WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

# Research on Surface Weathering Patterns of Glass Artifacts Based on Data Analysis and Machine Learning

Yonghe Peng [1, *], Junhe Yang [1], Rui Liu [2]

[1] School of Electromechanical Engineering, Guangdong University of technology, Guangzhou, China, 510006

[2] School of Information Engineering, Guangdong University of technology, Guangzhou, China, 510006

* Corresponding author: 3121000630@mail2.gdut.edu.cn

**Abstract.** With the deepening of cultural relics conservation and archaeological research, changes in the chemical and physical properties of ancient glass artifacts, especially the phenomenon of surface weathering, have become an important basis for the study of ancient civilizations and for guiding the authentication and preservation of cultural relics. To help reveal whether the glass surface is weathered or not and accurately classify the glass, this paper firstly preprocesses the sample point data of glass artifacts' surface weathering, establishes a multivariate linear model through the regression analysis method, and finally statistically identifies the weathering intervals of two types of glass. It provides a quantitative basis for the detection of the presence or absence of weathering on the glass surface. Further, using cluster analysis and similarity metrics, the glass samples were subclassified and quantitatively assessed for similarity. The research in this paper not only helps to deepen the understanding of the weathering phenomenon of ancient glass artifacts but also provides strong scientific support for the conservation and restoration of cultural relics).

**Keywords:** Regression analysis; Multiple linear modeling; Cluster analysis.

## 1. Introduction

With the deepening of heritage conservation and archaeological research, scientific studies on ancient glass artifacts have received increasing attention. As witnesses of history and carriers of culture, the changes in the chemical and physical properties of these precious relics have provided modern mankind with important clues revealing the life and technology of ancient civilizations, as well as their modes of communication. In particular, the surface weathering phenomenon of glass artifacts, as a direct reflection of the long-term natural environment and the influence of human factors, is closely related to the preservation status of artifacts and the identification of authenticity.

Although the importance of weathering research has been widely recognized, most of the traditional methods of weathering research rely on direct observation and description of the artifact surface, as well as simple chemical composition analysis [1]. Although these methods can provide some basic information, it is difficult to accurately quantify the weathering process and reveal its intrinsic mechanisms. With the advancement of science and technology and the rapid development of data analysis techniques, researchers and scholars must adopt more advanced and precise methods to conduct in-depth studies on the weathering phenomena of glass artifacts [2-7].

This paper aims to conduct a systematic study of the surface weathering phenomenon of ancient glass artifacts through data analysis and unsupervised cluster learning. Firstly, this paper will pre-process and quantitatively analyze the chemical composition data of the sampling points, and summarise the intervals where the weathering values are located when the surface of different glass types are weathered through regression analysis and other methods. This will provide direct data for accurately judging the preservation status and authenticity of cultural relics. Secondly, through unsupervised cluster learning, the glass samples are further sub-classified to reveal their intrinsic differences and similarities. This will contribute to a more in-depth understanding of the production techniques and historical background of glass artifacts.

This paper provides new perspectives and methods for the scientific study of ancient glass artifacts, as well as more scientific and effective guidance for heritage conservation and archaeological research.

## 2. Data Analysis of Glass Artifact Weathering

### 2.1. Sample point data pre-processing

In this paper, the sampling point is not detected by chemical composition caused by the original data blank missing part filled with 0 to complete. Then the proportion of chemical composition of each glass sample sampling point was accumulated, and it was considered that the cumulative sum of the sampling points in the 85% to 100% interval of the detection data was valid, and then screened out the data test is invalid for the elimination of processing. To eliminate the influence caused by the difference in the scale, the raw data were first processed without the scale, so that the mean value of the standardized variable was 0 and the variance was 1. The standardized data were obtained by the standardized processing method.

$$z = \frac{x_{ij} - \bar{x}}{\delta_x} \tag{1}$$

Where $\bar{x}$ represents the expected value of $x_{ij}$ and $\delta_x$ represents the standard deviation of $x_{ij}$ .

### 2.2. Chi-square test

The chi-square test can analyze variance by comparing a fixed category of variables with a fixed category of variables. The test method determines the size of the chi-square value through the degree of deviation between the actual observed value and the theoretical inferred value of the statistical sample, if the chi-square value is larger, the greater the degree of deviation between the two; otherwise, the smaller the deviation between the two. The category data of ancient glass products include the presence or absence of weathering on the glass surface, decoration, glass type, and color. In this paper, we set the variable X as the surface weathering, and the variable Y as the category data containing three kinds of decoration, type, and color, and through the interaction analysis, we get the results of the chi-square test, as shown in Table 1. According to the results of the chi-square test, for the surface weathering and glass type, the significant P-value is 0.01***, rejecting the original hypothesis, that there is a significant difference. Therefore, surface weathering can be used as a research object to determine the type of glass, providing a new solution for the analysis of ancient glassware.

**Table 1.** Surface weathering chi-square test

| | Name | Surface weathering | | $X^2$ | Calibrate | P |
|---|---|---|---|---|---|---|
| | | Yes | No | | | |
| Figure | C | 13 | 22 | 5.304 | 5.304 | 0.071* |
| | A | 14 | 14 | | | |
| | B | 0 | 6 | | | |
| Typology | High potassium | 14 | 6 | 11.271 | 9.516 | 0.001*** |
| | Lead and barium | 13 | 36 | | | |
| Color | Blue-green | 8 | 9 | 13.233 | 13.233 | 0.067* |
| | Light green | 8 | 16 | | | |
| | Violet | 2 | 4 | | | |
| | Dark green | 3 | 4 | | | |
| | Dark blue | 3 | 0 | | | |
| | Black | 0 | 8 | | | |
| | Light blue | 2 | 1 | | | |
| | Green | 1 | 0 | | | |

## 2.3. Regression analysis

In this paper, the results of binary regression analysis of glass surfaces with and without weathering in terms of chemical composition are obtained by the logistic regression analysis method, as shown in Table 2.

**Table 2.** Results of binary regression analyses

| | Regression coefficient | Standard error | OR | 95% confidence interval | |
|---|---|---|---|---|---|
| | | | | Limit | Lower limit |
| E | 7.006 | 32.437 | 1103.254 | 0 | 4.505 |
| $SiO_2$ | -0.067 | 0.329 | 0.935 | 0.491 | 1.781 |
| $Na_2O$ | -0.637 | 0.615 | 0.529 | 0.158 | 1.766 |
| $K_2O$ | -1.151 | 0.881 | 0.316 | 0.056 | 1.777 |
| $CaO$ | 0.130 | 0.760 | 1.139 | 0.257 | 5.053 |
| $MgO$ | -0.328 | 1.076 | 0720 | 0.088 | 5.931 |
| $Al_2O_3$ | -0.308 | 0.416 | 0.735 | 0.325 | 1.662 |
| $Fe_2O_3$ | -1.053 | 0.761 | 0.349 | 0.078 | 1.552 |
| $CuO$ | 0.598 | 0.383 | 0.349 | 0.078 | 1.552 |
| $PbO$ | 0.027 | 0.320 | 1.027 | 0.548 | 1.925 |
| $BaO$ | -0.348 | 0.482 | 0.706 | 0.274 | 1.817 |
| $P_2O_5$ | 0.232 | 0.542 | 1.261 | 0.436 | 3.646 |
| SrO | -2.348 | 2.686 | 0.096 | 0 | 18.489 |
| $SnO_2$ | 3.818 | 2.571 | 45.509 | 0.295 | 7020.725 |
| $SO_2$ | 0.311 | 0.278 | 1.364 | 0.791 | 2.352 |

Based on the regression coefficients in Table 2 the multivariate linear model can finally be derived as:

$$y = 7.006 + \sum_{i=1}^{14} a_i x_i \tag{2}$$

Where $a_i$ represents the regression coefficient of the ith chemical component and $x_i$ represents the content of the ith chemical component of the artifact. The weathering values of the chemical compositions of the two types of glass samples were calculated by the multivariate linear model, and the intervals in which the surface weathered and surface unweathered weathering values of the high-potassium and lead-barium glass samples were located were derived, as shown in Table 3.

**Table 3.** Distribution of weathering intervals for the two types of glass

| | No weathering of the surface | Surface weathering |
|---|---|---|
| High potassium glass | [-16.3, -1.096] | (-1.096, 2.229] |
| Lead and barium glass | [-5.487, 0.8287] | (0.8287, 8.416] |

According to the results in Table 3, the weathering value of lead-barium glass is greater than that of high-potassium glass. Therefore, it can be assumed that when the weathering value is in the interval of [-16.3, -1.096], the surface of high-potassium glass is not weathered; and when the weathering value is in the interval of (-1.096, 2.229], the surface of high-potassium glass is weathered. When the weathering value is in the interval of [-5.487, 0.8287], there is no weathering on the surface of lead-barium glass; when the weathering value is in the interval of (0.8287, 8.416], there is weathering on the surface of lead-barium glass.

## 3. Clustering Analysis and Similarity Quantification of Glass Samples

### 3.1. Similarity metric

In this paper, the Euclidean distance from the sample points of heritage weathering to the cluster centers is used as a measure to portray the proximity between the sample points and the cluster centers. The Euclidean distance, as a commonly used distance measure, can intuitively reflect the relative positional relationship of the sample points in the multidimensional space [8-10].

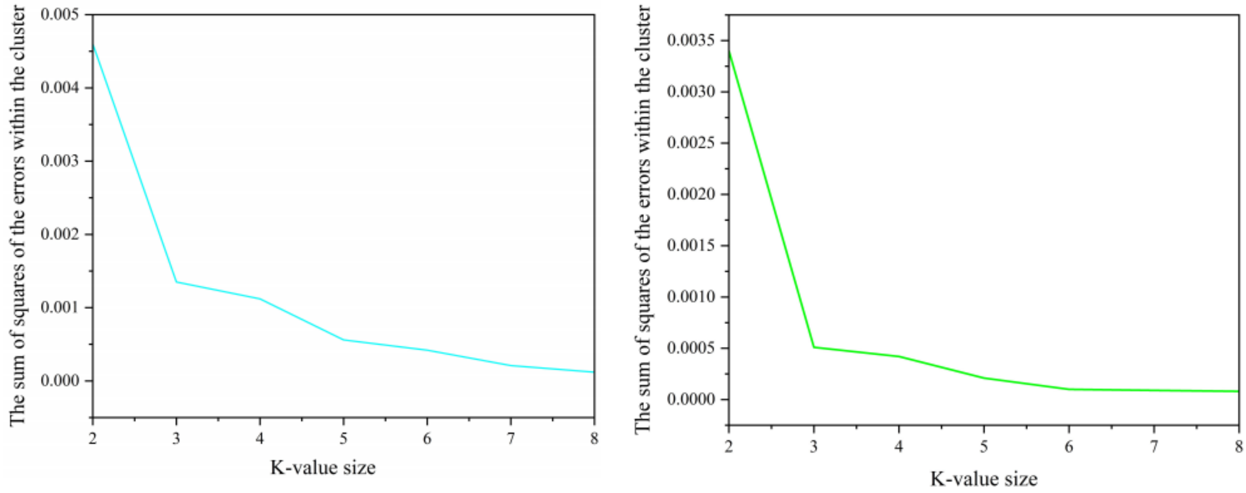
$$d(x_{ij}, x_k) = \sqrt{\sum_{i,j=1}^{14} (x_{ij} - x_k)^2} \tag{3}$$

Where $x_{ij}$ tabulates the sample data for the *j*-th chemical composition of the *i*-th glass and $x_k$ is the clustering point data.

In this paper, sample points that are close enough to each other are attributed to the same subclass, which enables effective glass classification with the help of glass weathering data. This process quantifies the sample points that are similar enough to more accurately describe the degree of association between them.

### 3.2. K-means cluster analysis

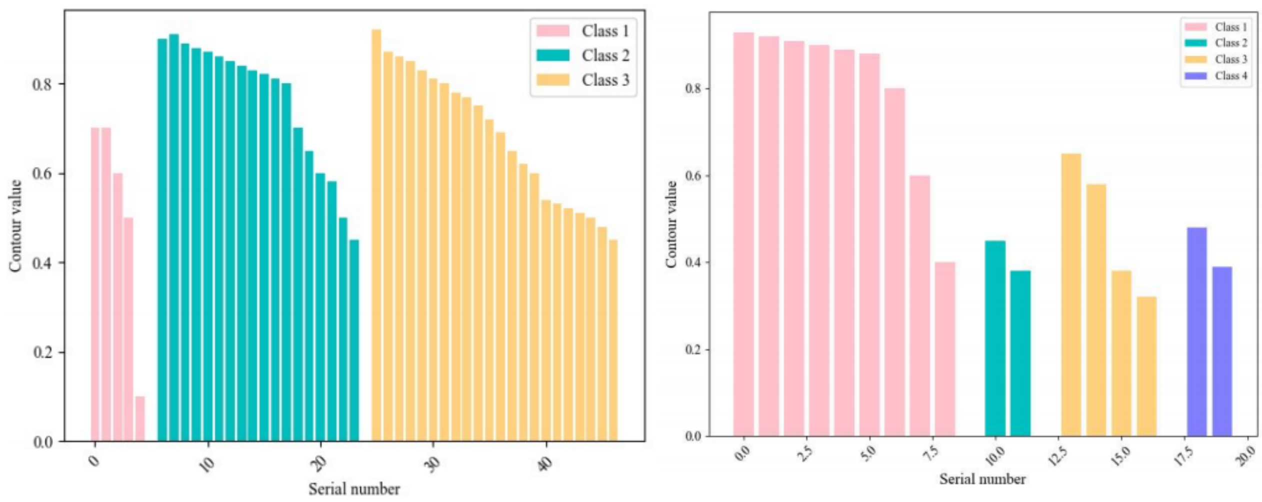
In cluster analysis, the selection of an appropriate number of cluster categories K is crucial. In this paper, the elbow method is used to constrain the choice of K. By observing the elbow diagram, the sum of squares of errors within clusters decreases with the increase of the K value, and there will be an obvious turning point at a certain point. The K value corresponding to this turning point is the ideal number of categories for subcategory continuation, as shown in Fig.1. In this study, for high potassium glass and lead barium glass, K=3 and K=4 is chosen as the ideal number of clustering categories for the two glasses, respectively.




(a) High potassium Glass (b) Lead and barium glass

**Figure 1.** Elbow diagram of two glass subclassifications

K-means clustering is performed by randomly selecting K sample point data as the initial clustering center, and then calculating the distances of other sample point data from these clustering centers. Through comparison and screening, the sample points with the smallest distance from the cluster center are classified as the category to which this cluster center belongs. The mean value of the coordinates of the sample points within each category is calculated as the new clustering center. This process is repeated and iterated until the category attribution of the sample points no longer changes. This iterative solving method allows the clustering results to be reproducible and to determine the category to which each sample point belongs. The subclassification contour values for high potassium glass and lead-barium glass are shown in Fig.2.




(a) High potassium Class (b) Lead and barium glass

**Figure 2.** Two glass subclassification profile value

According to Fig.2, when choosing K=3 and K=4 for clustering of high potassium glass and lead barium glass respectively, the contour value of the sample points of both types of glass is greater than 0, which indicates that the number of clustering categories chosen is reasonable. The results of the cluster analysis are valid.

## 4. Conclusion

According to the research of this paper, the weathering value of high potassium glass can be identified as no weathering on the surface at [-16.3, -1.096], and surface weathering at (-1.096,2.229]; while lead-barium glass has no weathering on the surface within [-5.487,0.8287], and surface weathering at (0.8287,8.416]. In the cluster analysis of subclassification of high-potassium glass and lead-barium glass, this paper found that high-potassium glass has the highest silica content in category 1, category 1, 3, and 4 glasses do not contain sodium oxide, and category 4 contains both lead oxide and barium oxide chemical compositions; whereas, lead-barium glass has the highest content of barium oxide and the lowest content of potassium oxide in category 1, the least amount of lead oxide in category 2, and the least amount of barium oxide in category 3, with the highest content of lead oxide.

This study not only reveals the quantitative criteria for the weathering characteristics of high-potassium glass and lead-barium glass but also analyses the distribution patterns of their chemical compositions in different subcategories. These findings provide a scientific and systematic research idea for the identification and analysis of ancient cultural relics and help us to judge the age, production process, and preservation status of cultural relics more accurately, thus promoting the further development of the field of cultural relics identification.

## References


[1] Liangzhi Zhou. Main factors affecting weathering of silicate glass [J]. Journal of Dalian Institute of Light Industry, 1984 (01): 34 - 44.

[2] Baoqian Wang, Jianjun Jiang. Identification of ancient glass types based on component data analysis and fuzzy pattern recognition [J]. Science and Technology Innovation and Application, 2024, 14 (07): 41 - 46.

[3] Mengting Zhang, Huiling You, Hetao Zhang, et al. Mathematical model for composition analysis and identification of glass relics [J]. Heilongjiang Science, 2024, 15 (04): 44 - 50.

[4] Hui Xie, Jianguo Zhang, Lei Guo, et al. Research on mathematical modeling of properties and chemical composition of ancient glass products [J]. Journal of Dezhou College, 2023, 39 (06): 6 - 14.

[5] Yang Lou, Lei Dou, Chaoyang Zhuo, et al. Research on ancient glass composition analysis and subclassification method [J]. Mathematical Modelling and its Applications,2023,12(04):73-83.

[6] Lei Yang, Tong Wei, Yiheng Zong. Mathematical modeling of compositional analysis and identification problems of ancient glassware [J]. Journal of Beijing Institute of Printing, 2023, 31 (09): 61 - 66.

[7] Han Dong, Minghua Zou, Lu Li, et al. Classification and prediction of ancient glass compositions [J]. Journal of Xianyang Normal College, 2023, 38 (04): 31 - 37.

[8] Huiting Huang, Chunming Li, Siyu Liu, et al. Compositional analysis and identification of ancient glassware based on compositional data [J]. Mathematical Modelling and its Applications, 2023, 12 (02): 52 - 62.

[9] Zhihao Chen, Jingmin Ji. Composition analysis and identification of ancient glassware based on hierarchical clustering model [J]. Modern Information Technology, 2023, 7 (08): 122 - 125.

[10] Shuafan Zhang, Ruijuan Sun, Hanzhuo Ren. Chemical composition analysis and correlation study of ancient glass products [J]. Chemical Management, 2023, (08):138 - 141.