Journal of Applied Mathematics and Computation, 2022, 6(4), 513-516
https://www.hillpublisher.com/journals/JAMC/
ISSN Online: 2576-0653
ISSN Print: 2576-0645

Analysis of Chemical Composition Content of Glass Artifacts Based on Clustering Algorithm

Xiaowen Fan

School of Mathematics and Science, Liaocheng University, Liaocheng, Shandong, 252000, China.

How to cite this paper: Xiaowen Fan. (2022) Analysis of Chemical Composition Content of Glass Artifacts Based on Clustering Algorithm. *Journal of Applied Mathematics and Computation*, 6(4), 513-516.
DOI: 10.26855/jamc.2022.12.014

Received: October 28, 2022
Accepted: November 25, 2022
Published: December 30, 2022

***Corresponding author:** Xiaowen Fan, School of Mathematics and Science, Liaocheng University, Liaocheng, Shandong, 252000, China.

Abstract

The development of civilizations over the centuries has left mankind with many valuable treasures, especially the Silk Road which carried the exchange of ideas and cultures between China and the West in ancient times, and ancient glassware was born in the process. This paper first analyses the correlation between whether the artefacts are weathered and the relationship between glass type, decoration and color to find the degree of correlation; then the variance of each component before weathering is used to find the substance with the most stable change as the eigenvalue, and the K-Means clustering algorithm is used to cluster the two categories of glass artefacts, and by exploring the partial least squares regression equation for each category, multiple By exploring the partial least squares regression equation for each class, the trend of change in chemical content before and after weathering was found; finally, the pre-weathering content was predicted by determining the class of the post-weathering eigenvalue data and substituting back into the regression equation for each class.

Keywords

Chemical content, correlation analysis, K-Means clustering algorithm

1. Background of the problem

With the gradual improvement of material life, the pursuit of spiritual civilization has gradually increased. The ancient glassworks that emerged during the Silk Road, which blended Western and Eastern aesthetics, are a valuable part of the historical and cultural heritage of human civilization. Due to the unique chemical nature of ancient glass objects, they are susceptible to weathering by environmental influences, which in turn leads to changes in the proportion of elements and chemical composition of the objects. By analyzing the information available on the artefacts and analyzing the relevant patterns, the chemical composition content before and after weathering can be obtained, which is of great importance for the subsequent conservation and restoration of the artefacts.

2. Data processing

This paper describes the number, decoration, type, color and surface weathering of 58 artefacts based on the basic information on glass artefacts provided by a website, but some artefacts have missing color information, so these artefacts are regarded as incomplete data. Ideally, the sum of the chemical composition of each artefact should be 100%, but due to errors caused by the means of testing and the complexity of the components being tested, artefacts with a sum of less than 85% and more than 105% of each component were considered invalid. In summary, after processing, valid data for 52 groups of artefacts were obtained and will be used in model building and solving where appropriate.

3. Modeling solution

3.1 Person test

The known data were statistically classified and the frequency of whether the glass objects were weathered or not was calculated separately and rearranged into a bar table of $r \times c$. Since only the relationship between weathering and classification information was discussed, only incomplete data were excluded. The results show that glass objects are independent of both decoration and color. Whether the glass artefacts are weathered or not is related to the type of glass and not to the decoration or color.The known data were statistically classified and the frequencies of whether the glass artefacts were weathered were calculated separately from the three sets of classified information, and each set of frequencies was rearranged into $r \times c$ the columnar table. As only the relationship between weathering and categorical information was discussed, only incomplete data were excluded. At this point it was assumed that there were independent events between the glass artefacts and the three pieces of information. The statistical value of

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(o_{ij} - e_{ij}\right)^2}{e_{ij}}$$

In order to obtain a more precise relationship between the glass artefacts and the three types of information, the significance test P-Value (later referred to as P) [1] was obtained using χ^2, considering P>0.05 as a strong correlation between information and artefacts, and otherwise a weak correlation between information and artefacts. The results of the Person χ^2 test showed that the significance test P was greater than 0.05 for glass artefacts and ornamentation, and for glass artefacts and color, so the original hypothesis that glass artefacts and ornamentation and color are both independent is accepted. However, the significance P value for glass artefacts and glass type is less than 0.05, rejecting the original hypothesis and leading to the conclusion that whether a glass artefact is weathered or not is related to glass type and not to decoration and color.

For high potassium glass artefacts, the total material content before weathering is between 97.25% and 100%. $c(SiO_2)$ Clearly the maximum, the $c(SO_2)$ minimum, and after weathering, the total substance content ranged between 99.81%~100%, with the exception of $c(SiO_2)$ there is a large increase, the rest of the material content is decreasing, especially$c(NaO_2), c(PbO), c(BaO), c(SrO), c(SnO_2), c(SO_2)$ has decreased to 0%.

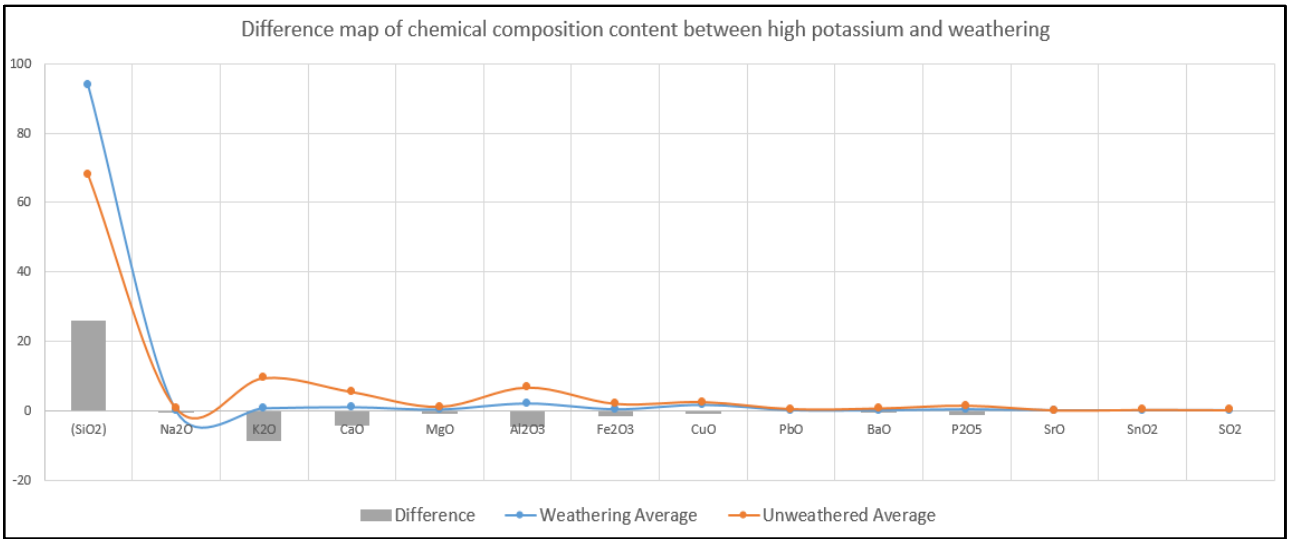

Fig. 1. Diagram showing the difference in chemical composition of high potash glass whether weathered or not [2].

For barium lead glass artefacts, before weathering, the total material content is between 88.41% and 99.98%. $c(SiO_2)$ Still the largest, the$c(PbO)$ followed by both with significantly greater content of material relative to the other, the $c(SnO_2)$ the smallest. After weathering, the total material content ranged from 90.17% to 99.89%. $c(SiO_2)$. There is a significant decrease in the $c(PbO)$ Significantly more and the most abundant material in the artefacts, with minor changes to the rest of the material, but almost no loss of material.

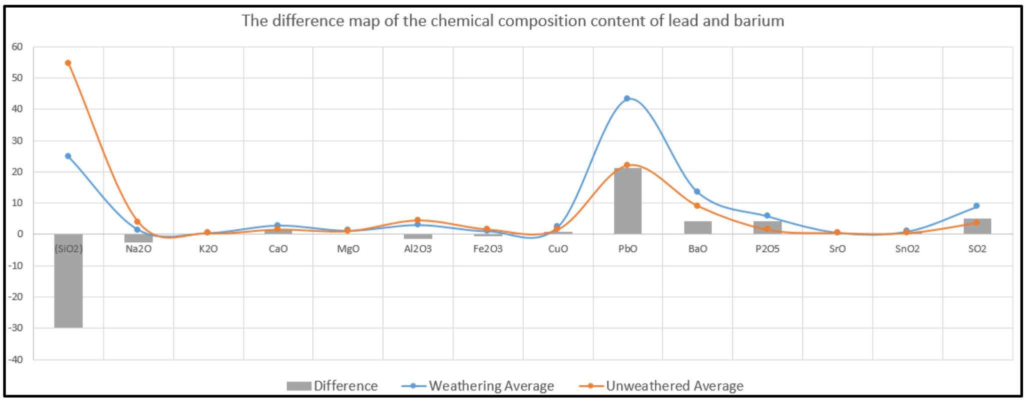

Figure 2. Difference in the chemical composition of lead and barium glass with and without weathering.

3.2 Prediction of chemical content

The data of high potassium glass and lead-barium glass without weathering were screened and the variance was found separately for each chemical content [3]: The k-values for the cluster analysis were determined as shown in Figure 3 below.

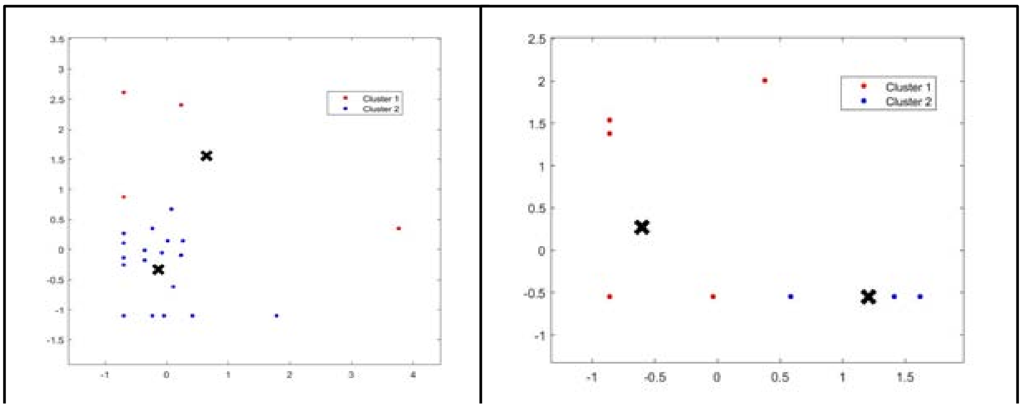

Figure 3. Clustering results for high potassium glass (left) and lead-barium glass (right).

For high potassium glass at k=2 and lead-barium glass at k=2, the degree of distortion of the SSE is greatly improved, with the elbow having the highest curvature for k values at this point and the best quality of clustering. Clustering of two classes of glass unweathered artefacts by a determined k-value. The Euclidean distance of each sample to each cluster center is calculated and each object is assigned to the cluster with the closest distance. In order to predict the chemical content before weathering, each of the categories obtained from the clustering above is analyzed to find the trend in the chemical content of each type of glass artefact when it is untethered, followed by partial least squares regression modelling [4] to derive the regression equation and make predictions. The weathered glass artefact data is then categorized to determine which of the above clusters it belongs to, and the weathered data is then added to the corresponding type of regression equation to obtain the predicted data. This paper uses$(\hat{y}_{ij}, y_{ij})$ as the coordinate value and plot the prediction for all sample points. $\hat{y}_{ij}$ is the first$j$ dependent variable indicator at the first i sample point (y_{ij}) of the predicted value to examine the model accuracy of the regression equation. Based on the predicted results, the sample points for both NaO and CaO are evenly distributed around the diagonal of the plot, then the fitted values for these two equations have very small differences from the original values and the equations are a good fit. In contrast, the PbO and BaO images do not make a diagonal line, so the fit of these two equations is not satisfactory.

4. Conclusion

This paper uses the Person test for variance analysis, and by establishing a correlation coefficient and significance analysis model, we determine the relationship between whether the glass artefacts are weathered and the three types of glass, decoration and color, respectively, according to the probability P-value; then we quantify the glass types: high potassium glass and lead-barium glass, based on which we establish a multiple regression model to find the statistical patterns between weathering and the content of several chemical components, including the range of chemical content before and after weathering, the size of the chemical content share, and the trend direction; finally, we predict the chemical content before weathering by means of partial least squares regression equations. The results were then quantified for the types of glass: high potassium glass and lead-barium glass, and then a multiple regression model was developed to find the statistical patterns between weathering and the content of several chemical components [5], including the range of chemical content before and after weathering, the size of the chemical content share, and the trend direction. The method proposed in this paper provides an effective method for the identification of glass types.

References

[1] Yin Pengda, Zhao Lina, Zhu Wenxu, Zhang Huihui, Jiao Yusheng, Zhao Guangwei, Sun Guangyu. Optimal fertilization of filled roasted tobacco based on partial least squares regression [J]. China Tobacco Science, 2011, 32(04):61-65.

[2] Wang Lei, Du Zhiqian, Wang Ling, Liu Ming, Wang Cong. Research on tillage depth measurement device and prediction analysis of multiple linear regression model [J]. Agricultural Mechanization Research, 2022, 45(06):139-145

[3] Li AY. Research on the prediction of particulate matter concentration in Urumqi based on RF-Kmeans-LIBSVM Environmental Science, 2022, 48(04):118-124.

[4] Cao Duanxi. The improvement of clustering algorithm and the study of clustering effectiveness index [D]. Nanjing University of Posts and Telecommunications, 2021.

[5] Long WJ, Zhang XF, Zhang Chain. A business process clustering method based on k-means and elbow rule [J]. Journal of Jianghan University (Natural Science Edition), 2020, 48(01):81-90.