Prediction Model for Various Elements of Glass Artifacts Based on Probability Distribution

Haiping Qiu[1, *], Wenjie Zhu[2], Sirun Zhu[1]

[1] School of Ocean Engineering, Harbin Institute of Technology, Weihai, China

[2] School of Information Science and Engineering, Harbin Institute of Technology, Weihai, China

* Corresponding Author Email: qhp13860002649@163.com

Abstract. During the weathering process of glass, its internal and external environments undergo complex exchange of elements, resulting in the possibility that archaeologists may misjudge the category of glass due to the changes in the percentage of its components. In this paper, the data used is from the C problem of the 2022 Contemporary Undergraduate Mathematical Contest in Modeling to fit the probability distribution of each component in glass using a logistic model, and established a prediction model based on the probability distribution. The results demonstrate that the model has a good interpretation in the prediction of the chemical composition content of glass.

Keywords: Logistic model, Glass artifacts, Probability distribution, Markov model.

1. Introduction

Glass was active on the Silk Road as important physical evidence of trade and cultural exchange between China and the West in ancient times, and ancient Chinese glass had different chemical composition and similar appearance under the influence of Western glass technology. Quartz sand is the main raw material of glass, its chemical composition is SiO_2, in order to reduce the melting temperature of pure quartz sand, often added to the refining of grass wood ash, lead ore and other fluxes. The chemical composition of different fluxes also varies, resulting in two types of lead-barium glass and potassium glass. Affected by the burial environment, ancient glass has different degrees of weathering. During this process, complex elemental exchanges between its internal and external environments occurred, resulting in the possibility of archaeological staff making erroneous judgments about the glass categories due to changes in their composition ratios.

Thus, many scholars have done research in this area. p-XRF devices were investigated by Yatsuk Oleh et al. for the detection of some major, minor and trace elements associated with the intentional addition of specific ingredients or impurities in the raw material of glass batches [1]. Rossano S et al. found variations of iron elements in manganese glass [2]. Medeghini Laura et al. considered elements that have an element that have a prominent influence on the glass composition and thus finely differentiate the compositional groups [3]. Zhou Xueqi et al. investigated new information about Chinese glassmaking recipes from the 10th to the 12th century and further elucidated the development of glass production in ancient China [4]. Mariangela Vandini et al. proposed a method based on easily accessible and widely used techniques (e.g., OM, SEM-EDS, μ Raman and XRPD) based on easily accessible and widely used techniques (e.g., OM, SEM-EDS, μ-Raman, and XRPD) for achieving in-depth characterization of coloring and opaque inclusions [5]. Saminpanya Seriwat et al. investigated the oxidation state of coloring elements and pigments in ancient rare glasses and the effect on glass color [6]. Tomomi Tamura et al. classified natron glasses into seven major types based on chemical composition and other minor types [7].

The main purpose of this paper is to analyze the statistical pattern of the chemical composition content of the surface of cultural relics with and without weathering, and to predict their chemical composition content before weathering based on the detection data of weathering points.

2. Model building and Solving

2.1. Method description

2.1.1 One-way ANOVA models

The basic idea of ANOVA is to determine whether controllable factors have a significant influence on the results of the study by analyzing the magnitude of the contribution of the variance from different sources to the total variance of the study. One-way ANOVA considers only the effect of an indicator due to one factor, keeping all other factors that affect the indicator constant.

A one-way variance model was developed in this paper. It is verified whether the degree of weathering on the surface of glass artifacts has a significant effect on glass type, color, and decoration, respectively.

1) Experimental group set up

Let factor B take m levels, denoted as B_1, B_2… B_m, where the factor has two levels. Under level B_i, the overall satisfaction $X_i \sim N(\mu_i, \sigma^2)$, (i = 1, 2..., m), and assume that X_i has the same variance. Conduct c_i independent trials at level B_i and record the experimental results as X_{ij}, and include them in Table 1.

Table. 1. one-way ANOVA experiment table

B_1	X_{11}	X_{12}	$\dots$	X_{1n_1}
B_2	X_{21}	X_{22}	$\dots$	X_{2n_2}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
B_m	X_{m1}	X_{m2}		X_{mn_m}

2) Hypothesis building

To examine whether the two levels of factors have significant effects on the type, color and decoration, the original hypothesis H_0 is as follows:

$\mu_1 = \mu_2 = \cdots \mu_m$;

The alternative hypothesis is:

$\mu_1, \mu_2, \cdots \mu_m$ are not all equal.

3) Factor level

Degree of weathering: Weathered or unweathered.

4) Model construction

According to the above assumption, there is $X_{ij} = \mu_i + \varepsilon_{ij}$, that is $\varepsilon_{ij} = X_{ij} - \mu_i \sim N(0, \sigma^2)$, ε_{ij} are random errors and are independent of each other.

Introduce variables:

$$\mu = \frac{1}{n}\sum_{i=1}^{m} n_i \mu_i, \quad n = \sum_{i=1}^{m} n_i, \quad \alpha_i = \mu_i - \mu, \text{ i=1,2,…,m} \tag{1}$$

In the above equation, μ is the average value of the whole, and α_i is the effect under the level of weathering degree B_m. A mathematical model is then constructed that can be expressed as:

$$\begin{cases} X_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \\ \sum_{i=1}^{m} n_i \alpha_i = 0, \\ \varepsilon_{ij} \sim N(0, \sigma^2), i = 1, \cdots, m, j = 1, 2, \cdots, n_i \end{cases} \tag{2}$$

In order to analyze the effects at different levels from the deviations of observations, three deviations are introduced:

$$X_{ij}-\overline{X},\ \ \overline{X}_i-\overline{X},\ \ X_{ij}-\overline{X}_i \tag{3}$$

The sum of squares of these three deviations has the following theorem:

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X})^2=\sum_{i=1}^{m}n_i(\overline{X}_i-\overline{X})^2+\sum_{i=1}^{m}\sum_{j=1}^{n_1}(X_{ij}-\overline{X}_i)^2 \tag{4}$$

Let

$$S_T=\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X})^2,$$
$$S_A=\sum_{i=1}^{m}n_i(\overline{X}_i-\overline{X})^2, \tag{5}$$
$$S_E=\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2,$$

The above formula can thus be expressed as follows:

$$S_T=S_A+S_E \tag{6}$$

S_T in the above equation is called the sum of squared total deviations, which reflects the degree of variation and dispersion of the total data X_{ij} with respect to $\overline{X}$. Call S_E the error sum of squares (or within-group sum of squares), which reflects only the magnitude of experimental error; S_A is called the effect sum of squares (or between-group sum of squares) of factor B, which contains experimental error in addition to reflecting the degree of dispersion of each level effect of factor B.

Similarly it is easy to prove that under the condition that H_0 holds, $\frac{S_T}{\sigma^2}$, $\frac{S_A}{\sigma^2}$, and $\frac{S_E}{\sigma^2}$ satisfy the cardinality distribution with degrees of freedom mn_i-1, $m-1$ and $m（n_i-1）$, respectively. and $\frac{S_A}{\sigma^2}$ and $\frac{S_E}{\sigma^2}$ are mutually independent.

2.1.2 Quantitative analysis model based on the probability distribution function

Since the dependent variable of the probability distribution function is a value in the range of 0 to 1 and has a monotonically increasing trend in the interval of the independent variable, the inverse function, exponential function and other models are not suitable for fitting, and the distribution of the chemical content is nearly normal, so this paper finally chooses to construct a logistic growth model for fitting.

The specific model is established as follows.

Based on the engineer's principle, the growth rate of the probability distribution is considered as a linear decreasing function with respect to x. Obtain:

$$\begin{cases} x(T_0)=x_0 \\ l(x)=l-sx \quad (l,s>0) \end{cases} \tag{7}$$

where, is the initial growth rate and s being the rate of change of growth rate.

When $x=x_m$, the growth rate $l（x_m）=0$, brought into the above equation, gives:

$$l(x) = l\left(1 - \frac{x}{x_m}\right) \tag{8}$$

Solve to obtain the logistic growth model:

$$x(T) = \frac{x_m}{\left(1 + \left(\frac{x_m}{x_0} - 1\right)e^{-l(T-T_0)}\right)} \tag{9}$$

2.1.3 Prediction model based on homogeneous probability distribution linkage

The conclusion that the probability distribution function of the content of the same chemical components with or without weathering as a variable has the same functional form, combined with relevant information and Markovian ideas, the process of weathering is recorded as a transfer of state, then the change in the percentage of each chemical component constitutes a probability transformation matrix, so this paper defines that the influence of the probability distribution function with or without weathering on the content of the same chemical components has isotropic linearity, that is, the influence of weathering on the content of the same chemical components with or without weathering The positive and negative effects of weathering on the change of the content of the same chemical components are consistent in magnitude.

In this regard, the following proof is given in this paper.

Suppose the following relationship exists under a certain probability distribution function:

$$P(x_1) \leqslant P(x) \leqslant P(x_2) \tag{10}$$

When influenced by a factor with isotropic linearity, it is clear that there is

$$P(kx_1 + b) \leqslant P(kx + b) \leqslant P(kx_2 + b) \tag{11}$$

where the coefficients k, b indicates the magnitude of the effect of the factors.

When there are an infinite number of inequalities connected, i.e.

$$P(kx_1 + b) \leqslant P(kx + b) \leqslant P(kx_2 + b) \tag{12}$$

In a comparative analysis, the change in the function $P(kx+b)$ of each dependent variable compared to $P(x)$ will tend to be infinitesimal. That is, they can be considered equal.

2.2. Build process and validation

In order to visually analyze the statistical law that gives the content of chemical components with and without weathering on the surface of the artifact samples, this paper first draws scatter plots of the content of various chemical components with and without weathering for the given data, respectively, and gives some of the scatter plots with significant differences, as follows.

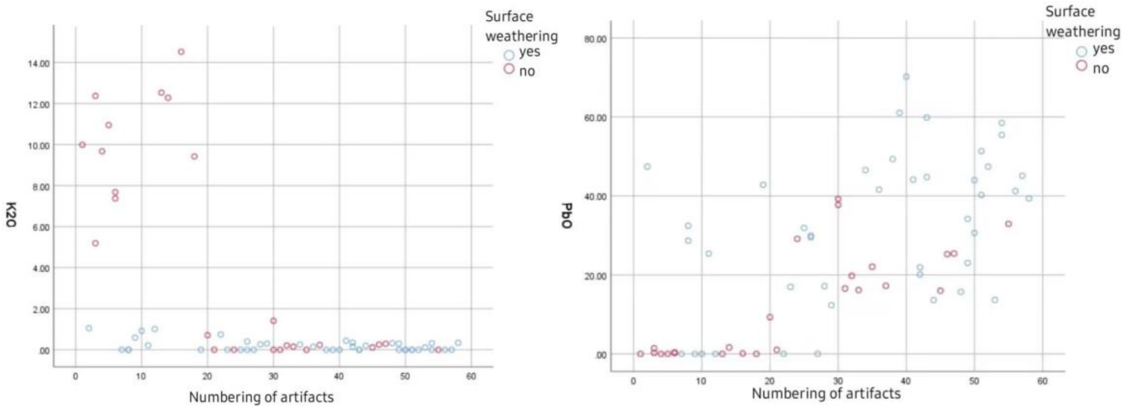

Figure 1. Scatter plot of chemical content with and without weathering

Analysis of the left diagram in Figure 1, such as numbers 1, 3, 4, and 5. Comparison of the data characteristics in Form 1 reveals that all of these glass artifacts are in the category of high potassium, and examination of the context of the problem and the relevant literature suggests a high probability that this is due to the residue of fluxes with high potassium content on the glass surface during the refining process. In contrast, the low potassium oxide content of surface weathered glass artifacts is related to flux loss due to surface weathering [8].

Analysis of the right diagram in Figure 1, the lead oxide content of the glass artifacts of type lead barium is much higher than that of the high potassium type because lead ore was added as a flux during its refining process. And the lead oxide content of the weathered lead-barium glass in the above graph is much higher than that of the unweathered lead-barium glass artifacts, and after reviewing the relevant literature, we learned the reason, the surface weathered lead-barium glass artifacts will form a large amount of lead carbonate in its outer glass crust, which may be generated by the reaction between the large amount of lead in the glass and carbon dioxide and water vapor in the air [9].

A search of the relevant literature combined with analysis of the relevant data led to the following conclusions.

1) During the weathering process of lead barium glass, the elements in the internal and external environment are exchanged in large quantities. The main manifestations are: SiO_2 is lost from inside to outside along the weathering layer, BaO reacts with external elements in the outermost crust to produce $BaCO_3$ and accumulates; S produces SO_2 and accumulates in the outermost layer; MgO and CaO are reduced in the surface layer due to the reaction of carbonate, and the content is higher the closer to the inner layer; CuO, the main color-rendering chemical component, is lost to the outside, while Fe elements accumulate in the outer layer in the form of Fe_2O_3 . The main color-emitting chemical component CuO is lost to the outside, while Fe elements accumulate in the outer layer in the form of Fe_2O; the monomeric P elements are oxidized to produce P_2O_5, K_2O, Na_2O, Al_2O_3, SrO, PbO, SnO_2 with no significant change in content [10].

2) The high potassium glass in the process of weathering shows that most of K2 O in the surface glass is precipitated and lost, SiO_2, Al_2O_3 is significantly increased, Fe_2O_3, Na_2O is generally reduced. Part of SiO_2 forms better weathering resistant silica-oxygen tetrahedra while retaining. MgO, CaO decreases in the surface layer due to the reaction to form carbonates, and its content is higher the closer to the inner layer. CuO, PbO, BaO, P_2O_5, SrO, SnO_2 have no significant changes.

Several single-factor analysis of variance (ANOVA) of weathering degree factors were carried out by using SPSS Software. The results are shown in Table 2:

In Table 2, the test of the effect of whether the surface is weathered on the grain finish was selected as an example, and the F-value of the ANOVA was 3.037, corresponding to a p-value of 0.057, at the level of 0.05, it is considered that there is no significant effect of whether the surface is weathered on the difference of the grain finish. Similarly, it can be seen that there is a significant effect of whether or not surface weathering has a significant effect on the difference in category, as well as no significant effect of whether or not surface weathering has a significant effect on color.

From the results of the one-way ANOVA model, only surface weathering has a significant effect on glass type. Therefore, this paper decided to first classify the data in Form 2 into four categories based on glass type and weathering degree, and use different continuous probability distribution functions to fit the chemical contents of each category separately to give the corresponding numerical analysis.

Using PyCharm software to call the statsmodels library function, the Logistic growth function was selected to fit the probability distribution functions of various chemical contents, and the probability distribution function plots and fitting equations of some chemical contents were given (as shown in Figure 2 and Table 3).

Table. 2 statistical table of analysis of variance results

ANOVA						
Indicators		Sum of squares	Degree of freedom	Mean Square	F	Significance
Color	Between groups	1.552	7	.222	.866	.540
	Of the project	11.781	46	.256		
	Total	13.333	53			
Type	Between groups	1.333	1	1.333	5.778	.020
	Of the project	12.000	52	.231		
	Total	13.333	53			
Decoration	Between groups	1.419	2	.710	3.037	.057
	Of the project	11.914	51	.234		
	Total	13.333	53			

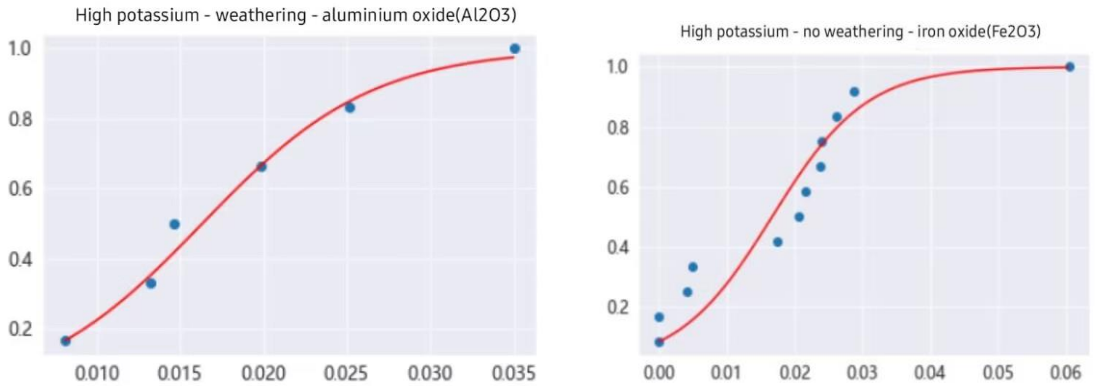

Figure. 2 Partial probability distribution function fit graph

Table. 3. Fitting equations for probability distribution functions

Glass surface weathering		No weathering on the glass surface	
Category-Chemical composition	Fitting equations for probability distribution functions	Category-Chemical composition	Fitting equations for probability distribution functions
High potassium - calcium oxide	$y=\frac{1}{1+\left(\frac{1}{0.1666}-1\right)e^{331.9581(x-0.0021)}}$	High potassium - calcium oxide	$y=\frac{1}{1+\left(\frac{1}{0.0833}-1\right)e^{46.7987x}}$
…	…	…	…
Lead barium-calcium oxide	$y=\frac{1}{1+\left(\frac{1}{0.0384}-1\right)e^{201.2986x}}$	Lead barium-calcium oxide	$y=\frac{1}{1+\left(\frac{1}{0.0434}-1\right)e^{201.2986x}}$

In the fitted model of Figure 2, the goodness of fit of the polynomials are all greater than 90%, which is considered to explain the probability distribution relationship of each chemical composition content under different categories better and verify the accuracy of the model. As for the chemical composition contents with the lack of sample point data, the distribution law cannot be accurately described, and the equation of the fitting function for this category is identified in this paper as $y=0$.

Combined with the analysis of the fitted curve equations in Table 3, a prediction model based on the homogeneous probability distribution was constructed as follows

$$P_1(x_1)=P_2(x_2) \tag{13}$$

Where, $P_1(x)$, $P_2(x)$ are the fitted expressions of the probability distribution function with and without surface weathering for the corresponding chemical composition content of the same category, respectively. By using PyCharm software, the weathering data were substituted to predict the chemical composition content before weathering.

3. Conclusions

This paper explores the statistical law of the chemical composition content of cultural relics with and without weathering on their surfaces, and predicts their chemical composition content before weathering based on the detection data of weathering points. The fitting and prediction results show that the prediction model based on probability distribution is a more ideal and feasible prediction method. The logistic function can describe the probability distribution of chemical content in glass more accurately, and the prediction model based on the probability distribution can unify the influence of different factors and improve the prediction accuracy of chemical content.

References

[1] Yatsuk Oleh; Ferretti Marco; Gorghinian Astrik; Fiocco Giacomo; Malagodi Marco; Agostino Angelo; Gulmini Monica Data from Multiple Portable XRF Units and Their Significance for Ancient Glass Studies [J] Molecules,2022.

[2] Rossano S., Khomenko V., Bedidi A., Muller C., Loisel C., Ferrand J., Sarrasin L., Bertin A. Glass colourations caused by Mn-Fe redox pair: Application to ancient glass technology [J]. Journal of Non-Crystalline Solids,2022,594.

[3] Medeghini Laura; Botticelli Michela; Cadena-Irizar Ana C.; Lepri Barbara; Ferrandes Antonio F.; Costa Mafalda; Barrulas Pedro Blue shadows of Roman glass artefacts [J]. Microchemical Journal,2022.

[4] Zhou Xueqi; Lv Hongshu; Cui Jianfeng; Dong Xinlin; Wang Ying Fluorite used in ancient Chinese glassmaking during the 10th to 12th centuries: Evidence from glass products excavated in the capital city site of the Liao dynasty [J]. Archaeometry,2022.

[5] Mariangela Vandini; Sara Fiorentino from Crystals to Color: A Compendium of Multi-Analytical Data on Mineralogical Phases in Opaque Colored Glass Mosaic Tesserae [J]. Minerals,2020.

[6] Saminpanya Seriwat, Saiyasombat Chatree, Thammajak Nirawat, Samrong Chanakarn, Footrakul Sirilak, Potisuppaiboon Nichanan, Sirisurawong Ekkasit, Witchanantakul Thumrongsak, Rojviriya Catleya. Shedding New Light on Ancient Glass Beads by Synchrotron, SEM-EDS, and Raman Spectroscopy Techniques [J]. Scientific reports,2019,9(1).

[7] Tomomi Tamura, Katsuhiko Oga. Archaeometrical investigation of natron glass excavated in Japan [J]. Microchemical Journal,2016,126.

[8] Bogdan Constantinescu, Daniela Cristea-Stan, Zoltán Szőkefalvi-Nagy, Imre Kovács, Ildikó Harsányi, Zsolt Kasztovszky. PIXE and PGAA – Complementary methods for studies on ancient glass artefacts (from Byzantine, late medieval to modern Murano glass) [J]. Nuclear Inst. and Methods in Physics Research, B,2018,417.

[9] Jakob König; Alberto Lopez-Gil; Paula Cimavilla-Roman; Miguel A. Rodriguez-Perez; Rasmus R. Petersen; Martin B. Østergaard; Niels Iversen; Yuanzheng Yue; Matjaž Spreitzer Synthesis and properties of open- and closed-porous foamed glass with a low density [J] Construction and Building Materials,2020.

[10] Won in K, Dararutana P. X-ray spectroscopy study of ancient glass beads at Hor-Ek, Thailand [J]. Journal of Physics: Conference Series,2021,1719(1).