Composition analysis and identification of ancient glass products

Wenyang Wang [*, #], Yunpeng Chen [#], Xinyu Gu [#]

College of Information Engineering, Yancheng Institute of Technology, Yancheng, China

* Corresponding author: 1240836196@qq.com

[#]These authors contributed equally.

Abstract. Aiming at the problem of composition analysis and identification of ancient glass products, a multiple linear regression model based on least square method was established, and the relationship between surface weathering and type, ornamentation and color was obtained.Firstly, the data were preprocessed, and the invalid data were eliminated by BP neural network, and the basic information of glass relics was quantified.Secondly, a multiple linear regression model was established, and the least square method was used to obtain that the surface weathering was negatively correlated with ornamentation, and positively correlated with color and type.Secondly, a component correlation model based on Spearman's coefficient was established to analyze the statistical rules of chemical components.Finally, JB test was used to verify whether the components obey the normal distribution, and the normal distribution test was completed.

Keywords: Ancient glass products, composition analysis, least squares method, BP neural network.

1. Introduction

Under the guidance of the Belt and Road Initiative, the issue of cultural exchanges between China and foreign countries has received wide attention.Glass products are the representative articles of ancient Western cultural exchanges. Glass originated from the Silk Road and was transmitted to China via West Asia, Egypt and other regions.China absorbs production technology and selects suitable local materials in combination with local conditions to produce ancient products with similar appearance and different composition.

Ancient glass generally refers to silicate mainly by artificial synthesis with amorphous structure of the material, often used in ornaments and ritual [1].Due to the influence of the environment where the glass products are located, the glass elements react with the external elements, and the glass will appear the phenomenon of weathering.Glass is made of quartz sand, mainly containing silica.In the process of melting, a cosolvent is needed to reduce its melting temperature.Now we need to study a batch of ancient glass products, which are known to be high-potassium glass and lead-barium glass.Lead-barium glass used lead ore as flux, so the content of lead oxide and barium oxide was relatively high.High potassium glass is mainly distributed in Lingnan, Southeast Asia and India, which uses plant ash as the flux.In this paper, a model was established and analyzed based on the basic information of the known types, colors, ornamentations and surface weathering of glass relics and the chemical composition of the classified glass relics at the sampling sites[2].

2. Data preprocessing

(1) Eliminate invalid data

The sum of components of cultural relics sampling points is between 85% and 105%, otherwise they will be regarded as invalid data. In this paper, the proportion of the main components in the sampling points of each cultural relic is added up, as shown in Table 1, Due to limited space, this paper only lists part of the data [3].

Table 1. Summation result table of component proportions

Cultural Relics Sampling Sites	Accumulation and	Heritage sampling sites	Accumulation and	Heritage sampling sites	Accumulation and
01	97.61	21	98.52	42 Unweathered points 1	98.01
02	99.89	22	100	42 Unweathered points 2	97.9
03 part 1	100	23 Unweathered points	96.5	43 part 1	94.68
03 part 2	98.88	24	98.88	43 part 2	96.67
04	96.06	25 Unweathered points	97.06	44 Unweathered points	99.12

According to the sum of composition ratios in Table 1, the sum of chemical composition ratio of glass relic 15 is 79.47, and the result of chemical composition ratio of glass relic 17 is 71.89, which are not in the range of 85% to 105%, so they are regarded as invalid data, and glass relics 15 and 17 are excluded [4].

(2) Filling of missing values

Firstly, the missing cultural relics were screened from Annex 1, and it was found that the missing cultural relics were all weathered lead-barium glass.Secondly, the information of weathered lead-barium glass was screened out from Annex 1, and its cultural relics were listed as 02, 08, 11, 19, 23, 25, 26, 28, 29, 34, 36, 38-44, 48-54, 56, 57, 58. Finally, 32 sampling points of weathered lead-barium glass cultural relics were screened out from Annex 2.

As the training set of BP neural network, the function relationship between color and cultural relics sampling points is fitted, and the missing cultural relics color information is obtained.

In the process of signal forward propagation, the sigmoid function is used as the excitation function, and its function expression is as follows [5].

$$\phi(x)=\frac{1}{1+e^{-\alpha x}} \tag{1}$$

The input net_k of the k node transmitted to the output layer is:

$$net_k=\sum_{i=1}^{q} w_{ki}y_j+a_k=\sum_{i=1}^{q} w_{ki}\phi(\sum_{j=1}^{M} w_{ij}x_j+\theta_i)+a_k \tag{2}$$

The output o_k of the output layer node k is:

$$o_k=\varphi(net_k)=\varphi(\sum_{i=1}^{q} w_{ki}y_j+a_k)=\varphi[\sum_{i=1}^{q} w_{ki}\phi(\sum_{j=1}^{M} w_{ij}x_j+\theta_i)+a_k] \tag{3}$$

The expected error T_k is calculated based on the output value and the actual true value. The error of the first sample is:

$$E=\frac{1}{2}\sum_{k=1}^{L}(T_k-o_k)^2 \tag{4}$$

The error gradient descent method is used to adjust the weight and threshold change error continuously, and the expected output is gradually approximated. According to the given learning rate η, the weight correction quantity is [6]:

$$\Delta w_{ki} = \eta \delta_k y_i \tag{5}$$

The correction amount of threshold is:

$$a_k = -\eta \frac{\partial E}{\partial a_k} = \eta \delta_k \tag{6}$$

The values of the colors are assigned, and the values of black, blue-green, light blue, light green, dark green and purple are 1, 2, 3, 4, 5 and 6, respectively.BP neural network was used to fit the function diagram of color and cultural relic sampling points, as shown in FIG. 1[7].

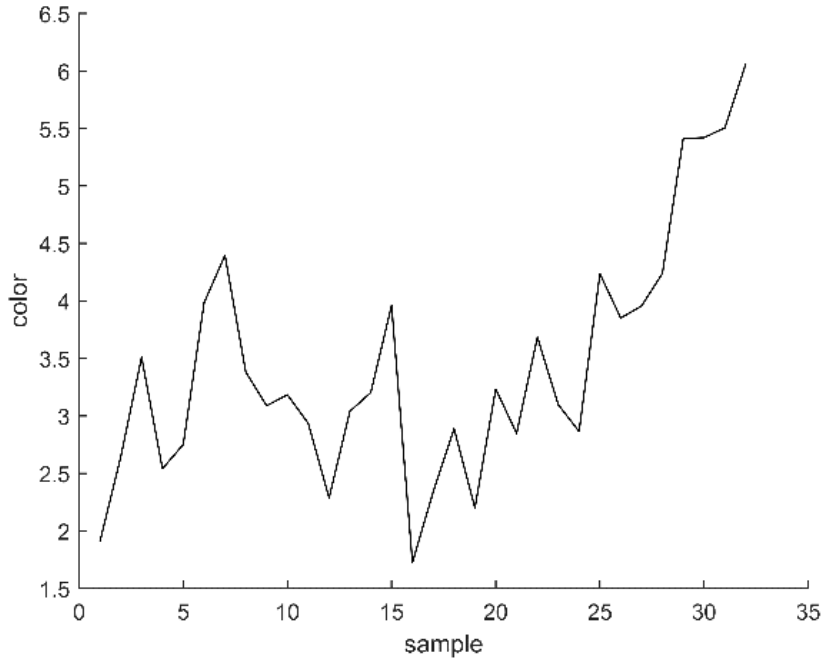


Figure 1. Fitting function diagram of color and cultural relic sampling points

Finally, the missing color information can be obtained as shown in Table 2.

Table 2. Missing color information Table

Artifact Sampling Points	type	Surface weathering	Color value	color
19	Lead barium	weathering	2.7596	Light green
40	Lead barium	weathering	4.1899	Dark green
48	Lead barium	weathering	2.9917	Light green
58	Lead barium	weathering	2.8893	Light green

That is, the missing color information is light green for artefact 19, dark green for artefact 40, light green for artefact 48, and light green for artefact 58.

(3) Standardization of data

Due to the large difference in the order of magnitude between each chemical element index, the subsequent data processing has a certain impact.Therefore, the data were standardized to make the evaluation indexes in the same magnitude system, and the data were distributed in the range of 0~1 for comparison[8].

Methods of standardized processing:

$$U_{ij}=\begin{cases}\frac{x_{ij}-minx_{ij}}{maxx_{ij}-minx_{ij}}, x_{ij}\ shows\ positive\ effect\\ \frac{maxx_{ij}-x_{ij}}{maxx_{ij}-minx_{ij}}, x_{ij}\ shows\ negative\ effect\end{cases} \tag{7}$$

3. Quantification of glass cultural relic information

The ornamentations, types, colors and surface weathering of cultural relics are all expressed in words, so it is necessary to convert their information into numerical variables to analyze the relationship between variables.

(2) Surface weathering

The weathering conditions of glass cultural relics in Annex 1 are weathering and no weathering, so 0-1 variable is introduced, so that

$y_i = 0$ Denotes that there is no weathering on the surface of the cultural relic I,

$y_i = 1$ Denotes that there is weathering on the surface of item I, where, $i = 1, 2, \cdots, 58$.

(2) glass type

From the firing process of glass, lead ore flux is added to lead-barium glass in the firing process, while potassium glass uses substances with high potassium content as flux. Therefore, the chemical composition of lead-barium glass containing lead oxide and barium oxide is higher, and the potassium content of high-potassium glass is higher.

Therefore, the ratio of the proportion of chemical components contained in the glass cultural relics was introduced[9]

$$\alpha_i = \frac{c_K}{c_K + c_{Pb} + c_{Ba}} \quad i = 1, 2, \cdots, 67 \tag{8}$$

Where, i represents the number of the cultural relic, c_K represents the proportion of chemical composition containing potassium in the glass cultural relic, c_{Pb} represents the proportion of chemical composition containing lead in the glass cultural relic, c_{Ba} represents the proportion of chemical composition containing barium in the glass cultural relic. The larger the number α_i, the closer it is to 1, the type of glass is high potassium glass; α_i the smaller, the closer to 0, the type of glass is lead-barium glass.

In the process of processing, it is found that some glass products in the data do not contain three elements, that is, the denominator will be 0.At this time, if the glass type is high potassium glass, take $x_{1i} = 1$; If the glass type is lead-barium glass, take $x_{1i} = 0$. Cultural relics 7 and 27 do not contain the three elements, and both are high potassium glasses, so the value of glass type is 1.

(3) Ornamentation

Firstly, the data of ornamentation were statistically analyzed, and the results are shown in Table 3. [10]

Table 3. Statistics of glass texture data

		Class A grain	Class B grain	Class C grain
type	High potassium glass	6	6	6
	Lead barium glass	16	0	24
color	Light green	2	0	4
	Light blue	10	0	11
	green	0	0	1
	Blue, green	6	6	2
	Deep blue	2	0	8
	purple	0	0	4
	black	2	0	0
weathering	Weathering on the surface	11	6	17
	No weathering on the surface	11	0	13

According to the statistics in Table 3, 11 of the 22 cultural relics with type A ornamented surface are weathered; 6 cultural relics with category B ornamentation all have surface weathering;Among the 30 cultural relics with category C ornamentation, 17 cultural relics had surface weathering.

From the above data, it can be seen that ornamentation and surface weathering have a great correlation. Therefore, the weathering rate of ornamentation is used to represent ornamentation data, which is specifically expressed as:

$$q_i = \frac{m_i}{n_i} \quad i = 1, 2, \cdots, 68 \tag{9}$$

Where, denotes the number of cultural relics with category I ornamentation, and denotes the number of cultural relics with Category I ornamentation weathered. n_i i m_i i There are three types of ornamentation: A, B and C.The texture quantification table is shown in Table 4.

Table 4 Texture Quantization Table

Texture types	quantitative	Cultural relics number
A	0.5	02 03 04 05 06 18 19 20 21 23 28 29 30 42 44 45 46 47 48 49 50 53
B	1	07 09 10 12 22 27
C	0.629	01 08 11 13 14 15 16 17 24 25 26 31 32 33 34 36 37 38 39 40 41 43 51 52 54 55 56 57 58

(4) Glass color

According to the color depth, the values of light green, light blue, green, blue green, dark green, dark blue, purple and black are successively assigned as 1, 2, 3, 4, 5, 6, 7 and 8 from light to dark.

4. Model establishment and solution

4.1. Establishment of multiple regression model

The general form of multiple linear regression model is

$$\eta(u) = \beta_1\varphi_1(u) + \beta_2\varphi_2(u) + ... + \beta_m\varphi_m(u) \tag{10}$$

Make

$$y = \beta_1\varphi_1(u) + \beta_2\varphi_2(u) + ... + \beta_m\varphi_m(u) + \varepsilon \tag{11}$$

Where, ε is the random error, and $\varepsilon \sim N\left(0,\sigma^2\right)$; $\varphi_i(u)$ Is the explanatory variable and is a known function.

Suppose that an experiment is performed and the group observation value is:

$$\begin{bmatrix} u_1 & y_1 \\ u_2 & y_2 \\ \vdots & \vdots \\ u_n & y_n \end{bmatrix} \tag{12}$$

Substituting it into the regression formula with error, it can be obtained

$$\begin{cases} y_i = \beta_1\varphi_1(u_i) + \beta_2\varphi_2(u_i) + ... + \beta_m\varphi_m(u_i) + \varepsilon_i, \\ i = 1, 2, \cdots, n, \end{cases} \tag{13}$$

Where ε_i are the independent random errors in the second trial, and $\varepsilon_i \sim N\left(0,\sigma\right)$.

It is required to analyze the relationship between the surface weathering of glass relics and its glass type, ornamentation and color, that is, it is required to indicate that one variable is affected by multiple variables, so a multiple linear regression model with three regressors is established as [3]: $y = \beta_0 + \beta_1 x + \beta_2 x + \beta_3 x + \varepsilon$

Where, is the glass type of the cultural relic, is the ornamentation of the cultural relic, is the color of the cultural relic, is the regression coefficient, and is the error term. $x_1\ x_2\ x_3\ \beta\ \varepsilon$

4.2. Solution of multiple regression model

Using the least squares method to solve the regression coefficients in the multiple linear regression model, the least squares criterion can be obtained as follows for the given 56 groups of glass sample data:

$$s\left(\beta_0,\beta_1,\beta_2,\beta_3\right) = \sum_{i=1}^{56}(y_i - \beta_0 - \sum_{j=1}^{2}\beta_j x_{ij})^2 \tag{14}$$

Calculate and solve the least squares estimator on the regression coefficients, which is transformed into a simple linear regression model using a matrix: $\beta_0, \beta_1, \beta_2, \beta_3$,and $y = X\beta + \varepsilon$

In the model,

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{56} \end{pmatrix},\ X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{561} & x_{562} & x_{563} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{56} \end{pmatrix},\ \varepsilon = \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_{56} \end{pmatrix} \tag{15}$$

The least squares estimate is: β

$$\beta = \left(X^{'}X\right)^{-1}X^{'}y \tag{16}$$

Substituting the processed numerical variables into the multiple linear regression model, the results are shown in Table 5.

Table 5. Regression model results

variable	The coefficient of	P values	Lower limit of confidence interval	Upper confidence interval
Glass types	0.647	0	0.378	0.916
grain	1.904	0	2.718	1.091
color	0.010	0.713	0.044	0.064
Constant term	1.331	0	0.808	1.855

The surface weathering of glass relics as a function of glass type, ornamentation and color is obtained as follows.

$$y = -0.647x_1 + 1.904x_2 - 0.010x_3 + 1.331 \tag{17}$$

As can be seen from Table 5, the relationship between surface weathering glass types, ornamentation and color of glass samples is as follows:

(1) Surface weathering and glass type of cultural relics: The P-value of glass type is less than 0.05, which means that the null hypothesis is rejected at the 95% confidence level, indicating that the influence of glass type on surface weathering is significant.In the case of glass weathering, the regression coefficient of glass type is -6.47, indicating a negative correlation between glass type and weathering degree. In other words, the smaller the content of potassium in glass, the more similar the glass is to lead-barium glass, and the more likely the glass surface is to be weathered.It is speculated that the high-potassium glass has certain advantages in weathering resistance, and the technology is closer to the modern technology.

(2) Surface weathering and ornamentation of cultural relics: The P-value of ornamentation is less than 0.05, which means that the null hypothesis is rejected at the 95% confidence level, indicating that ornamentation has a significant effect on surface weathering.The regression coefficient of ornamentation was 1.904, indicating that ornamentation was positively correlated with the degree of weathering, indicating that ornamentation B was the most prone to surface weathering of glass, followed by ornamentation C, and ornamentation A was the least prone to surface weathering.It is speculated that B ornamentation is more three-dimensional, more fragile, and more susceptible to water and wind erosion during burial.

(3) Surface weathering and color of cultural relics: the p value of color is 0.713, greater than 0.05, which cannot reject the null hypothesis, indicating that the influence of color on glass weathering is insignificant and can be ignored.The regression coefficient of color was 0.010, indicating a positive correlation between color and weathering degree. That is, the lighter the color of B, the more likely it is to be weathered. It is speculated that weathering produced chemicals with lighter colors.

4.3. Conclusions of the model

The surface weathered glass type, ornamentation, and color of the glass sample are related as follows:

(1) Surface weathering of cultural relics and glass type: the influence of glass type on surface weathering is significant, and there is a negative correlation between glass type and weathering degree. That is, the smaller the content of potassium in glass, the closer the glass is to lead-barium glass, and the more prone it is to surface weathering.

(2) Surface weathering and ornamentation: ornamentation has A significant effect on surface weathering, and there is a positive correlation between ornamentation and weathering degree. Ornamentation B is the most prone to surface weathering, ornamentation C is the next, and ornamentation A is the least prone to weathering.

(3) Surface weathering and color of cultural relics: the influence of color on the weathering of glass is insignificant, and the color is positively correlated with the weathering degree. The lighter the color of B, the more likely it is to be weathered.

5. Jarque - Bera inspection

Based on the type and surface weathering of glass, the statistical rule of the chemical composition content of glass samples was obtained, which belongs to the data analysis problem. The Jarque-Bera test was used for normal distribution, and the data were classified for descriptive statistics, and the composition correlation model based on Spearman coefficient was established.Jarque-Bera test is a goodness-of-fit test, which is applicable to whether a large sample population obeys normal distribution. The calculation formula is as follows:

$$J=\frac{N}{6}\left[\left(\frac{\frac{1}{N}\sum_{i=1}^{N}(X_i-\overline{X})^2}{\sigma^3}\right)+\frac{1}{4}\left(\frac{\frac{1}{N}\sum_{i=1}^{N}(X_i-\overline{X})^4}{\sigma^4}-3\right)^2\right] \tag{18}$$

Where, J is the Jarque-Bera coefficient, X_i is the possible value of the percentage of chemical composition of glass sample; $\overline{X}$ is the average value of the chemical composition percentage of glass samples; σ Is the standard deviation of glass sample composition; N is the number of glass samples.

The normal distribution of chemical composition can be obtained by Jarque-Bera test as shown in Table 6.

Table 6. Normal distribution of chemical components

Chemical composition	Whether normal distribution is satisfied
silica	is
Sodium oxide	no
kali	no
Calcium oxide	no
magnesia	is
alumina	no
Iron oxide	no
Copper oxide	no
Lead oxide	is

6. Conclusion

The statistical law of weathering chemical content on the surface of glass samples was described in detail in terms of data distribution, concentration, difference and correlation, based on the classification of glass type and weathering degree.The unweathered and weathered chemical components were arranged in ascending order, and the unweathered and weathered components were combined as the results of the reaction of different chemical components over time, which could be better predicted from the perspective of time.

References

[1] ZHAO Hualing. Study on the Development of Ancient Glass History in Zibo Area [D]. Soochow University, 2008, 1.

[2] ZHU Rongjie. Study on spatial-temporal prediction and pollution warning model of river water quality based on improved BP neural network [D]. Harbin Institute of Technology, 2021, 9 - 12.

[3] FU S. Research on security evaluation of cryptographic algorithms based on side channel analysis [D]. Beijing University of Posts and Telecommunications, 2020, 38.

[4] Yu Qun, Huo Xiaodong, He Jian, Li Lin, Zhang Jianxin, Feng Yuyao. [4] YU Q, HE J, Li L, FENG Y Y. Trend prediction of power outage accidents in China power grid based on Spearman correlation coefficient and system inertia [J/OL]. Proceedings of the CSEE :1-12, 2022 - 09 - 16.

[5] Zhang Yuanning, Huang Jingguang, Zhang Yanhua, Yu Nan, Chen Yong, Weng Hanli, Jia Yongbo, Li Zhenxing. [5] High Order Statistical Principle of Identifying transformer inrush Current with Dual Characteristic Quantity [J]. Protection and Control of Electric Power Systems, 2021, 49 (24): 21 - 30.

[6] Zhou Xiuya. Descriptive analysis of factors influencing the high suicide rate in Nunavut, Canada, 2001-2011[D]. Beijing Foreign Studies University,2015. 2019, 07: 51 - 52.]

[7] WANG S. Research and application of time series prediction algorithm based on deep learning [D]. University of Electronic Science and Technology of China, 2022, 10 - 12.

[8] JIANG Na. Research on Performance Degradation Prediction and Maintenance Strategy of Turbofan Engine [D]. Harbin Institute of Technology, 2020, 90.

[9] Che Yujian. Research on Health State Prediction of Power Battery of Pure Electric Vehicle [D]. Jilin University, 2022, 56 - 60.

[10] YU Weiheng. Research on retrieving large earthquake rupture scale based on Fisher linear identification method [D]. Institute of Engineering Mechanics, China Earthquake Administration, 2016, 9 - 11.