluster analysis results

Glass type	Contour Coefficient	DBI	CH
High Potassium Glass	0.682	0.464	62.489
Lead and barium glass	0.582	0.505	92.348

The clustering effect of high potassium glass and lead-barium glass was assessed by three evaluation indexes, namely, profile coefficient, DBI and CH. As can be seen from Table 4, these metrics are within the appropriate range, indicating a better clustering effect. Specifically, high potassium glass exhibits a slightly higher profile coefficient, suggesting a higher degree of similarity within its sample and more significant differences between categories. Both DBI and CH metrics indicate relatively good clustering effects, with high potassium glass with a lower DBI suggesting a relatively higher degree of tightness within its clusters. Taken together, the chosen clustering algorithm presents a relatively good rationality for differentiating and classifying the samples of high-potassium glass and lead-barium glass.

3.3. Sensitivity analysis

In terms of data sensitivity, if a random data is adjusted by 5 per cent, basically one and only the boundary value will change the subcategory classification, so it is not sensitive to the change of the original data which is similar to the data used in the model, so it is reasonable for the subcategory classification.

4. Conclusion

This study provides a comprehensive taxonomic study by analysing the composition of ancient glass objects in depth, leading to a series of meaningful results. Firstly, it was found that colour and ornamentation did not have a significant effect on the degree of weathering of the glass, however, the glass types showed significant differences in surface weathering, with high-potassium glass in particular showing more significant differences relative to lead-barium glass. Secondly, analyses for high-potassium glasses showed significant differences in SiO₂ content, which may reflect the inhomogeneity of mineral distribution during weathering. Meanwhile, Na₂O and K₂O showed

significant changes in weathering, providing important clues for us to interpret the weathering process. For unweathered lead-barium glasses, the study showed that several components decreased significantly after weathering, especially Na_2O , K_2O , Fe_2O_3 , SnO_2 , and SO_2 . In the weathering process of high-potassium glasses, all components generally showed a decreasing trend, which provided key information about the loss of SiO_2 , Na_2O , K_2O , and Fe_2O_3 . Finally, the study used K-means cluster analysis to classify the subclasses of high-potassium glass and lead-barium glass, and two and three subclasses were obtained, respectively. The reasonableness of the clustering results was verified by the chi-square test, and the good classification effect was demonstrated by the accuracy analysis. Among them, the accuracy rate of lead-barium glass without weathering is 69.23%, and the accuracy rate of weathering is 84%; the accuracy rate of high potassium glass without weathering is 91.67%, and the accuracy rate of weathering is 71.92%. This study provides a preliminary method for the classification of ancient glass artefacts, but it still needs to face the challenges and opportunities. With the continuous progress of science and technology, the advanced means such as spectral analysis and imaging technology will provide a more accurate tool for the study, which will help to dig deeper into the information of glass artefacts. Therefore, future research should be committed to continuous innovation of data collection methods, continuous optimisation of modelling techniques, and the combination of emerging technological means, to create a broader prospect for the research and protection of ancient glass artefacts.

References

- [1] Journal of Archaeological Science: Reports. (2021), Did ancient glassware travel the Silk Road? X-ray fluorescence analysis of a Sasanian glass vessel from Okinoshima Island, Japan. Abe, Y.; Shikaku, R.; Murakushi, M.; Fukushima, M.; Nakai, I., 40,
- [2] Yiming, X.; Weimei, Z. Analysis and Identification of Ancient Glass Relics Based on Clustering Algorithm and Random Forest. In Proceedings of the 2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS), 12-14 May 2023, 2023; pp. 386-393.
- [3] Highlights in Science, Engineering and Technology. (2023), Prediction Model for Various Elements of Glass Artifacts Based on Probability Distribution. Qiu, H.; Zhu, W.; Zhu, S., 42:1-8,
- [4] Guo, Y.; Zhan, W.; Li, W. Application of Support Vector Machine Algorithm Incorporating Slime Mould Algorithm Strategy in Ancient Glass Classification. Applied Sciences 2023, 13, doi:10.3390/app13063718.
- [5] Journal of Archaeological Science: Reports. (2019), Chemical analyses and production technology of archaeological glass from Athienou-Malloura, Cyprus. Lin, Y.; Liu, T.; Toumazou, M.K.; Counts, D.B.; Kakoulli, I., 23:700-713,
- [6] Highlights in Science, Engineering and Technology. (2023), Glass weathering chemical composition prediction model. Liao, J., 34:443-449,
- [7] Highlights in Science, Engineering and Technology. (2022), Based on the analysis and identification of ancient glass products. Liu, S.; Zhang, Y., 21:212-221,
- [8] Sustainability. (2021), Spatial Layout and Coupling of Urban Cultural Relics: Analyzing Historical Sites and Commercial Facilities in District III of Shaoxing. Zhou, X.; Zhang, X.; Dai, Z.; Hermaputi, R.L.; Hua, C.; Li, Y., 13,
- [9] IEEE Sensors Journal. (2023), Component Correction and Category Identification for Weathered Silicate Glass with Small Datasets. Tang, S.; Yao, L., 23:23169-23178,
- [10] Molecules. (2023), Molecular-Composition Analysis of Glass Chemical Composition Based on Time-Series and Clustering Methods. Zou, Y., 28,
- [11] Nongye Jixie Xuebao/Transactions of the Chinese Society for Agricultural Machinery. (2022), Grading Detection Method of Grape Downy Mildew Based on K-means Clustering and Random Forest Algorithm. Li, C.; Li, Y.; Tan, H.; Wang, X.; Zhai, C., 53:225-236 and 324,
- [12] Journal of Advanced Transportation. (2021), Passenger Flow Prediction of Integrated Passenger Terminal Based on K-Means-GRNN. Tan, Y.; Liu, H.; Pu, Y.; Wu, X.; Jiao, Y.; Shriniwas, A., 2021:1-14,