Data analysis of factors related to weathering of ancient Chinese glass

Xiaoyu Xiong #, Zhaochun Zhuang #*, #, Peiwen Luo #

College of Electronic Engineering (College of Artificial Intelligence), South China Agricultural University, Guangzhou 510642, China

* Corresponding Author Email: ymtksy@stu.scau.edu.cn

#These authors contributed equally.

Abstract. Ancient glass is an important part of the history of cultural relics, and it is important to study the factors related to glass weathering. Ancient Chinese glass is divided into two categories: lead-barium glass and potassium glass, and the excavated glass artifacts are diverse in terms of decoration and color due to the different production processes and the environment in which the glass is located. This paper analyzes data for both types of glass, analyzes the correlation between glass and each index from data and statistics, and uses mathematical methods to establish models to explore the correlation between whether glass is weathered or not and the decoration, color and category to facilitate smoother archaeological research.

Keywords: Glass Weathering, Person’s Chi-square Test, Binary Logistic Regression Analysis.

1. Introduction

Glass has a long history and is one of the first artificial materials to be created in the history of human invention [1], and is of great significance to the study of human historical development and ethnic integration. Weathering is one of the most common phenomena in ancient glass and is a major factor affecting the quantitative analysis of glass composition [2], and the long-term burial of relic glass in the ground can severely damage the surface of buried glass artefacts under certain conditions [3]. It has been concluded with the aid of scanning electron microscopy that hydrolysis plays a dominant role in the weathering of glass, and that conditions such as the composition, geology and hydrological state of the environment of the glass artefacts on the basis of hydrolysis can affect the weathering condition of the glass [4].

The main chemical composition of glass varies from one glass to another. Ancient glass within China mainly includes lead-barium glass, high-potassium glass and soda-lime glass, with lead-barium and high-potassium glass being the purer domestic glass [5]. It has also been documented that some elements are lost from the surface of the glass during long-term weathering. [5]. For example, samples of lead-barium glass gradually lose Pb content during weathering, and the main colour-emitting element Cu is lost outwards, so the colour of the sample changes [6]. Potassium glass loses significant amounts of K during weathering [7].

The chi-square test is a hypothesis test used to determine whether two factors are correlated [8], counting the degree of deviation between the actual observations of a sample and the theoretical inferences. Tang Meiyian [9] et al. analyzed the main influencing factors of the average monthly residential electricity bill by means of chi-square test and logistic regression, and obtained the ranking of the degree of influence of five different factors. Bi Ran [10] et al. examined the plausibility of the statistical properties of the channel based on the cardinality goodness-of-fit test, and investigated the influence of channel performance by key simulation parameters.

The logistic regression model is a simple and efficient classification model, which can be used for binary classification or multi-classification. Guo Wenfeng et al. conducted a MOOC dropout prediction based on a binary logistic regression model [11], and Yin Qiuju et al. used a logistic regression model to evaluate the geological hazard of landslides [12]. In this study species analysis of whether glass will weather is a dichotomous problem, so the logistic regression model was chosen to be a binary logistic regression.
Existing studies have mainly focused on the chemical composition and cultural value of ancient glass, and less on the mechanism of weathering of ancient glass through the establishment of mathematical models. For this reason, this study analyses the correlation between weathering and glass type, colour and decoration based on chi-square tests and binary logistic regression.

2. Methodology

The correlations between weathering and type, color and ornamentation of the theoretical values were analyzed by using the Person's chi-square test. Finally, the categorical variables were treated as dummy variables, and binary logistic regression analysis was performed for those variables with significant correlations.

2.1. Data preprocessing

The data are shown in Table 1 to Table 3 shows. Among them, the categorical variable of color was missing in artifact number $N = \{19, 40, 48, 58\}$, so it was not included in the statistics.

<table>
<thead>
<tr>
<th>Surface weathering/glass type</th>
<th>High potassium (percentage)</th>
<th>Lead Barium (percentage)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weathering</td>
<td>6 (17.6%)</td>
<td>28 (82.4%)</td>
<td>34</td>
</tr>
<tr>
<td>No weathering</td>
<td>12 (50.0%)</td>
<td>12 (50.0%)</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Surface weathering/texturing</th>
<th>A (%)</th>
<th>B (%)</th>
<th>C (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weathering</td>
<td>11 (32.4%)</td>
<td>6 (17.6%)</td>
<td>17 (50.0%)</td>
<td>34</td>
</tr>
<tr>
<td>No weathering</td>
<td>11 (45.8%)</td>
<td>0 (0.0%)</td>
<td>13 (54.2%)</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>6</td>
<td>30</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Surface weathering/color</th>
<th>Black</th>
<th>Purple</th>
<th>Green</th>
<th>Light blue</th>
<th>Light green</th>
<th>dark blue</th>
<th>Dark green</th>
<th>Blue Green</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weathering</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>30</td>
</tr>
<tr>
<td>No weathering</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>20</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>15</td>
<td>54</td>
</tr>
</tbody>
</table>

2.2. Person’s chi-square test

For verifying the correlation between categorical variables, which in this paper refers to the correlation between surface weathering and glass type and decoration respectively, the chi-square distribution is introduced to verify whether the variables are independent (correlation is not significant) because the provided data are discrete and do not conform to normal distribution. The test procedure (using glass type as an example) is as follows.

1) Put forward the hypothesis:

H0: The correlation between glass type and surface weathering is not significant.

H1: The correlation between glass type and surface weathering is significant.

The hypothesis H0 holds in the analysis.

2) significance level $\alpha = 0.05$, $r = 2$, $c = 2$;

3) Calculation of chi-square test statistics $\chi^2$

$E_{i,j} O_{i,j}$, where is the expected value, is the actual value, $r$ is the type of glass, and $c$ is whether the surface is weathered or not.
\[ E_{i,j} = \frac{\left( \sum_{n_i=1}^{c} o_{i,n_i} \right) \cdot \left( \sum_{n_j=1}^{r} o_{n_j,i} \right)}{N} \]  

\( \chi^2 \) \text{Calculate the chi-square according to equation (2) and the degrees of freedom df according to equation (3).}

\[ \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}} \]  

\[ df = (r - 1)(c - 1) \]

4) Find the critical value of the cardinality assignment according to the obtained results for comparison, and infer whether the null hypothesis can be rejected.

The column-linked data of categories and tattoos were imported into SPSS for Person’s chi-square test, and the data are shown in Table 4.

<table>
<thead>
<tr>
<th>Table 4. Person’s chi-square test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation with weathering</td>
</tr>
<tr>
<td>ornamentation</td>
</tr>
<tr>
<td>color</td>
</tr>
<tr>
<td>Category</td>
</tr>
</tbody>
</table>

After concluding that the correlation between category and surface weathering was significant, the Person coefficient \( V = 0.344 \) was further introduced to analyze the strength of the correlation between glass category and surface weathering \( \pi = 0.326 \), which is a general correlation.

2.3. Binary logistic regression analysis

The dependent variable \( W \) of the dichotomous logistic regression model is a dichotomous variable, and the dependent variable of this question is a categorical variable, so it is first taken to be 0 and 1.

\[ W = \begin{cases} 
0 & (\text{Weathering}) \\
1 & (\text{No weathering}) 
\end{cases} \]  

Where 1 represents the surface of glass products without weathering, 0 represents the surface with weathering, assuming that there are \( w \) impression factors that affect the value of \( W \). Combined with the model solution results of 5.1.2, and then use parameter coding to convert the independent variable can be \( w = 1 \), that is, \( X_1 = \text{lead barium glass} \). At this point, the probability of \( w = 1 \) is recorded as the probability of no weathering \( p_1 \), and the probability of \( w = 0 \) is recorded as the probability of weathering \( 1 - p_1 \). To establish the equation relationship between the degree of weathering and glass category, it is necessary to logit transform the probability \( p_1 \), that is

\[ \logit(p_1) = \ln \left( \frac{p_1}{1 - p_1} \right) b_0 + b_1 X_1 \]  

After a simple transformation, we can obtain.

\[ p_1 = \frac{e^{b_0 + b_1 X_1}}{1 + e^{b_0 + b_1 X_1}} \]
Further calculation of the probability of no weathering on the surface of lead-barium glass/none on the surface of high potassium glass.

\[ OR_1 = \exp(\beta_i) \]  

The OR value can be interpreted as an absolute case: OR<1, it can be concluded that the probability of weathering is greater for lead-barium glass compared with high potassium glass; OR>1, the probability of weathering is smaller for lead-barium glass compared with high potassium glass.

3. Results

The calculated variables of the model were derived using SPSS, and the data are shown in Table 5. For the OR diagram shown in Figure 1.

<table>
<thead>
<tr>
<th>β₁</th>
<th>OR</th>
<th>95% confidence interval of OR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X1</td>
<td>-1.54</td>
<td>0.214</td>
</tr>
</tbody>
</table>

From the model, it is concluded that OR= 0.214 < 1, then lead barium glass is more prone to surface weathering and the OR value lies in the interval with high reliability. As shown in Figure 1.

4. Conclusion

For the glass artifacts provided in the form, the mathematical modeling allows the following results.
1) the correlation between surface weathering and glass color and glass decoration is not significant.
2) the correlation between surface weathering and glass type is significant, with a correlation coefficient of 0.326.
3) Of the two glass types, lead-barium glass is more susceptible to surface weathering.

References


