Composition analysis and identification of ancient glass products

Wenyu Sun *, Xiaoyi Wang #

School of Statistics and Mathematics, Shandong University of Finance and Economics, Jinan, China

* Corresponding author: s15650593973@163.com
# These authors contributed equally.

Abstract. Based on the proportion of various components of ancient glass products and the classification method, the composition, change and relation of different types of cultural relics under different weathering conditions are analyzed in detail in this paper. The statistical rules and the basis for the classification of cultural relics are given, and on this basis, the unknown types of glass cultural relics are identified. Firstly, a statistical model is established, and Chi-square test is used to analyze the factors that affect the weathering of cultural relics. It is concluded that the weathering of cultural relics is related to the type of glass and the type of decoration, and the type of glass has a greater influence on the weathering. The chemical composition of the weathered cultural relics before weathering is predicted. Then, after the classification rule of glass is preliminarily explored, the grey correlation model is established to calculate the correlation degree between each chemical component and the two kinds of glass, and then the K-means clustering model is used to divide the glass into subclasses. Thirdly, BP neural network was used to analyze the chemical composition of glass relics of unknown category, identify their types, and analyze the sensitivity of classification results. Finally, according to the idea of discrete Frechet distance algorithm, the cultural relics are grouped according to the type of glass, and the discrete Frechet distance of the broken line of the change rate of the content of different chemical components between cultural relics is used as the basis to judge the correlation relationship, and the correlation relationship between the two substances is analyzed.

Keywords: Chi-square Test, BP Neural Network, Grey Correlation Model, Discrete Frechet Distance Algorithm.

1. Introduction

For more than half a century, Chinese cultural relics and archeological circles have analyzed and discussed the shape, pattern and texture of ancient glass relics in different regions and periods of China. It is believed that after the Hantong Western regions, ancient Chinese glass products and technologies were introduced from the West through the Silk Road. [1] Ancient Chinese glass has a unique style in the world glass history, and its chemical composition is different from that of foreign glass, which has become the concrete evidence of the cultural and technological exchanges between ancient China and foreign countries. The weathering degree of ancient glass is related to its burial environment, and the internal element exchange during the weathering process will change its composition ratio, thus affecting the judgment of the type of glass. The ancient glass marked with obvious color and pattern is not weathered on the surface, but it does not exclude the local shallow weathering; There are also unweathered areas on the surface of obviously weathered relics. The existing glass relics need to be analyzed according to the known information such as their chemical composition.
2. Statistical rules were used to analyze and predict the categories of cultural relics

2.1. The relationship between the surface weathering of glass relics and the type, color and ornamentation of glass was analyzed by Chi-square test

Chi-square test is the deviation degree between the actual observed value and the theoretical inferred value of the statistical sample. The deviation degree between the actual observed value and the theoretical inferred value determines the Chi-square value. The larger the chi-square value is, the greater the deviation degree between the two values is. On the contrary, the deviation between the two is smaller; If the two values are exactly equal, the chi-square value is 0, indicating that the theoretical values are in complete agreement [2].

The values in form 1 are assigned as follows. Figure 1 is the Chi-square test assignment table

<table>
<thead>
<tr>
<th>Weathering</th>
<th>After</th>
<th>Before</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glass</td>
<td>High potassium</td>
<td>Lead barium</td>
</tr>
<tr>
<td>Assignment</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Figure</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Assignment</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Colour</td>
<td>Black</td>
<td>Blue-green</td>
</tr>
<tr>
<td>Assignment</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 1. Chi-square test assignment table

Figure 2. shows a Chi-square test of glass type and weathering

<table>
<thead>
<tr>
<th>Chi-Square Tests</th>
<th>Value</th>
<th>df</th>
<th>Asymptotic Significance (2-sided)</th>
<th>Precision Significance (2-sided)</th>
<th>Precision Significance (1-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>3.861</td>
<td>1</td>
<td>0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuity Correction</td>
<td>2.758</td>
<td>1</td>
<td>0.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>3.853</td>
<td>1</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear</td>
<td>3.786</td>
<td>1</td>
<td>0.052</td>
<td></td>
<td>0.049</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.77.
b. Calculation only for 2x2 tables

Figure 2. Type of glass and chi square test for weathering

According to Figure 2, the correlation coefficient between glass type and weathering is 0.049, which is less than 0.05, so it can be considered that weathering of cultural relics is related to glass type. Among them, the weathering rate of high-potassium glass and lead-barium glass is calculated according to the requirements. It can be seen that the weathering rate of high-potassium glass is 33%, and that of lead-barium glass is 67%. Therefore, the relics made of lead-barium glass are more likely to be weathered than those made of high-potassium glass.
Figure 3. Chi-square test of color type and presence of weathering

As can be seen from Figure 3, the correlation coefficient between color type and weathering is 0.428, so it can be considered that weathering of cultural relics has nothing to do with color.

Figure 4. Chi-square test for type of ornament and whether it is weathered

As can be seen from Figure 4, the correlation coefficient between decorative types and weathering is 0.057, slightly greater than 0.05. It can be considered that weathering of cultural relics is related to ornamentation, but the relationship is weak. According to the data available, the weathering rate of pattern A is 0.45, that of pattern B is 1, and that of pattern C is 0.54. Therefore, B is the most susceptible to weathering, C is in the middle, and A is the least susceptible to weathering.

In conclusion, through the Chi-square test, this paper believes that the weathering of cultural relics is related to the type of glass and the type of decoration among the types, colors and patterns of glass. And the glass type is more likely to affect the weathering of cultural relics than the decorative type. The relics of lead-barium glass type are more easily weathered than those of high-potassium type. Among the decorative types, B is the most easily weathered and A is the least easily weathered.

2.2. The type of glass and its weathering condition and the statistical law of weathering chemical composition content

The mean ratio of each chemical composition before and after weathering was calculated, and the cultural relics were divided into four types: high-potassium glass weathered, lead-barium glass weathered, high-potassium glass unweathered, and lead-barium glass unweathered. The mean value of the proportion of each chemical component in each type was calculated respectively, and the statistical rules of weathering and weathering chemical component contents of different glass types and cultural relics were summarized.
Figure 5. Chemical composition range of high potassium glass before and after weathering

The analysis of Figure 5 shows that the content proportion of silica in the weathered high-potassium glass relics is higher than that before weathering, while the proportions of other chemical components all decrease or have little difference after weathering.

Figure 6. Chemical composition range of lead-barium glass before and after weathering

Figure 6 summarizes the statistical rules of weathering and unweathering chemical components of different glass types are summarized. After weathering of high-potassium glass, the content of silica increases, while the proportion of other components decreases slightly. After weathering, the contents of sodium oxide, potassium oxide, calcium oxide, magnesium oxide, copper oxide, lead oxide and phosphorus pentoxide increase, while other components decrease slightly.

2.3. Prediction of weathering weathering before content

(1) The change rate of the mean ratio of each chemical component of glass before and after weathering was calculated

For the four cultural relics separated in 2.2, $x_i$ is the mean chemical composition ratio before weathering, $x'_i$ is the mean chemical composition ratio after weathering, and $\beta_i$ is the change rate of the mean chemical composition ratio before and after weathering.

$$\beta_i = \frac{x'_i - x_i}{x_i}$$

When the mean value of chemical composition before and after weathering is 0, $\beta_i$ is denoted as 0.

(2) To predict the chemical composition of glass before weathering

For the glass labeled as weathered cultural relics, the contents of each chemical component $y_i$ before weathering were predicted by using the above change rate $\beta_i$ and the mean value of chemical component ratio $y'_i$ corresponding to the weathered cultural relics. The calculation is as follows:

$$y_i = \frac{y'_i}{1+\beta_i}$$

In particular, when $\beta_i$ is 0, the corresponding chemical content is also 0.

The prediction of chemical composition contents of the two kinds of glass labeled as weathered before weathering is shown in Figure 7 and 8.
3. Classification of glass types based on K-means

3.1. Establish grey correlation model and select clustering index

To select chemical composition as clustering index, it is necessary to know the correlation degree between each chemical composition and glass type. In this paper, using the grey correlation method of high potassium and lead barium glass and correlation of various kinds of chemicals. Grey correlation analysis is a very active branch of grey system theory. Its basic idea is to judge whether the connection between different sequences is close according to the geometric shape of sequence curves [3].

(1) Determine analysis sequence

The parent sequence was first determined. It is necessary to conduct quantitative analysis of high potassium glass relics and lead barium glass relics. The proportion values of high-potassium glass relics and lead-barium glass relics in all cultural relics were used as the analysis sequence of the two kinds of glass. There were 45 cultural relics after data cleaning, 16 of which were high-potassium glass relics, accounting for 0.36. There are 29 lead barium glass cultural relics, accounting for 0.64. That is, the parent sequence element value of high potassium glass is 0.36, and the sub-sequence element value of lead barium glass is 0.64.

The subsequence is composed of the ratio of each chemical component. Take high-potassium glass cultural relics as an example to continue the following operations. The parent sequence of high-potassium glass cultural relics is \( w_0 \) and the subsequence is \( w \).

(2) Preprocess the data

First find the mean of each index, and then divide each element of the index by its mean.
The processed parent sequence:

\[ w_0 = (w_0(1), w_0(2), \ldots, w_0(n))^T \]  \hspace{1cm} (3)

And the processed subsequence:

\[
\begin{aligned}
w_1 &= (w_1(1), w_1(2), \ldots, w_1(n))^T \\
w_2 &= (w_2(1), w_2(2), \ldots, w_2(n))^T \\
&\quad \vdots \\
w_m &= (w_m(1), w_m(2), \ldots, w_m(n))^T 
\end{aligned}
\]  \hspace{1cm} (4)

(3) Subsequence are calculated respectively and the two female sequence correlation

Denote \( a = \min \min |w_0(k) - w_i(k)| \) as the minimum difference between the two levels and \( b = \max \max |w_0(k) - w_i(k)| \) as the maximum difference between the two poles.

\[ y(w_0(k), w_i(k)) = \frac{a + \rho b}{|w_0(k) - w_i(k)| + \rho b} \quad (i = 1, 2, \ldots, m; k = 1, 2, \ldots, n) \]  \hspace{1cm} (5)

The grey correlation degree defined is shown in Equation (6):

\[ y(w_0, w_i) = \frac{1}{n} \sum_{k=1}^{n} y(w_0(k) - w_i(k)) \]  \hspace{1cm} (6)

3.2. Two glass to pick three strongest correlation chemical composition as a clustering index

It is necessary to cluster the inner parts of high-potassium glass and lead-barium glass according to their chemical composition. Firstly, an appropriate clustering algorithm should be selected. As the research data based on this paper is relatively simple, it is decided to adopt the K-means clustering algorithm based on partition. K-means clustering algorithm is an iteratively solved clustering analysis algorithm. Its step is to select K objects as the initial clustering center, then calculate the distance between each object and each clustering center, and assign each object to the nearest clustering center [4]. In Matlab simulation [5], the learning and classification of data samples are from the initial center distance to the recalculation of each type of center. The samples are gradually aggregated to various centers and clustered into appropriate classes according to data samples and actual needs until the algorithm converges to the optimal solution. Determine the number of clustering categories.

The high potassium glass and lead barium glass were grouped into 2 and 3 categories respectively. The clustering results were compared and the good clustering effect was selected as the number of clustering categories.
(1) K-means clustering was performed using SPSS. Parameter setting: set resolution coefficient $\rho = 0.5$.

![Figure 9](image1)

Figure 9. Clustering effect with different number of clusters

As can be seen from Figure 9, the clustering effect of high-potassium glass and lead-barium glass is smaller when they are clustered into three categories. And when the cluster is divided into two groups, the distance between some members and the cluster center is relatively far. Therefore, the effect of clustering into two categories is not as good as that of clustering into three categories. Finally, three categories are selected for clustering of both kinds of glass.

The clustering results refer to the sub-classification results of high-potassium glass and lead-barium glass, as shown in Figure 10 and 11.

![Figure 10](image2)

Figure 10. Two kinds of glass subgroups cluster center
Figure 11. Concrete results of two kinds of glass subclass division

(2) Model sensitivity analysis
On the premise of control variables, the value of clustering index is adjusted, and the classification changes after adjustment are shown in Figure 12 and 13. The sensitivity of the two kinds of glass to the change of chemical composition was analyzed through the classification change.

Figure 12. Changes in the subclass division of high potassium glass
The sections highlighted in black indicate that the subclass division has changed. It can be found that when the silica content of high-potassium glass cultural relics increases or decreases by 15%, the classification situation will change, while the original category remains unchanged when the alumina...
and copper oxide change. This indicates that the subclass division of high potassium glass is more sensitive to the change of silica content, but less sensitive to the change of alumina and copper oxide content.

Figure 13. Changes in the subclass division of high potassium glass

In Figure 13, the sections highlighted in black indicate that the subclass division has changed. It can be found that when the content of silica and lead oxide in lead-barium glass cultural relics increases or decreases by 15%, the classification situation will change, while the original category remains unchanged when the content of alumina changes. This indicates that the subclassification of lead-barium glass is more sensitive to the change of silica and lead oxide content, but less sensitive to the change of alumina content.

4. Qualitative analysis of unknown cultural relics based on BP neural network

The learning of neural network, also known as training, refers to the process of adjusting the free parameters of the neural network through the stimulus of the environment in which the neural network is located, so that the neural network can react to the external environment in a new way. The ability to learn from the environment and improve their own performance in learning is the most meaningful property of neural networks. Neural networks learn more about their environment over time [6]. BP (Back Propagation) neural network [7] is a kind of multilayer feedforward neural network. Its name
comes from the fact that the adjustment rule of network weight adopts the back propagation learning algorithm, namely BP learning algorithm.

The BP neural network algorithm is used to use 45 valid data according to the requirements, and 35 of them are randomly selected as the training set, and the remaining 10 data are used as the test set to train the neural network. The type of cultural relic can be identified by bringing the required qualitative cultural relic data into the neural network. BP parameters are set in Table 1 below.

<table>
<thead>
<tr>
<th>Table 1. BP parameter setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of enter layer nodes</td>
</tr>
<tr>
<td>Number of hidden layer nodes</td>
</tr>
<tr>
<td>Number of output layer nodes</td>
</tr>
<tr>
<td>Training number</td>
</tr>
<tr>
<td>Learning rate</td>
</tr>
<tr>
<td>Minimum error of training target</td>
</tr>
</tbody>
</table>

Using the set parameters, the accuracy of the results is higher than 90% after several tests on the test set, so the neural network is feasible. The identification results are shown in Figure 14. The number 1 represents high potassium glass and the number 2 represents lead barium glass.

![Figure 14. Identify the results](image)

The trend of mass data in power system provides a basis for load characteristic analysis and prediction model establishment, but the classical load forecasting method cannot afford such a huge time and computing resource consumption. The problem of over fitting in large sample set will affect the prediction accuracy. In this paper, a power load forecasting model is built by using the BP neural network model, making full use of the powerful data processing function of Clementine and preventing the over fitting function. The experimental results show that the BP neural network model has good predictability and robustness, and has a certain practical application value.

5. **Based on the qualitative analysis of cultural relics under Frechet**

   (1) Classify artifacts by type of glass.

   (2) In addition to the first cultural relic of each group, the rate of change formed by subtracting the chemical composition content of each cultural relic and the previous cultural relic and then dividing the chemical composition content of the above cultural relic is calculated respectively, and the matrix of the chemical composition change rate of cultural relic under different glass types is obtained.

   (3) The rate of change of each chemical component is wired to obtain the corresponding rate of change curve.

   (4) The Frechet distance, commonly known as the man-dog distance model, can be graphically explained in Figure 15. The Frechet distance is a basic measure of the similarity between two curves [8]. Frechet distance believes that in the sequence of two different sampling points, we should try to find a path that can make the distance sum of the values matched with each other minimum, and this path is the core path.
Let the discrete ordered point string $S_1$ and $S_2$ obtained after discretization of continuous curves $f$ and $g$ be $<S_{1,1}, S_{1,2}, \ldots, S_{1,n}>, <S_{2,1}, S_{2,2}, \ldots, S_{2,m}>$, and the discrete Frechet distance between $S_1$ and $S_2$ be $D_{df}$. \[9\]

The $D_{df}$ between $S_1$ and $S_2$ can be solved recursively by equation (7)

\[
D(S_1, S_2) = \max \begin{cases} d_d(S_{1,n}, S_{2,n}) \\ \min \begin{cases} <S_{1,1}, S_{1,2}, \ldots, S_{1,n-1}>, <S_{2,1}, S_{2,2}, \ldots, S_{2,m}>, \forall n \neq 1 \\ <S_{1,1}, S_{1,2}, \ldots, S_{1,n}>, <S_{2,1}, S_{2,2}, \ldots, S_{2,m-1}>, \forall m \neq 1 \\ <S_{1,1}, S_{1,2}, \ldots, S_{1,n-1}>, <S_{2,1}, S_{2,2}, \ldots, S_{2,m-1}>, \forall n \neq 1 \text{and} \forall m \neq 1 \end{cases} \end{cases}
\]

(7)

According to the derivation process of discrete Frechet distance calculation, the higher the similarity of two broken lines, the smaller the calculated curve distance [10]. Using this idea, the curve of the change rate of chemical composition was analyzed, and the distance was calculated to form the distance matrix. The discrete Frechet distance of the broken line of the change rate of the contents of different chemical components between cultural relics is used as the basis to judge the correlation. If the Frechet distance value is less than 1, it is considered that there is a strong correlation between the two chemical components. The correlation between chemical components was analyzed according to the above criteria.

According to the analysis of the results according to the reference value, it can be found that in the cultural relics with high potassium glass, there is a strong correlation between silica and potassium oxide, alumina, iron oxide, lead oxide, strontium oxide, potassium oxide and iron oxide, lead oxide, magnesium oxide [6] and copper oxide, and iron oxide and lead oxide, while the correlation between other pairings is weak. Among the cultural relics of lead-barium glass type, there is a strong correlation between silica and magnesium oxide, sodium oxide and lead oxide, potassium oxide and iron oxide, lead oxide, magnesium oxide and lead oxide, iron oxide and lead oxide, and lead oxide and barium oxide and strontium oxide, while the correlation between other pairings is weak. The discrete Frechet distance values between phosphorus pentoxide and other components in lead-barium glass cultural relics are very large, that is, the correlation relationship is very weak, so we can guess that phosphorus pentoxide is an essential component in lead-barium glass.

6. Conclusion

In this paper, the classification of cultural relics is firstly judged by the classification law, and then the classification of cultural relics is identified by BP neural network algorithm. Finally, according to
the idea of discrete Frechet distance algorithm, the correlation between different chemical components is converted into the similarity of two broken lines, which is easy to understand and analyze. However, when judging the correlation between different chemical components, the reference value of Frechet distance is the observation value, which may be slightly inaccurate.

References