Composition Analysis and Category Prediction Model of Ancient Glass Based on Decision Tree

Jieyin Chen

School of Management, Shenzhen University, Shenzhen, China

2020044016@email.szu.edu.cn

Abstract. Ancient glass is easily affected by environmental factors, which cause chemical changes between glass composition and environmental substances and leads to weathering. It is important to analyze the chemical composition of ancient glass for the study of the Silk Road. To study the composition of glass cultural relics, the relationship between surface weathering and glass type, ornamentation and colour was analyzed first, and then qualitative analysis was carried out through a frequency distribution map, and grouping was carried out according to glass type and surface weathering type, the statistical distribution law of each chemical composition was studied, and the loss of chemical composition with or without weathering was analyzed; Finally, the chemical composition content of the weathered sample before weathering is predicted according to the chemical composition ratio of the weathered and non-weathered cultural relics. The experimental results show that the model has good robustness and can be used to prepare and quickly predict the composition and classification of ancient glasses.

Keywords: Ancient glass, Silk Road, weathering, chemical composition, qualitative analysis.

1. Introduction

The main raw material of glass is quartz sand, and the difference in the chemical composition of glass relics belonging to different glass systems affects the composition and change of chemical composition during burial, resulting in different degrees of weathering [1]. Ancient glass is highly sensitive to temperature, humidity and environmental atmosphere (environmental factors including water and carbon dioxide, which react with glass chemicals). And other factors, cause chemical changes between the glass composition and the surrounding environmental substances, resulting in weathering [2]. Among them, the higher the temperature, the heavier the degree of weathering, and different humidity under different acid gas ratios will also cause different degrees of weathering [3]. From the development period and region of glass types, the potash glass system was developed in the Lingnan area in the middle and late Western Han Dynasty. From the end of the Spring and Autumn Period to the beginning of the Warring States Period to the Western Han Dynasty, lead-barium glass was mainly found in Chu tombs in Hunan Province and extended to the northern region [4]. It is important to analyze the chemical composition of ancient glass for the study of the Silk Road.

In this paper, a comprehensive prediction model for ancient glass relics is established, first of all, qualitative analysis is carried out to judge the factors affecting the surface weathering of glass relics, then the loss of components with or without weathering is analyzed, the distribution of chemical components of glass types is studied, and finally to improve the prediction accuracy. The chemical composition content of the weathered sample before weathering was predicted by the chemical composition ratio within the group with and without weathering. The results show that the model is robust and can be used to analyze the chemical composition of ancient glasses.

2. Model establishment

The model first judges whether there is weathering according to data 1, combines the cultural relic number according to data 2, judges the relationship between weathering and glass type, and then carries out a chemical composition analysis. The flow chart is shown in Figure 1.

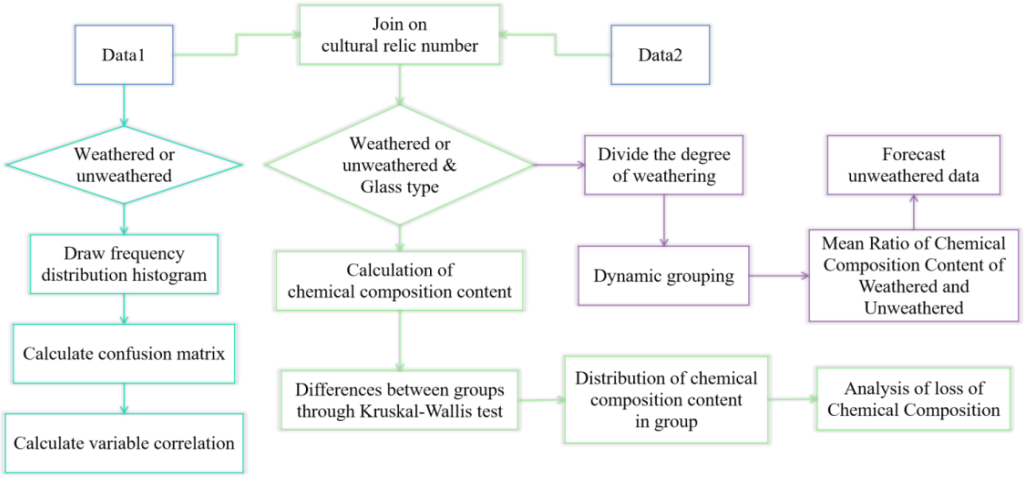


Figure 1. Model flow chart

2.1. Analysis of the relationship between glass surface weathering and glass properties

The relationship between glass surface weathering and texture, glass type and colour was studied, the weathering frequency distribution of each category in the three attributes was counted, and the surface weathering proportion of each category was observed. Since the three attributes and surface weathering variables are discrete variables, the traditional Spearman and Pearson correlation coefficients cannot be used to describe the correlation, so the Kendall correlation coefficient is considered in this paper. Firstly, the three attributes are treated as dummy variables to facilitate the calculation of the confusion matrix, and then the Kendall correlation coefficient between each category and weathering is calculated, and the thermodynamic map is drawn to observe the cultural relics category that each attribute of glass is most likely to be marked as weathering, and finally, the surface weathering characteristics of each category are analyzed and confirmed by combining with the literature [5].

Draw the frequency distribution map and thermodynamic map of surface weathering under different textures, as shown in Figure 2, where the sum of weathering frequencies of all textures is 1, and the sum of non-weathering frequencies is 1.

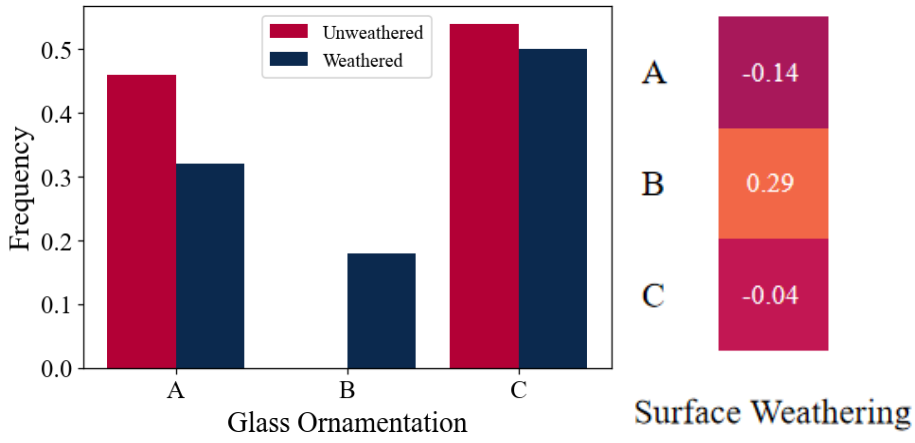


Figure 2. The relationship between different ornamentation and weathering

According to the analysis [6], both ornamentation A and ornamentation C have the same proportion of weathering and non-weathering, and the surface weathering of ornamentation B is marked as weathering, indicating that the cultural relics of ornamentation B are the easiest to weather. Further analysis shows that the proportion of weathering in ornamentation C is higher than that in ornamentation A, indicating that the cultural relics of ornamentation C are more easily weathered than those of ornamentation A. Looking at the thermodynamic diagram, ornamentation B is more highly correlated with weathering, followed by ornamentation C and finally ornamentation A.

According to the literature [7], different regions have different preferences for ornamentation, and regional environmental atmosphere differences have a certain degree of impact on glass surface weathering, so different ornamentation of glass may cause different proportions of surface weathering.

The frequency distribution and thermal map of surface weathering under two different glass types are drawn, as shown in Figure 3, where the sum of weathering frequencies of all glass types is 1, and the sum of non-weathering frequencies is 1.

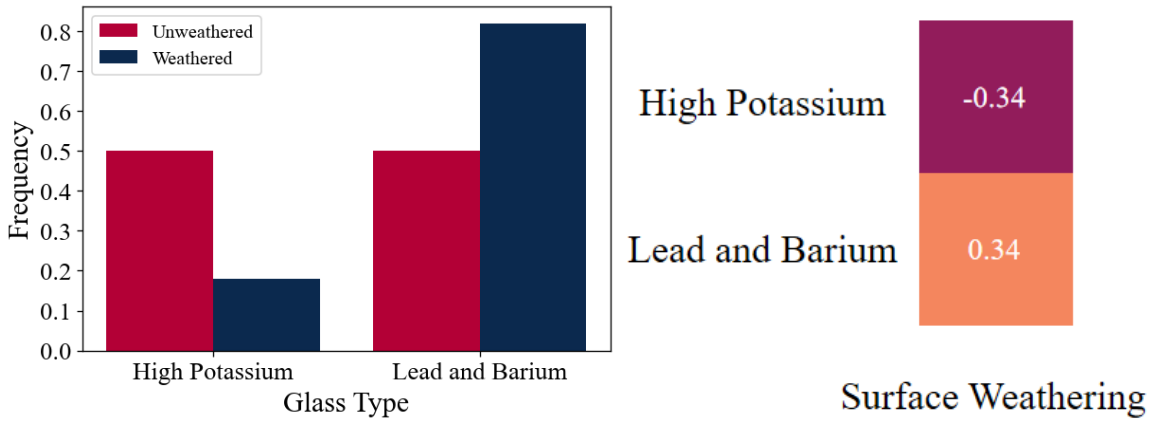


Figure 3. Relationship between different glass types and weathering

Through the analysis, it can be seen that there are different proportions [8] of weathered and unweathered marks in the two kinds of glass, and the proportion of cultural relics marked as weathered in lead-barium glass is much larger than that in high-potassium glass, which indicates that lead-barium glass is easier to weather than high-potassium glass to some extent. The Kendall correlation coefficient for Pb-Ba glass and weathering is 0.34. The Kendall correlation coefficient between high-K glass and weathering is -0.34, indicating that Pb-Ba glass is easier to weather than high-K glass. According to the literature, the weathering corrosion of ancient Chinese glass is concentrated on the glass with high lead content, in which the glass with high lead content is very easy to corrode in the environment, which confirms the analysis result that the proportion of lead-barium glass relics marked as weathering is high.



Figure 4. Thermodynamic map of different glass colors and weathering

The Kendall correlation coefficient between green, dark blue and weathering is smaller, indicating that the surface weathering of the cultural relics with this colour is less, and the Kendall correlation coefficient between black and weathering is the largest, indicating that the cultural relics with black colour are more marked as weathering. In essence, different colors of glass reflect the chemical substances corresponding to different colorants of glass, and different chemical compositions may cause different surface weathering of glass. As shown in Figure 4.

2.2. Statistics of chemical composition content by type

Based on the glass type and weathering type [9], explore the statistical law of chemical composition content in data 2. Firstly, the glass type and weathering type are cross-divided into four categories, namely, high-potassium non-weathering, high-potassium weathering, lead-barium non-weathering and lead-barium weathering. Then the mean value and variance are counted respectively, as shown in Figure 5 and Figure 6.

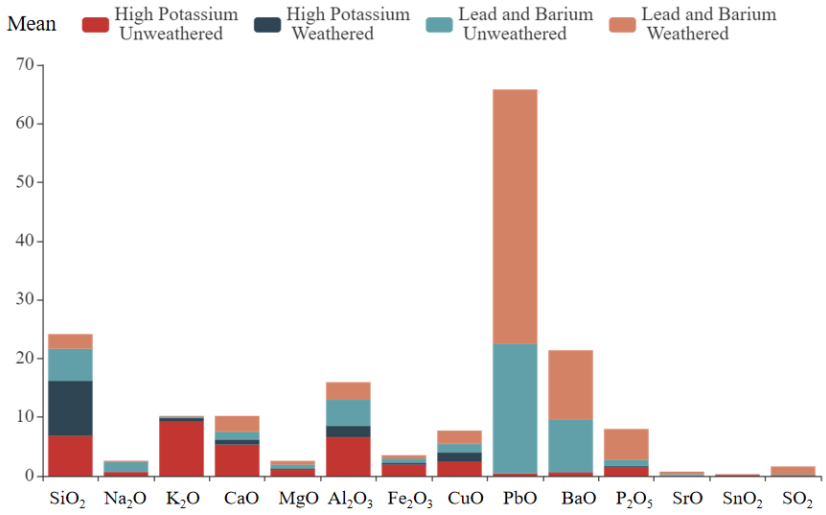


Figure 5. Mean values of chemical composition in different groups

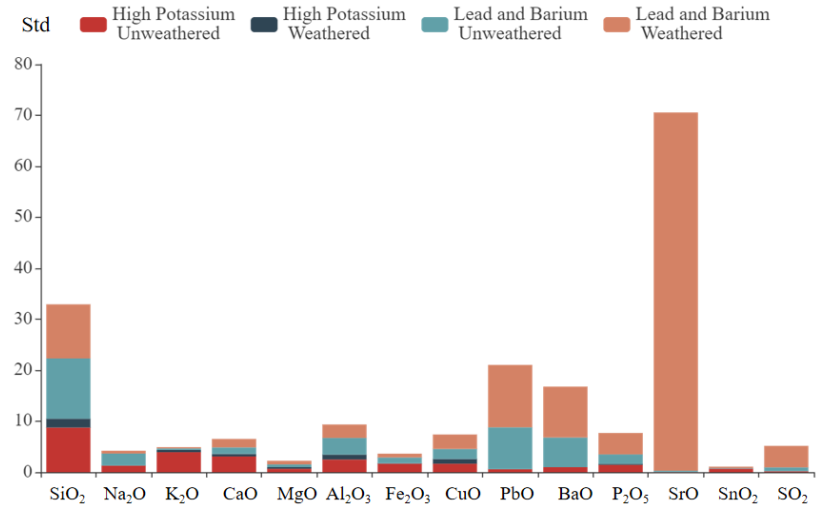


Figure 6. Mean variance of chemical components in different groups

Since the content of SiO_2 is much higher than that of other chemicals, the percentage of SiO_2 content is multiplied by one-tenth to allow simultaneous observation of the mean and variance of chemical content in different categories. It can be seen from the Figure 5 that under the four groups, the content of SiO_2 is the most, in which the high-potassium type is greater than the lead-barium type, and the high-potassium weathering is greater than other types. At the same time, the mean value of each chemical element in different groups is quite different, but the standard deviation is not necessarily small. Therefore, the Kruskal-Wallis test is carried out for each chemical element to verify the difference in each chemical component content in each group, as shown in Table 1.

Table 1. Results of the Kruskal-Wallis test

Analysis item	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	F e_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
P	0.000 ***	0.028 **	0.000 ***	0.000 ***	0.038 **	0.000 ***	0.025 **	0.075 *	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.813	0.286
Cohen's f Value	0.288	0.099	0.261	0.187	0.12	0.143	0.133	0.049	0.255	0.169	0.144	0.176	0.053	0.054

Note: * * *, * *, * represents the significance level of 1 %, 5 %, 10 % respectively

Among them, Cohen's f value represents the effect size, which is used to measure the difference between groups. It can be seen that most of the chemical components show significant differences in different groups, only copper oxide, tin oxide, and the sulfur dioxide P value is greater than 5%, showing no signs. Therefore, when we study the law of chemical composition with or without weathering, we should discuss different glass types separately.

To further observe the distribution law of the content of each chemical component in different groups [10], it is considered that the chemical components are drawn in a box and displayed in a chart to facilitate the internal comparison of the groups. It can be seen from Table 1 that among the chemical components, SiO_2 is the main component of the glass, and its magnitude is larger than that of other components, so it is necessary to draw a box diagram separately, as shown in Figure 7; while potassium oxide, lead oxide and barium oxide are the main substances of the two glass types, and their proportions in each category are analyzed separately, as shown in Figure 7.

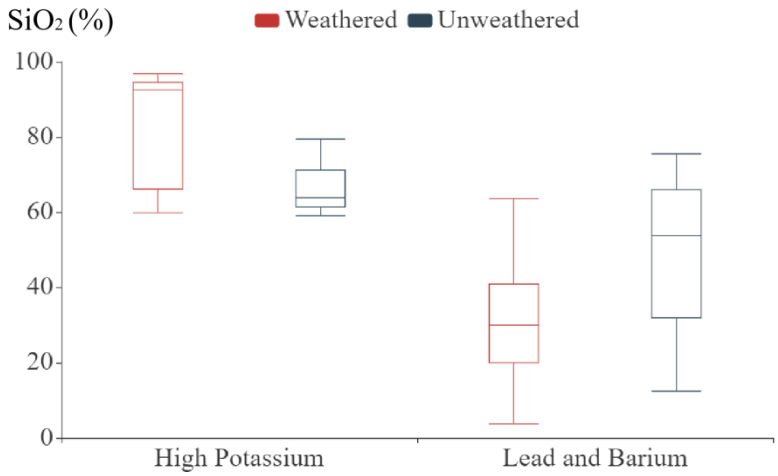


Figure 7. Silica content of each group

Observing the box plots of the four groups of silica in figure 8. The average content of silica is more than 20%, and the average value of high-potassium glass is higher than that of lead-barium glass, the average value of high-potassium weathered glass is higher than that of high-potassium unweathered glass, and the average content of lead-barium weathered glass is lower than that of unweathered glass. From the fluctuation of the data, the content of silica in the high-potassium weathering category is the most stable in the sampling sites of cultural relics, while other types of data show different degrees of dispersion.

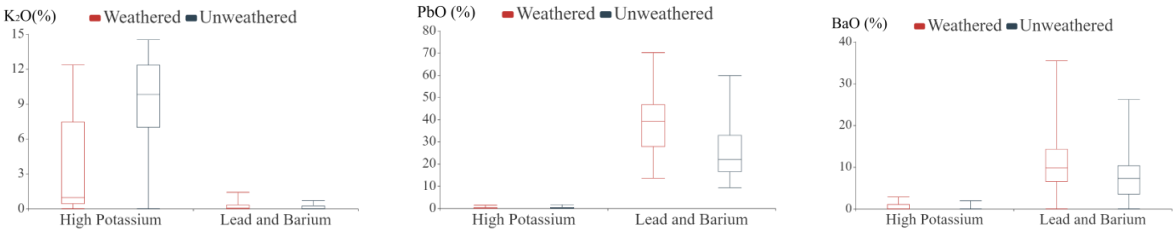


Figure 8. Box plot of K2O, PbO and BaO in different types

Similarly, for high-K glass, the content of sodium oxide, calcium oxide, magnesium oxide, aluminium oxide, iron oxide, copper oxide and phosphorus pentoxide decreases as a whole after weathering. For lead-barium glass, the content of sodium oxide and aluminium oxide decreases after weathering; the content of calcium oxide, copper oxide and phosphorus pentoxide increases. Before weathering, the content of calcium oxide and aluminium oxide in high-potassium glass is higher than that in lead-barium glass; after weathering, the content of calcium oxide, copper oxide and phosphorus pentoxide in lead-barium glass is higher than that in the high-potassium glass.

2.3. Pre-weathering chemical composition prediction based on group mean

Because of the different decorative patterns, colorants and fluxes used in cultural relics, the different colours and glass types lead to different weathering degrees of cultural relics after a long time. Therefore, in this paper, the physical attributes of a class of cultural relics are fixed by dividing the cultural relics into groups, to ensure the consistency of the attributes of cultural relics to a certain extent, and then the chemical composition changes of the cultural relics before and after weathering

in the unified group are consistent, and then the average ratio of the weathered and unweathered chemical composition of the group is calculated. The chemical composition of the weathering sample is multiplied by the mean ratio to predict the chemical composition content before weathering.

In this paper, the weathering degree of cultural relics is marked first. It can be seen from the above that silica is one of the main components of glass, and its component content has the highest effect value in the differences of different groups, and it is a medium effect index. Therefore, the component content of silica is used as an index to divide the weathering degree. Because of the small within-class variance of the high-K weathering categories in the above analysis and the overall concentration of silica, we focused only on lead. Barium weathering samples are marked with weathering degree. In this paper, the lower quartile and the upper quartile are taken as the threshold of severe weathering and shallow weathering respectively, and the weathering degree is divided. The division results are shown in Table 2.

Table 2. Weathering labels for different glass types

	No weathering	Weathering		
High potassium glass	0	1		
Lead-barium glass	0	1	2	3

Let $w(SiO_2)$ be the content of SiO_2, the label is 1 when $w(SiO_2) \le 30.2$ $18.2 < w(SiO_2) \le 30.2$, the label is 3 when $w(SiO_2) \le 18.2$.

After the weathering degree is calculated, the data will be grouped according to the four physical attributes of ornamentation, colour, type and weathering degree, and compared with the data grouped by weathering degree alone. The variance within the group of each chemical component will be calculated to prove the validity of the grouping. The colour attribute of the sample with missing colour will not be considered. The grouping results are shown in Table 4, and the mean change of variance within the group is shown in Figure 9.

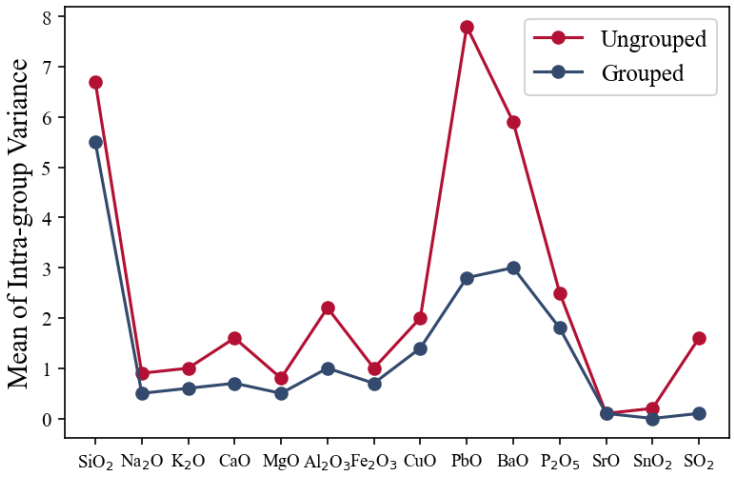


Figure 9. Variance mean curve of each chemical component before and after grouping

It can be seen from the figure 9 that after grouping the physical attributes, the average within-group variance of each chemical component is reduced.

Table 3. Grouping

	Ornamentation	Color
Lead-barium glass	A	Light blue, dark blue, bluish green, black, none
	C	Light green, light blue, dark green, purple, green, bluish green, none
High potash glass	A	Dark blue, bluish green
	B	Blue and green
	C	Light blue, dark green, bluish green

By observing the number of samples in the group after Table 3, we find that for each weathered sample, under the corresponding type, we may not be able to find the same sample as its decoration and colour. Therefore, we do not use the method of grouping first and then predicting, but when predicting the chemical composition of each sample before weathering, we use the method of "maximum grouping" to screen the samples for calculating the mean ratio. For example, the sample with the cultural relic number 23 has the type of lead and barium, the colour of blue and green, and the ornamentation of A, but there is no non-weathered sample with the same type, colour and ornamentation. Therefore, we will screen out the non-weathered samples with the type of lead and barium and the ornamentation of A to calculate the mean value, and then calculate the mean value ratio to the weathered samples under the same conditions for prediction.

Table 4. Predicted results of chemical composition before weathering at the weathering point

Cultural relics sample	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
02	53.01	1.68	0.70	1.55	1.25	5.70	1.99	0.23	27.63	9.00	1.10	0.19	0.05	0.16
07	67.02	0.70	9.33	6.56	1.08	6.79	1.24	5.09	0.41	0.60	3.06	0.04	0.20	0.10
08	44.80	1.68	0.22	0.68	0.64	2.24	0.74	5.39	14.57	23.63	0.65	0.28	0.05	1.09
08 Severe Weathering Point	19.38	1.68	0.22	1.37	0.64	2.50	0.74	2.13	14.80	18.55	1.28	0.21	0.05	0.54
09	68.75	0.70	10.13	3.80	1.08	4.53	2.33	2.43	0.41	0.60	1.75	0.04	0.20	0.10

The chemical composition of the final predicted weathering point before weathering is shown in Table 4.

3. Conclusion

In this paper, a comprehensive prediction model for ancient glass relics is established, first of all, qualitative analysis is carried out to judge the factors affecting the surface weathering of glass relics, then the loss of components with or without weathering is analyzed, the distribution of chemical components of glass types is studied, and finally in order to improve the prediction accuracy. The chemical composition content of the weathered sample before weathering was predicted by the chemical composition ratio within the group with and without weathering. The results show that the model is robust and can be used to analyze the chemical composition of ancient glasses.

The model fully considers the relationship between the physical properties and chemical composition of glass relics, as well as the correlation between chemical compositions, and has strong interpretability. Prediction accuracy: The sampling points are innovatively grouped according to the physical properties of glass, which improves the accuracy of chemical composition prediction before weathering. However, due to the small amount of data available in the data set, the accuracy of the prediction model may not be high, so the data set should be expanded in the later period to improve the accuracy of the prediction model.

References

[1] Han Hongchang. Alumina is used as a raw material in the glass industry. Light Metals, 6, 1981.

[2] Conscience Zhou. The main factors affecting the weathering of silicate glass. Journal of Dalian Institute of Light Industry, 1, 1984.

[3] An Jiayao. Early Chinese glassware. Acta Archaeologica, (4): 413 – 448, 1984. 4.1 [4] Shi Meiguang, He Ouli, Zhou Fuzheng. Study on Potash Glass Unearthed from a Number of Han Tombs in China. Acta Silicate Sinica, 14 (03): 307 – 313, 1986.

[4] Cui Jianfeng, Wu Xiaohong, Tan yuanhui, Wang Yongbiao. Composition Analysis of Ancient Glassware Unearthed from Chu Tombs of the Warring States Period in the yuanshui Valley of Hunan Province. Acta Silicata Sinica, 37 (11): 1909 – 1913, 2009.

[5] Shao Hui, Zhou Hongqing, Liu Min, Zhu Haikui, Shen Xiaodong. Effect of alumina addition on barium silicate composite glass ceramics. Chinese Journal of Ceramics, 32 (1): 17 – 20, 2011.

[6] Wang Shengtie, Liu Yutang. Summary of Chu Culture. BEIJING BOOK CO. INC., 2013.

[7] Liu Song, Lv Liangbo, Li Qinghui, Xiong Zhaoming. Glass Beads Unearthed from Han Tombs in the South of the Five Ridges and the Sino-Foreign Exchange on the Maritime Silk Road of the Han Dynasty. Conservation and Archaeological Science, 31 (4): 18 – 29, 2019.

[8] Geological Dictionary Office, Ministry of Geology and Mineral Resources. Geological Dictionary: Geological Publishing House, 2005.

[9] Wei Jie. Research on improved K-means clustering algorithm [J]. Information Communication, 2018 (5): 14-15.

[10] Chen Min. Factors affecting the weathering of lead glass [D]. Dalian University of Technology, 1987.