

Prediction of Chemical Composition of Ancient Glass Relics before Weathering

Jiehua Sun¹, Huazhou Chen^{2*}, Yao Liu^{1,3}, Hongquan Lin¹, Huiwen Zheng¹, Yingzhen Qiu¹

¹School of Tourism Data, Guilin University of Tourism, Guilin, China

²School of Science, Guilin University of Technology, Guilin, China

³Guangxi Key Laboratory of Culture and Tourism Smart Technology, Guilin Tourism University, Guilin, China

Email: *sunjehua428@163.com

How to cite this paper: Sun, J.H., Chen, H.Z., Liu, Y., Lin, H.Q., Zheng, H.W. and Qiu, Y.Z. (2023) Prediction of Chemical Composition of Ancient Glass Relics before Weathering. *Open Journal of Applied Sciences*, 13, 1565-1580.

<https://doi.org/10.4236/ojapps.2023.139124>

Received: August 16, 2023

Accepted: September 19, 2023

Published: September 22, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Ancient glass relics are easily weathered by the influence of buried environment, and the internal elements exchange with the environmental elements in large quantities, resulting in changes in their composition ratio. Archaeological research can often detect the component content of glass relics after weathering, but it is difficult to obtain the corresponding component content before weathering. It is necessary to predict the chemical composition of glass relics before weathering in order to accurately identify the type of glass relics and repair them. To solve this problem, this paper proposes a distributed matching strategy, and studies the influence of weathering on the composition content of glass through compositional correlation analysis and linear regression statistical methods, so as to build a prediction model of the composition content of glass relics before weathering. The results show that the composition prediction model of glass cultural relics constructed by the distribution matching strategy has a good prediction ability, which is consistent with the change trend of the composition ratio of linear regression analysis. Moreover, the model is simple and easy to operate, which is convenient for popularization and application, and provides theoretical basis and reference value for further research on the composition and accurate classification of glass cultural relics.

Keywords

Glass Relics, Weathering, Chemical Composition, Correlation Analysis, Distribution Matching, Model Prediction

1. Introduction

As a kind of precious vessel, glass is defined as the precious material evidence of

the early trade between China and the West in ancient times [1]. The main raw material of glass is quartz sand, and its chemical composition is SiO_2 . The difference in the flux added when making glass will lead to the difference in the main chemical composition of the final glass. In ancient China, there were mainly two types of lead barium glass and high potassium glass. The content of PbO and BaO in lead-barium glass will be increased by adding the flux lead ore in the firing process. High potassium glass is made from plant ash or other substances with high potassium content as a flux [2].

Ancient glass products have been buried for a long time, which is highly susceptible to the influence of its surrounding environment. When the internal elements of glass exchange a large amount of chemical elements in the surrounding environment, weathering occurs, which will lead to changes in its chemical composition [3]. In archaeological research, the unearthed glass relics are all in the state after weathering, so it is difficult to accurately judge the chemical composition content before weathering, which seriously affects the type identification of glass relics. In order to accurately identify the type of glass relics and then carry out restoration work, it is necessary to measure/predict the chemical composition content of glass relics before weathering.

It is a hot topic to identify the type of glass cultural relics by quantitative analysis of chemical composition [4] [5]. In the early period, scholars did some research on the method of determining the chemical composition of glass relics. Gan Fuxi *et al.* [6] analyzed the chemical composition of glass beads by proton excited X-ray fluorescence and inductively coupled plasma emission spectroscopy; Li Qinghui *et al.* [7] [8] proposed a method of proton induced X-ray emission (PIXE) for the determination of the chemical composition of ancient glass. In recent years, with the development of information technology, neural networks, multi-layer perceptrons, decision trees, random forests, feature selection and other machine learning methods have also been applied to the research on the type identification of glass relics. Cao Yuxuan *et al.* [9] used GA-BP neural network to study the relationship between the internal chemical composition of different types of ancient glass and their own weathering degree, and established models for predicting the types of ancient glass and measuring the weathering degree; Shi Baoming *et al.* [10] established a multi-layer perceptron network model to predict the categories of ancient glass products; Lv Fei *et al.* [11] established decision tree model and random forest model to identify the types of ancient glass products. In addition, there are few quantitative studies on the prediction of various chemical components of glass cultural relics before weathering. Zou Ying [12] established a multivariate time series model to predict the contents of various chemical components of glass cultural relics before weathering.

Different from the methods of other scholars, this paper combined statistical correlation analysis and linear regression method to propose a strategy of data distribution matching. Assuming that there are two populations with normal

distribution before and after glass weathering, the mean value and standard deviation of these two populations are used to build a prediction model for the contents of various chemical components of glass relics before weathering.

In order to better demonstrate the research idea of this paper, **Figure 1** is used to show the general framework of this paper.

At the same time, the symbolic descriptions are shown in **Table 1**.

2. Data Source and Processing

According to the chemical composition of glass cultural relics and other detection methods, archaeologists collected a number of Chinese ancient glass cultural relics related data (2022 National College students Mathematical Competition in Modeling C), including the classification information of glass cultural relics and the proportion of the corresponding main components. However, due to the detection methods and other reasons, there are missing data or zero data, and the composition ratio of most samples is not accurate to 100%. Therefore, we first carry out necessary pre-processing of the data.

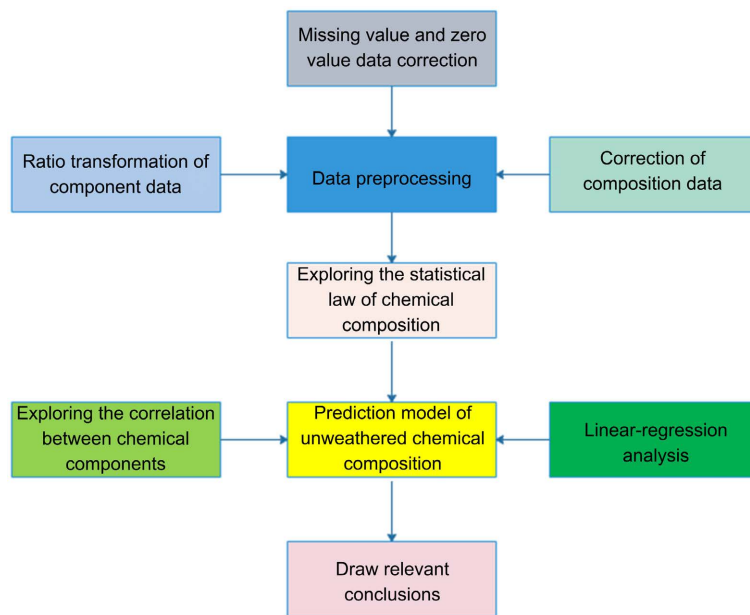


Figure 1. Overall thinking process.

Table 1. Symbols.

Symbols	Implication	Symbols	Implication
x	Component vector	X	Weathered glass population
y	Transformed component vector	Y	Unweathered glass population
ρ_{xy}	Correlation coefficient between variables x and y	$CLR(X)$	Weathered chemical composition content
σ_x	Standard deviation of population X	$CLR(Y)$	Unweathered chemical composition content
σ_y	Standard deviation of population Y	μ_y	Mean of population Y
μ_x	Mean of population X		

2.1. Missing Value and Zero Value Data Correction

The missing value is mainly generated because the content of one or several elements in the sample is very low and does not reach the detection limit, or the absence of artificial data entry cannot be excluded. There are three commonly used data missing value processing methods: 1) Directly delete the sample or variable containing the missing value; 2) Assign an arbitrary value below the detection limit to the missing value; The missing value is estimated based on the component analysis of the associated sample (e.g., likelihood estimation). In view of the fact that the original data is the component data, the missing values are too numerous and too scattered, and it is not suitable to directly delete or directly assign arbitrary values below the detection limit. Therefore, we use the component analysis based on the associated samples to estimate the missing values. Specifically, the number of missing components is calculated for each sample, and the missing data is filled by the expected maximum likelihood estimation method combined with the ideal condition of 100% of the proportion of each component.

The absence of an element in the sample or the amount of an element that does not reach the detection limit may result in a zero value in the recorded data. Considering the small number of zero values of the original data in the sample and the scattered distribution, we choose to modify the zero values recorded in the original data to a small positive number (0.001), which is convenient for subsequent data processing and modeling calculation.

2.2. Correction of Composition Data

The component data is non-negative, and the sum of each individual element content in sample is 1 [13], that is $W_{N \times D} = \{W_{ij} \geq 0 \mid i = 1, 2, \dots, N; j = 1, 2, \dots, D\}$, and $W_{i1} + W_{i2} + \dots + W_{iD} = 1 (i = 1, 2, \dots, N)$. Therefore, if the cumulative sum of component data is biased, the correlation analysis results between elements will be biased, which will further affect the covariance and correlation matrix between components [14].

The direct data studied in this paper is the content proportion of each chemical component, that is, the component data, which should theoretically meet the constraint of cumulative sum of 1. However, due to the accumulation of the proportion of each chemical component and non-100% deviation caused by the detection means, it is necessary to correct the deviation. If the cumulative sample component content of the original data is between 85% and 105%, it is considered that the fixed sum deviation of the data is not large, and the sample data is valid and can be corrected. If the component content of the original data accumulates and exceeds the range of 85% to 105%, the sample is considered invalid data and cannot be corrected. For valid sample data that can be corrected, the sum is fixed according to formula (1):

$$x'_{ij} = \frac{x_{ij}}{\sum_{j=1}^m x_{ij}} \quad (1)$$

where x_{ij} ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$) represents the original data of the indicator j of the sample i , x'_{ij} represents the converted data. After conversion, the proportion of each component of each sample is 100%.

The descriptive statistical results of the data after modification are shown in **Table 2**. It can be seen from **Table 1** that the skewness coefficients of component indicators in the sample are all greater than 0, indicating that there are more data on the right side of the mean value. The skewness coefficient of SiO₂ is close to 0, indicating that the data of SiO₂ is close to symmetric distribution. The kurtosis coefficient of 2 component indices in the sample is less than 0, indicating that the distribution is flatter at the top or thinner at the tail than the normal distribution. The kurtosis coefficients of the remaining components are greater than 0, indicating that the distribution is sharper at the top or thicker at the tail than the normal distribution. In general, the content of most components in the data does not conform to the normal distribution.

2.3. Ratio Transformation of Component Data

Since the ratio of component data variables is not restricted by the “constant sum” restriction, and the logarithm of the ratio usually follows the characteristics of normal distribution, this paper adopts the methods of additive log-ratio transformation (ALR) and central log-ratio transformation (CLR) [15], so that the component data presents the situation of multivariate normal distribution after transformation. Considering that the vector after the additive log-ratio transform (ALR) cannot correspond to the original vector one-to-one, the central log-ratio transform [16] is adopted:

Let component vector $x = (x_1, x_2, \dots, x_D), (x_i > 0)$, and $\sum_{i=1}^D x_i = 1$, let

$$y_i = \ln \frac{x_i}{\sqrt[D]{\prod_{i=1}^D x_i}}, i = 1, 2, \dots, D. \quad (2)$$

This transformation is called the central log-ratio transformation (CLR).

The corresponding inverse transformation is:

$$x_i = \frac{\exp(y_i)}{\sum_{i=1}^D \exp(y_i)}, i = 1, 2, \dots, D. \quad (3)$$

The transformed component corresponds to the original component one by one, and further enhances the interpretability of the variable. In addition, the central log-ratio transformation can make the data more stable, reduce the influence of outliers, and reflect the real situation of the data more accurately.

Use x_1, x_2, \dots, x_{14} to indicate the chemical composition of glass relics SiO₂, Na₂O..., SO₂. According to formula (2), for the component vector $x = (x_1, x_2, \dots, x_{14}) \in S^{14} (x_i > 0)$, the data transformed by CLR is:

$$y = (y_1, y_2, \dots, y_{14}) = \left(\log \frac{x_1}{\sqrt[14]{\prod_{i=1}^{14} x_i}}, \dots, \log \frac{x_{14}}{\sqrt[14]{\prod_{i=1}^{14} x_i}} \right) \quad (4)$$

Table 2. Descriptive statistics of data after modification.

Variable	Sample Capacity	Minimum Value	Maximum Value	Mean Value	Standard Deviation	Skewness	Kurtosis
SiO ₂	69	3.720	96.701	50.180	23.682	0.095	-0.799
Na ₂ O	69	0.001	8.203	1.079	1.726	1.893	3.901
K ₂ O	69	0.010	14.749	2.047	3.882	1.946	2.423
CaO	69	0.010	9.935	2.849	2.550	1.093	0.365
MgO	69	0.001	2.912	0.735	0.626	0.664	0.512
Al ₂ O ₃	69	0.450	14.354	4.213	3.157	1.547	2.284
Fe ₂ O ₃	69	0.001	5.938	0.872	1.154	2.028	5.262
CuO	69	0.001	10.584	1.968	2.221	2.181	5.394
PbO	69	0.001	71.185	24.526	20.200	0.311	-1.010
BaO	69	0.000	35.475	7.837	8.506	1.626	2.487
P ₂ O ₅	69	0.010	14.568	2.704	3.636	1.518	1.459
SrO	69	0.001	1.155	0.266	0.268	1.124	0.955
SnO ₂	69	0.000	2.365	0.118	0.416	4.586	21.245
SO ₂	69	0.001	15.961	0.605	2.659	5.368	28.748

According to formula (4), the data after logarithm transformation of the content center of high potassium and lead barium glass can be calculated. In view of the rationality of the hypothesis of normal distribution of data, it is necessary to test whether the data distribution pattern conforms to normal distribution. There are many testing methods for normal distribution, including Shapiro-Wilk test, Kolmogrov-Smirnov test, skewness test and kurtosis test [17]. In this paper, the normal distribution of data is tested by skewness and kurtosis test. In statistics, skewness is used to describe symmetry [18], and kurtosis represents the characteristic number of the height of kurtosis at the mean of the probability density distribution curve. The skewness and kurtosis of the data can be calculated through the software. Limited by space, only the skewness and kurtosis of the lead-barium glass data are shown (Table 3). When the skewness and kurtosis values are closer to 0, the data tends to be more normally distributed. In order to intuitively understand the influence of the center log-ratio transformation on the normal distribution of data, the original data frequency histogram and the data frequency histogram after the center log-ratio transformation were made for the data in this paper (Figure 2 and Figure 3). Limited by space, only some data frequency histograms of lead-barium glass are shown.

As can be seen from Table 3 and Figure 2 and Figure 3, the skewness and kurtosis of the composition content data of lead-barium glass are high, so most of the composition content in the original data does not conform to the normal distribution. After the central logarithmic ratio transformation, skewness and kurtosis of most component content data are reduced, and the processed data tend to be more normal distribution.

Table 3. Test table of skewness kurtosis of lead-barium glass.

Variable	Unweathered				Weathering			
	Original Data		Clr data		Original Data		Clr data	
	Skewness	Kurtosis	Kurtosis	Kurtosis	Kurtosis	Kurtosis	Kurtosis	Kurtosis
SiO ₂	-0.432	-0.379	0.668	0.328	0.189	1.111	-0.976	1.763
Na ₂ O	1.309	0.863	0.327	-1.915	2.663	6.708	1.851	2.261
K ₂ O	2.853	9.502	-0.299	-1.279	2.513	7.750	0.636	-1.072
CaO	1.356	1.060	-1.810	2.523	0.426	-0.508	-2.755	10.534
MgO	0.081	-1.247	-0.599	-1.536	1.129	1.610	-0.244	-1.786
Al ₂ O ₃	1.915	3.924	0.352	-0.521	2.756	10.243	0.067	-0.787
Fe ₂ O ₃	2.038	4.582	0.231	-1.873	1.445	1.722	-0.385	-1.599
CuO	2.378	6.139	-1.057	2.143	2.050	3.895	-1.567	3.209
PbO	0.641	-0.545	-0.050	-0.639	-0.174	0.048	0.044	0.024
BaO	1.879	3.923	-0.381	1.210	1.184	0.597	-1.955	3.200
P ₂ O ₅	2.220	3.957	-0.155	-1.054	0.386	-0.892	-1.606	2.265
SrO	1.291	2.090	-0.601	-0.722	0.559	0.721	-0.779	1.083
SnO ₂	2.648	5.902	1.913	2.836	4.337	19.528	2.483	6.471
SO ₂	4.796	23.000	4.032	17.971	3.260	9.619	1.939	2.451

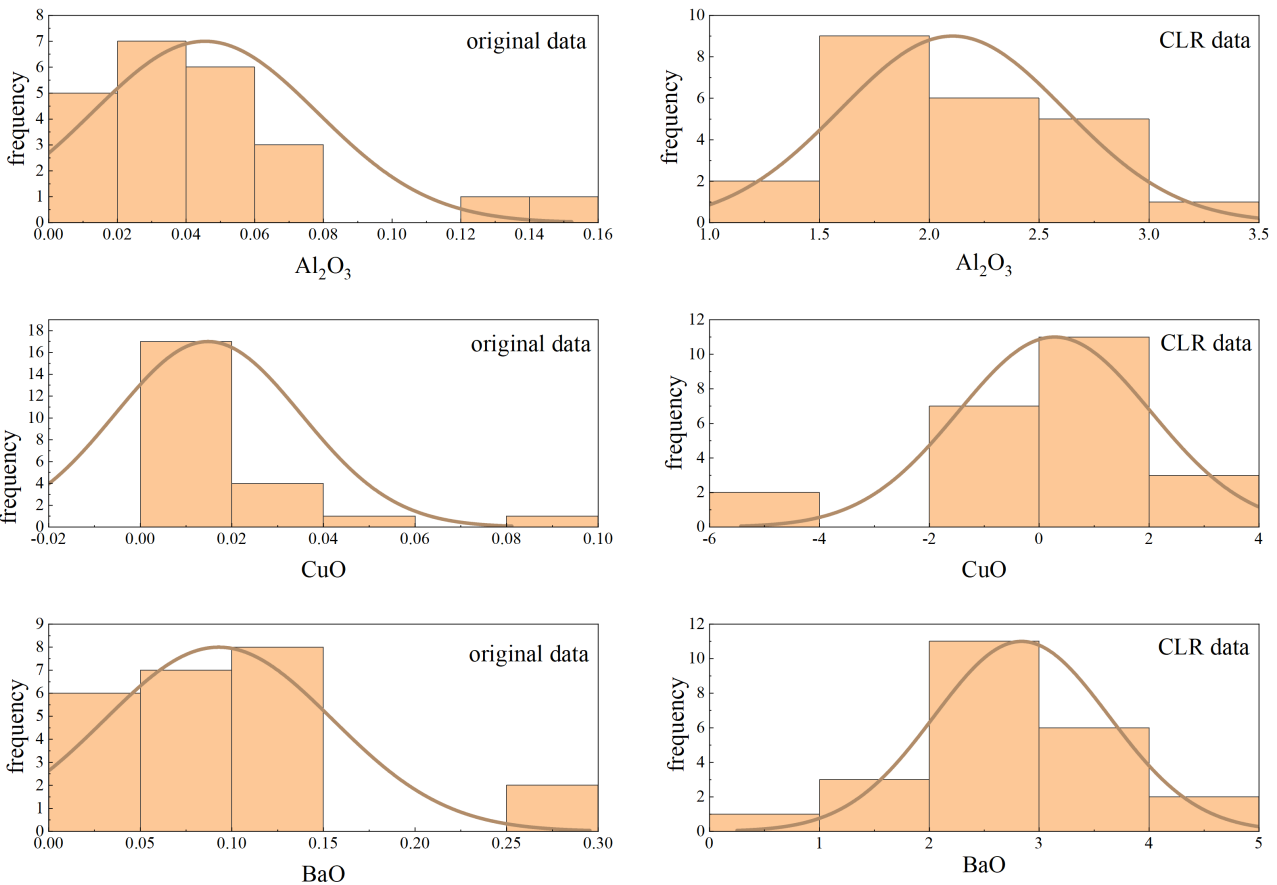


Figure 2. Histogram of the frequency of Al₂O₃, CuO and BaO of the unweathered lead barium population.

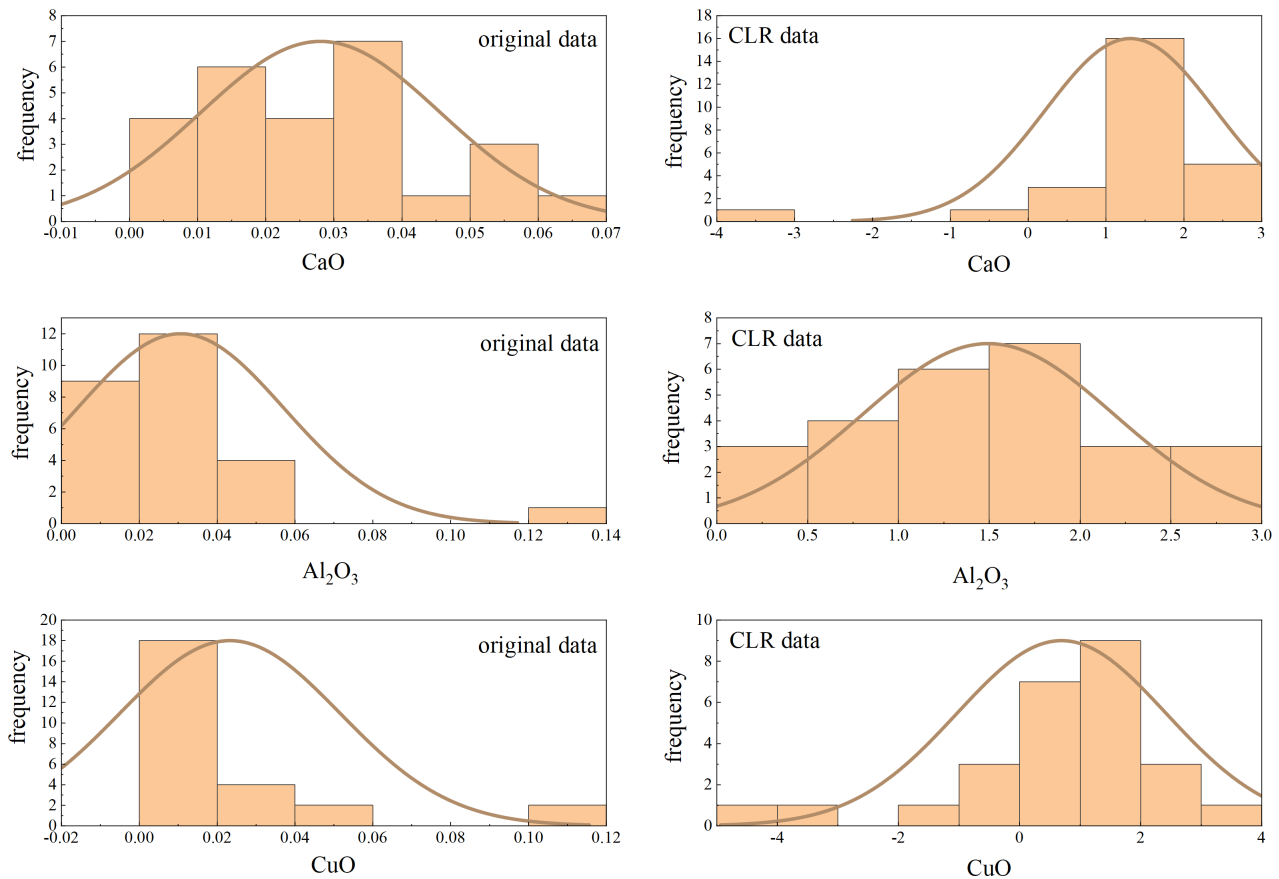


Figure 3. Histogram of CaO, Al₂O₃ and CuO frequencies of Pb barium weathering.

3. Model Design

It is assumed that the glass relics are weathered under natural conditions and the chemical composition is measured accurately. According to the data characteristics of chemical components of glass relics, Pearson correlation coefficient was used to study the correlation of chemical components of glass relics, and the correlation degree and difference between different types of glass relics and their chemical components were discussed. Secondly, linear regression method was used to discuss the influence of weathering state on the content of chemical components of glass relics. At last, the prediction model of the chemical composition contents of glass relics before weathering was constructed by the idea of distribution matching.

3.1. Correlation Analysis of Chemical Components of Glass Relics

The “closure effect” caused by the constant value of sum will lead to the deviation of the analysis results of the correlation between variables. Therefore, the Pearson correlation coefficient between the variables of high potassium and lead barium glass is calculated using the data after the transformation of the central logarithm ratio. The calculation formula is as follows:

$$\rho_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

where, x_i, y_i respectively represents the chemical composition variables of glass relics, \bar{x}, \bar{y} respectively represents the average value of the two variables, and n is the sample size. The closer the calculation result is to 1, the stronger the positive correlation between the two variables is, and the closer the result is, the stronger the negative correlation is. According to the calculation results of equation (5), we obtained the heat map of the correlation coefficient of high-potassium and lead-barium glasses, and the results are shown in **Figure 4** and **Figure 5**. The results show that there is a high correlation between the glass type and the proportion of some chemical components, and the correlation between different types of glass and their chemical components is different. For example, for high potassium glass, SiO_2 is positively correlated with K_2O , CaO , MgO , Al_2O_3 , CuO , P_2O_5 and SnO_2 , and negatively correlated with Na_2O and SrO . SiO_2 has a very significant positive correlation with Al_2O_3 and a very significant negative correlation with Na_2O . For lead-barium glass, SiO_2 has a very significant positive correlation with Na_2O and Al_2O_3 , and a very significant negative correlation with P_2O_5 and SrO . It can also be concluded that for high potassium glass, SiO_2 has a positive correlation with P_2O_5 , and a positive correlation with Na_2O , while the lead barium glass is the opposite. Therefore, the correlation between variables of different glass types is quite different.

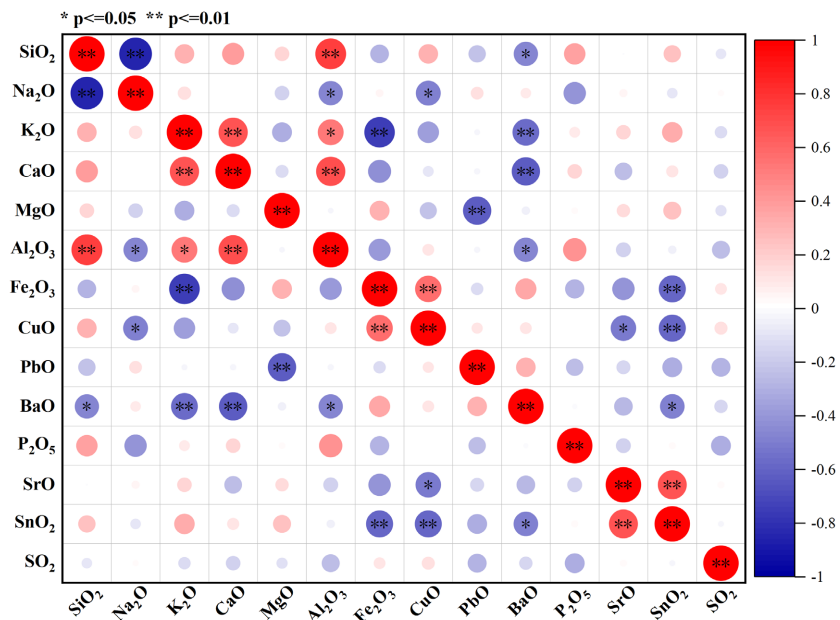


Figure 4. Heat map of correlation coefficient of each chemical composition of high potassium glass. Note: In the **Figure 4**, red represents the positive correlation, blue represents the negative correlation, the size of the circle represents the absolute value of the correlation coefficient, and the asterisk on the circle represents the existence of significance, one asterisk means significant, and two asterisks mean very significant.

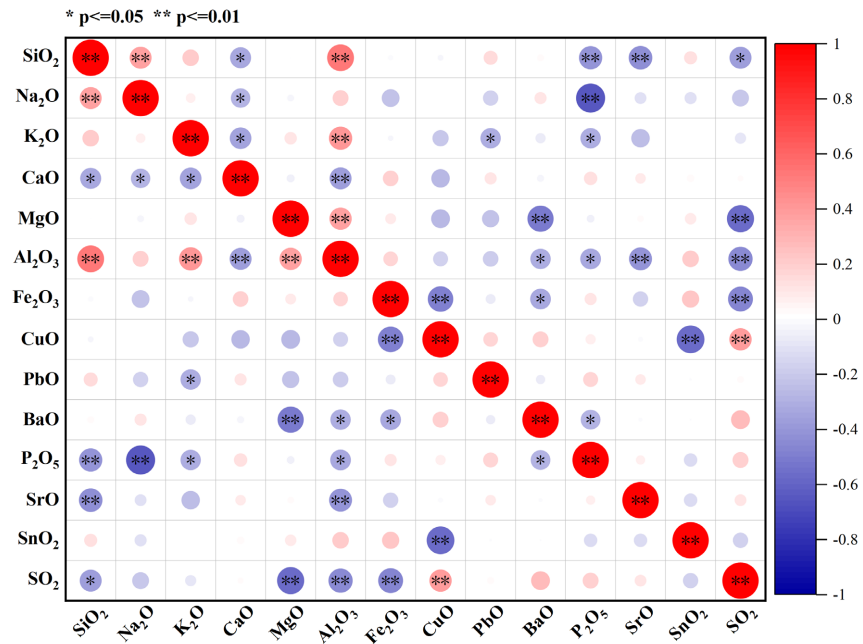


Figure 5. Heat map of correlation coefficient of chemical composition of lead barium glass. Note: In the **Figure 5**, red represents the positive correlation, blue represents the negative correlation, the size of the circle represents the absolute value of the correlation coefficient, and the asterisk on the circle represents the existence of significance, one asterisk means significant, and two asterisks mean very significant.

3.2. Analysis of the Relationship between Weathering State and the Content of Chemical Components of Glass Relics

Since the weathering state is divided into weathering and unweathering, which belong to categorical variables, and the chemical composition content of glass relics belongs to continuous variables, linear regression analysis can be carried out by setting dummy variables. With no weathering as a reference variable and weathering state as a dummy variable, linear regression models of weathering state of high-potassium glass and lead-barium glass on the content of each chemical component were established according to the data after CLR transformation. The results are shown in **Table 4**.

The results showed that weathering had different effects on the contents of chemical components of different glass types. 1) Weathering can cause significant changes in the proportion of SiO₂ and Na₂O components of high potassium glass, p value is 0.000 (<0.05), and can cause changes in the proportion of CuO components to a certain extent. According to the linear model primary term coefficient, surface weathering causes the proportion of SiO₂ and CuO components to increase, and the proportion of Na₂O components to decrease significantly. 2) Weathering can cause significant changes in the proportions of SiO₂, Na₂O, K₂O, CaO, Al₂O₃, PbO and P₂O₅ of lead-barium glass to a certain extent, and the p value is less than 0.05. According to the linear model primary term coefficient, the surface weathering causes the proportion of SiO₂, Na₂O, K₂O and Al₂O₃ to decrease, while the proportion of CaO, PbO and P₂O₅ to increase.

Table 4. Linear regression results of high potassium and lead barium glasses.

Type	Variable	Unnormalized Coefficient	t	p value	Variable	Unnormalized Coefficient	t	p value
High Potassium	SiO ₂	2.114	4.625	0.000	CuO	1.931	2.589	0.019
	Na ₂ O	-4.941	-7.237	0.000	PbO	-1.522	-1.665	0.113
	K ₂ O	-0.572	-1.317	0.205	BaO	-0.886	-0.855	0.404
	CaO	-0.240	-0.586	0.565	P ₂ O ₅	0.487	0.617	0.545
	MgO	1.222	2.077	0.052	SrO	0.123	0.200	0.844
	Al ₂ O ₃	0.501	1.225	0.225	SnO ₂	0.398	0.389	0.702
	Fe ₂ O ₃	0.960	1.174	0.256	SO ₂	0.424	0.532	0.601
Lead Barium	SiO ₂	-1.008	-5.422	0.000	CuO	0.414	0.830	0.411
	Na ₂ O	-1.896	-2.778	0.008	PbO	0.572	3.032	0.004
	K ₂ O	-1.031	-2.282	0.027	BaO	-0.566	-1.132	0.263
	CaO	0.986	2.509	0.016	P ₂ O ₅	2.339	4.062	0.000
	MgO	-0.411	-0.685	0.497	SrO	0.626	1.717	0.093
	Al ₂ O ₃	-0.612	-3.477	0.001	SnO ₂	-0.249	-0.781	0.439
	Fe ₂ O ₃	0.236	0.376	0.709	SO ₂	0.600	1.091	0.281

3.3. Construction of a Prediction Model for the Contents of Each Chemical Component of Glass Relics before Weathering

Due to the lack of detection data before and after weathering, the problem of predicting the chemical composition content of glass before weathering from the detection data of weathering points and unweathered cultural relics needs to analyze the statistical laws of weathering and unweathered populations, and make prediction through the relationship between the two populations. Therefore, the idea of distributed matching can be used for prediction. According to the fact that the ratio of the component data variables is not restricted by the “fixed sum” and the logarithm of the ratio usually follows the characteristics of normal distribution, it is assumed that the data after the transformation of the central logarithm follows the normal distribution. Considering two populations of weathering and unweathering, namely weathered glass population $X \sim N(\mu_x, \sigma_x^2)$ and unweathered glass population $Y \sim N(\mu_y, \sigma_y^2)$, the data after the transformation of the central logarithm ratio were classified and processed to obtain four types of data, namely, high potassium weathering (6), high potassium unweathering (14), lead barium weathering (26) and lead barium unweathering (23). Here, the detected data at its unweathered point in the weathering sample is regarded as unweathered, such as 49 unweathered points in the weathering sample of lead barium is regarded as unweathered category. The mean and standard deviation of each chemical component content of weathered and unweathered glass relics were analyzed respectively.

According to the weathered data $CLR(X)$, the pre-weathering data of high potassium and lead barium glass were calculated by the formula

$$CLR(Y) = \mu_y + \frac{\sigma_y}{\sigma_x} (CLR(X) - \mu_x).$$

The inverse transformation formula (3) of the central log-ratio transformation is then used to restore the unweathered component data (the result is multiplied by 100). The predicted contents of each chemical component of high-potassium glass and lead-barium glass before weathering were calculated, as shown in **Table 5** and **Table 6**.

In order to intuitively understand the change trend of the predicted data, the calculation results are visualized as shown in **Figure 6** and **Figure 7**. As can be seen from **Figure 6**, weathering will cause the proportion of SiO₂, CuO and P₂O₅

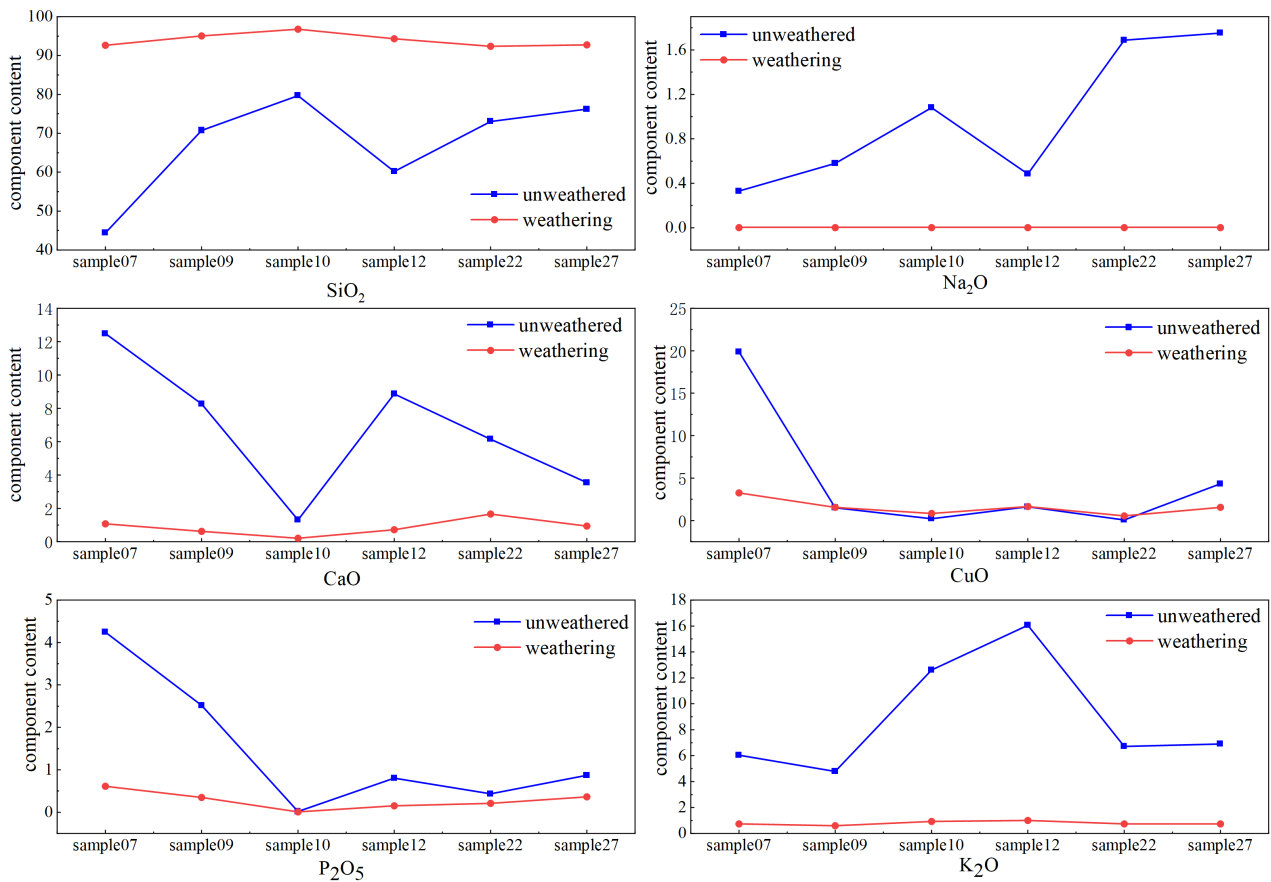


Figure 6. Visualization of composition prediction of high-potassium glass before weathering (part).

Table 5. Prediction data of high-potassium glass before weathering.

Number	SiO ₂	Na ₂ O	K ₂ O	CaO	MgO	Al ₂ O ₃	Fe ₂ O ₃	CuO	PbO	BaO	P ₂ O ₅	SrO	SnO ₂	SO ₂
07	44.379	0.330	6.038	12.489	0.416	10.691	0.051	19.883	0.738	0.430	4.242	0.101	0.122	0.091
09	70.734	0.579	4.776	8.273	0.545	7.930	1.066	1.490	1.040	0.626	2.517	0.125	0.179	0.121
10	79.663	1.080	12.603	1.304	0.172	2.257	0.763	0.201	0.947	0.638	0.016	0.076	0.188	0.092
12	60.190	0.485	16.062	8.863	1.368	8.149	0.609	1.631	0.915	0.547	0.805	0.114	0.156	0.107
22	73.057	1.689	6.705	6.144	1.437	6.927	3.550	0.051	0.002	0.001	0.436	0.001	0.000	0.001
27	76.245	1.751	6.902	3.546	1.047	4.905	0.397	4.331	0.002	0.001	0.871	0.001	0.000	0.001

components to increase, and the proportion of Na₂O, CaO and K₂O components to decrease. As can be seen from **Figure 7**, weathering causes the proportion of SiO₂ and Al₂O₃ in lead-barium glass to decrease, while the proportion of CuO, CaO and PbO increases, while BaO does not change significantly before and after weathering. It is highly consistent with the above linear regression analysis results of weathering state and chemical composition, indicating that the prediction model constructed has a good prediction effect.

Table 6. Prediction data of lead-barium glass before weathering.

Number	SiO ₂	Na ₂ O	K ₂ O	CaO	MgO	Al ₂ O ₃	Fe ₂ O ₃	CuO	PbO	BaO	P ₂ O ₅	SrO	SnO ₂	SO ₂
2	65.426	0.032	2.272	0.754	1.288	6.526	1.860	0.133	20.535	0.795	0.290	0.074	0.010	0.007
8	55.836	0.039	0.027	0.461	0.016	2.746	0.005	6.771	17.097	16.278	0.354	0.198	0.012	0.158
08(1)	34.527	0.056	0.041	2.109	0.025	3.667	0.007	3.036	28.779	25.406	1.256	0.453	0.018	0.620
11	65.913	0.026	0.442	1.175	0.768	3.841	0.004	2.501	12.938	11.403	0.822	0.150	0.009	0.008
19	61.963	0.027	0.022	0.908	0.652	4.803	1.163	1.798	19.630	8.172	0.776	0.069	0.010	0.008
26	55.735	0.041	0.028	0.470	0.017	1.730	0.005	7.102	17.801	16.345	0.315	0.261	0.013	0.138
26(1)	31.199	0.044	1.499	1.594	0.023	3.843	0.006	3.250	26.854	29.639	0.868	0.485	0.017	0.677
34	65.491	0.042	0.604	0.185	0.015	2.676	0.424	0.877	21.210	8.334	0.022	0.103	0.011	0.007
36	19.351	71.837	0.091	0.016	0.004	0.728	0.071	0.105	5.324	2.442	0.001	0.027	0.003	0.002
38	29.222	54.301	0.011	0.064	0.007	1.740	0.105	0.188	10.226	4.020	0.014	0.094	0.005	0.003
39	59.025	0.073	0.032	0.511	0.019	1.298	0.006	0.666	30.575	7.158	0.128	0.488	0.015	0.007
40	49.881	0.088	0.037	1.316	0.021	1.354	0.239	0.008	38.491	7.651	0.242	0.645	0.017	0.008
41	52.508	0.023	0.954	1.758	2.926	5.019	1.597	0.096	21.887	12.398	0.621	0.195	0.010	0.009
43(1)	43.352	0.047	0.030	3.603	1.325	4.294	0.894	3.846	32.301	9.851	0.000	0.434	0.014	0.009
43(2)	56.822	0.028	0.024	3.042	1.121	5.129	1.313	0.831	22.380	7.789	1.286	0.216	0.010	0.009
48	61.312	14.207	0.388	0.384	0.951	8.367	0.446	0.003	5.797	7.164	0.036	0.050	0.890	0.006
49	61.202	0.023	0.021	1.579	1.506	6.434	2.554	0.336	16.276	8.922	0.947	0.183	0.009	0.008
50	54.404	0.037	0.028	1.387	0.657	3.630	0.288	0.721	24.493	13.264	0.681	0.390	0.012	0.009
51(1)	58.387	0.020	0.020	0.952	1.199	6.477	0.882	0.639	18.779	11.388	0.613	0.140	0.495	0.008
51(2)	56.762	0.066	0.033	4.597	2.263	4.554	0.518	0.575	28.459	0.830	1.316	0.004	0.015	0.008
52	36.302	41.127	0.014	0.378	0.382	1.333	0.090	0.220	13.461	6.277	0.290	0.115	0.006	0.005
54	54.777	0.030	0.734	1.143	1.476	5.702	0.004	0.451	25.497	9.345	0.361	0.462	0.010	0.008
54(1)	51.300	0.093	0.038	0.001	2.009	6.440	0.008	1.205	34.083	0.794	2.796	1.208	0.018	0.008
56	62.829	0.082	0.034	0.637	0.019	3.436	0.007	0.622	22.938	9.034	0.336	0.004	0.016	0.007
57	58.655	0.118	0.039	0.982	0.022	4.101	0.008	1.067	26.280	8.697	0.001	0.005	0.019	0.007
58	62.712	0.020	0.648	0.919	0.781	4.585	0.589	1.428	17.577	9.970	0.678	0.077	0.008	0.008

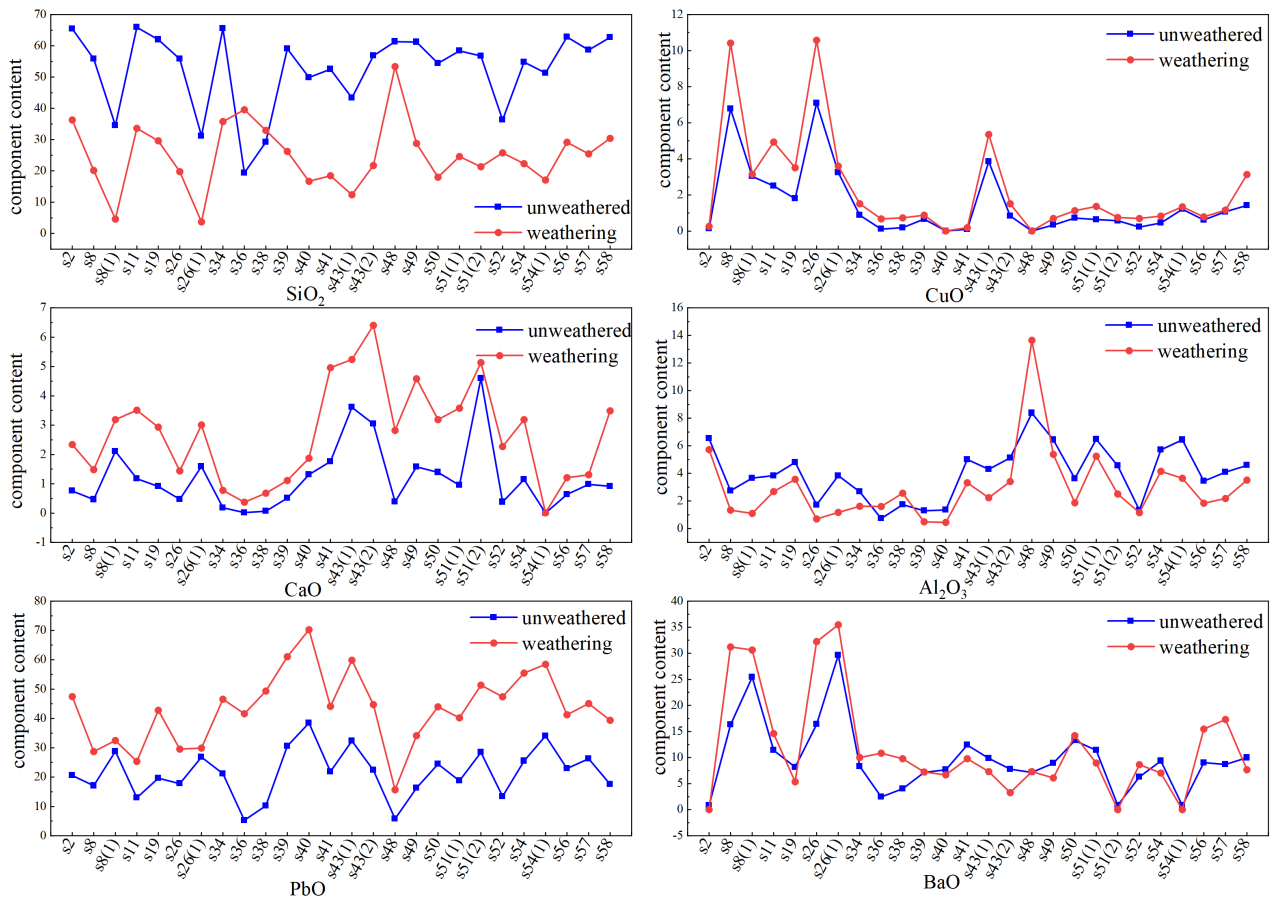


Figure 7. Visualization of composition prediction of lead-barium glass before weathering (part).

4. Conclusion

In this paper, a distributed matching prediction model is proposed. Firstly, the correlation between the chemical components of the glass relics is analyzed by using Pearson correlation coefficient, and the difference of the correlation between the chemical components of different types of glass relics is compared. The correlation between the variables of different glass types is quite different. For high potassium glass, SiO₂ is positively correlated with K₂O, CaO, MgO, Al₂O₃, CuO, P₂O₅ and SnO₂, and negatively correlated with Na₂O and SrO₂. For lead-barium glass, SiO₂ is positively correlated with Na₂O and Al₂O₃, and negatively correlated with P₂O₅ and SrO. It can also be concluded that for high potassium glass, SiO₂ has a positive correlation with P₂O₅, and a negative correlation with Na₂O, while the lead barium glass is the opposite. Secondly, the influence of weathering state on the content of various chemical components of different glass types was investigated by linear regression analysis method. The degree of influence of weathering on the contents of chemical components of different glass types is different. Weathering can increase the proportion of SiO₂, CuO and P₂O₅, and decrease the proportion of Na₂O, CaO and K₂O. At the same time, weathering can cause the proportion of SiO₂ and Al₂O₃ to decrease, while the

proportion of CuO, CaO and PbO to increase. Finally, the prediction model of each chemical component content of glass relics before weathering is constructed. The results of model prediction are highly consistent with those of linear regression analysis, which indicates that the prediction model based on distribution matching proposed in this paper has good validity.

Acknowledgements

This paper is the research result of the project “Multi-field coupling numerical simulation of the formation process of hidden ore body”. Project number: 2021AC19224.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Gan, F.X. (2006) Ancient Silk Road and Ancient Chinese Glass. *Chinese Journal of Nature*, **28**, 253-260.
- [2] Wang, Y.J. and Zheng, X.D. (2022) Research on the Conservation and Management Strategies of Cultural Relics in Archaeological Excavations. *Chinese National Expo*, **6**, 183-185.
- [3] Gu, L.L. and Zhou, J. (2021) Some Questions about the Study of Ancient Chinese Glass Art. *Art and Design*, **2**, 112-114.
- [4] Cho, T., Mantani, H., Ohta, T. and Li, G. (2019) Evaluation of Cretaceous Hinterland Weathering and Climate in the Sichuan Basin, SW China. *Open Journal of Geology*, **9**, 696-699. <https://doi.org/10.4236/ojg.2019.910078>
- [5] Tognonvi, M., Tagnit-Hamou, A., Konan, L., Zidol, A. and N'Cho, W. (2020) Reactivity of Recycled Glass Powder in a Cementitious Medium. *New Journal of Glass and Ceramics*, **10**, 29-44. <https://doi.org/10.4236/njgc.2020.103003>
- [6] Gan, F.X., Li, Q.H., Gu, D.H., *et al.* (2003) Study on Early Glass Beads Unearthed from Baicheng and Tacheng, Xinjiang. *Journal of the Chinese Ceramic Society*, **31**, 663-668.
- [7] Li, Q.H., Zhang, B., Cheng, H.S., *et al.* (2003) Application of Proton Excited X-Ray Fluorescence Technique to the Composition Analysis of Ancient Chinese Glass. *Journal of the Chinese Ceramic Society*, **31**, 950-954.
- [8] Zhao, H.X., Li, Q.H., Gan, F.X., *et al.* (2007) Proton-Stimulated Fluorescence Analysis of Han Dynasty Ancient Glass Unearthed in Hepu Area of Guangxi. *Nuclear Techniques*, **30**, 27-33.
- [9] Cao, Y.X. and Sui, G.R. (2023) Study on the Prediction and Weathering Degree Measurement Model of Ancient Glass Based on GA-BP. *Software Engineering*, **26**, 12-16.
- [10] Shi, B.M., Zhao, X., Wang, H.T., *et al.* (2023) Category Prediction of Ancient Glassware Based on Multi-Layer Perceptron Networks. *Journal of Lanzhou University of Arts and Sciences (Natural Science Edition)*, **37**, 57-62.
- [11] Lv, F., Fu, H.W. and Liu, C.L. (2023) Composition Analysis and Identification of Ancient Glass Products Based on Machine Learning. *Information & Computer*, **4**,

98-102.

- [12] Zou, Y. (2023) Molecular-Composition Analysis of Glass Chemical Composition Based on Time-Series and Clustering Methods. *Molecules*, **28**, Article 853. <https://doi.org/10.3390/molecules28020853>
- [13] Ferrers, N.M. (1866) An Elementary Treatise on Trilinear Coordinates. Macmillan, London, 31-35.
- [14] Zhou, D. (1998) Statistical Analysis of Geological Composition Data—Difficulty and Exploration. *Earth Science*, **23**, 147-152.
- [15] Aitchison, J. (1986) The Statistical Analysis of Compositional Data. Chapman and Hall, London, 66-72.
- [16] Fu, Y.L. (2019) Study on Composition Data Processing Methods: A Case Study of 1:200, 000 Lithic Geochemical Data in Jianshan-Pingkouxia Area, Gansu Province. Ph.D. Thesis, Chang'an University, Xi'an.
- [17] Yang, B. (2015) Comparison of Several Methods of Normality Test. *Statistics and Decision*, **14**, 72-74.
- [18] Lu, Y.L., Huang, C.M., Liu, C.W., *et al.* (2018) In-Door Positioning Algorithm of Wireless Local Area Networks Position Fingerprint Based on Skewness-Kurtosis Test. *Science Technology and Engineering*, **18**, 1-6.