ISSN 2616-5880 Vol. 4, Issue 3: 47-50, DOI: 10.25236/AJMC.2023.040307

# A study on the composition analysis of ancient glass products based on logistic model

**Jinglin Gong[#], Qing Cai[#], Jianfei Hang[*,#]**

*Yangzhou University, Yangzhou, Jiangsu, 225009, China*
*[*]Corresponding author:154147309@qq.com*
*[#]These authors contributed equally to this work.*

***Abstract:*** *As the first items traded on the Silk Road, the study of the chemical composition of glass before and after weathering is of great importance. In order to analyses the classification law of high potassium glass and lead-barium glass, this paper first normalizes the content of each chemical composition, then establishes a logistic regression model to solve the classification law of the two types of glass before and after weathering, and carries out the test of the logistic model, and finds that the accuracy of the model is as high as 83.3%, with good results. Therefore, the classification of this paper is considered reasonable.*

***Keywords:*** *Logistic model, Classification study, Ancient glassware*

## 1. Introduction

Ancient glass making in China was made locally after absorbing techniques from the West Asian and Egyptian regions, so the appearance of glass products is like that of foreign ones, but the chemical composition is different. Stabilizers such as grass wood ash, natural alkali, saltpeter and lead ore, and limestone were added during the firing process [1]. Ancient glass is highly susceptible to weathering by the environment in which it is buried. In the process of weathering, the chemical composition of the glass will change due to the massive exchange of internal elements with environmental elements, which will affect the judgement of its original category and is not conducive to the study of ancient glass artefacts by our archaeologists [2]. This paper therefore hopes to analyses the relationship between the weathering of the surface of glass artefacts and the type, decoration and color of the glass itself, to analyses the presence or absence of weathering chemistry on the surface of artefact samples, to predict the chemical composition content of glass before weathering and to find rules and methods of classifying glass, and then to be able to identify the properties of glass through the chemical composition of glass artefacts [3]. This paper analyses the classification patterns of high potassium glass and lead-barium glass based on the data. Further, for each category, appropriate representative chemical compositions are selected for subclass analysis, and the specific classification methods and results are sought, with reasonable descriptions and sensitivity analyses of the results.

## 2. The fundamental of classification models

This paper requires an analysis of the classification laws for high potassium glass and lead-barium glass based on the annexed data, i.e. the classification laws according to which glass classification results can be obtained for two types of glass - high potassium glass and lead-barium glass. In order to obtain a mathematical model of the relationship between glass classification results and their influencing factors, the main influencing factor, i.e. the content of the 14 chemical components of the artefact, is recorded as the independent variable in this paper, but it can be seen from the table that the size of the different chemical components varies greatly, and the glass classification results will be more influenced by factors with larger content, so it is considered to first normalise the content of all chemical components and use the one-to-one correspondence The data are normalized to the content of each chemical component, and the dependent variable is the result of the glass classification.$x_1, x_2, \dots, x_{14}$ The dependent variable is the result of the glass classification, denoted as Y, i.e. we have Y for Glass classification results (1 - high potassium glass, 0 - barium lead glass);

$x_1, x_2, \dots, x_{14}$Indicates the normalised content of silica, sodium oxide, potassium oxide, calcium oxide, magnesium oxide, aluminium oxide, iron oxide, copper oxide, lead oxide, barium oxide, phosphorus pentoxide, strontium oxide, tin oxide and sulphur dioxide respectively.

As the dependent variable is a discrete two outcome, we can build a logistic regression model.

The basic form of a logistic regression model:

$$P(Y=1|x_1,x_2,\dots x_{14})=\frac{e^{\beta_0+\beta_1x_1+\cdots\beta_{14}x_{14}}}{1+e^{\beta_0+\beta_1x_1+\cdots\beta_{14}x_{14}}}, \tag{1}$$

where$\beta_0$, $\beta_1$, ..., $\beta_{14}$ is similar to the regression coefficient in a multiple linear regression model. The equation represents the probability of the independent variable being 1 when the variable is$x_1,x_2,\dots,x_{14}$ the probability that the independent variable is 1 when the variable is

Remember

$$p=P(Y=1|x_1,x_2,\dots x_{14}). \tag{2}$$

We log-transform equation (1) to give

$$ln\frac{p}{1-p}=\beta_0+\beta_1x_1+\cdots+\beta_{14}x_{14}. \tag{3}$$

When we take the dependent variable$p$ logarithmically transformed, according to$ln\frac{p}{1-p}$ the form of a whole, the logistic regression problem can be transformed into a linear regression problem.

Since the$p$ takes only values of 0 and 1, in practical problems it is not possible to directly regress$p$ regression, we first define a monotonically continuous probability function $\pi$ :

$$\pi=P(Y=1|x_1,x_2,\dots,x_{14}),0<\pi<1. \tag{4}$$

We eventually developed a logistic regression model of the relationship between glass classification results and the influencing factors.

$$ln\frac{\pi}{1-\pi}=\beta_0+\beta_1x_1+\cdots\beta_{14}x_{14},0<\pi<1. \tag{5}$$

At this point, only a reasonable mapping process of the original data is needed to obtain the regression coefficients by linear regression, and finally the regression coefficients can be obtained according to$\pi$ and$p$ the mapping relationship is reflected to obtain the$p$ value [4].

## 3. Results

### *3.1 The establishment of simulation model*

By building a logistic regression model, a generalised linear regression model is obtained.

$$logit(y)\sim 1+x_1+x_2+x_3+x_4+x_5+x_6+x_7,$$

Considering that the weathering or non-weathering of glass can also have a different effect on the chemical content of the glass, it is necessary to discuss the classification of whether the glass is weathered or not [5].

### *3.2 Analysis of experimental results*

*Table 1: Estimates of coefficients when glass is unweathered*

| | Coefficient | Standard error | T-test value | P-value |
|---|---|---|---|---|
| (intercept) | 1027.5 | 5.5932e+08 | 1.837e-06 | 1 |
| $x_1$ | -990.97 | 3.4821e+08 | -1.8291e-06 | 1 |
| $x_2$ | -165.28 | 3.4287e+08 | -4.7467e-07 | 1 |
| $x_3$ | -107.48 | 4.7014e+08 | -2.2862e-07 | 1 |
| $x_4$ | 1567.3 | 8.0054e+08 | 1.9579e-06 | 1 |
| $x_5$ | -1583.8 | 1.546e+09 | -1.0244e-06 | 1 |
| $x_6$ | -523.89 | 3.307e+08 | -1.5853e-06 | 1 |
| $x_7$ | -718.26 | 9.8824e+08 | -7.268e-07 | 1 |
| $x_8$ | 455.48 | 5.6696e+08 | 8.0337e-07 | 1 |
| $x_9$ | -363.17 | 1.3321e+08 | -2.7264e-06 | 1 |
| $x_{10}$ | -62.119 | 2.0453e+08 | -3.0372e-07 | 1 |
| $x_{11}$ | 255.25 | 5.3382e+08 | 4.7815e-07 | 1 |
| $x_{12}$ | 4525.6 | 4.3156e+09 | 1.0487e-06 | 1 |
| $x_{13}$ | 4022.3 | 2.6774e+09 | 1.5023e-06 | 1 |
| $x_{14}$ | -1226.5 | 1.2107e+09 | -1.013e-06 | 1 |

The raw data were screened to obtain 33 groups of unweathered glass. The 14 chemical components

ISSN 2616-5880 Vol. 4, Issue 3: 47-50, DOI: 10.25236/AJMC.2023.040307

of the first 30 groups of glass were selected as independent variables and their classification results as dependent variables, and the estimates of the coefficients in the generalized linear regression model GM when the glass was unweathered were calculated using the fitglm function in Matlab, as shown in Table 1.

At this point, the overall p-value is 0.000486, a very small value and the model is significant.

The analysis then focuses on the quantitative analysis of the factors affecting the glass classification results and the relationships between them. The absolute values of the coefficients before the independent variables are approximated as weights, i.e. the quantified values of the importance of the influencing factors, where the positive or negative sign only indicates a positive or negative correlation, while it is the absolute value of the coefficients that reflects the level of relative importance.

Based on the above quantitative analysis, we ranked the absolute values of the coefficients before the independent variables from largest to smallest as shown in Table 2, and obtained the ranking of the corresponding independent variables which is the ranking of the factors affecting the glass classification results when the glass is unweathered from highest to lowest.

*Table 2: Table of pre-dependent coefficients in order from largest to smallest absolute value when glass is unweathered*

| Independent variable | Meaning | Absolute value of coefficients |
|---|---|---|
| $x_{12}$ | Strontium oxide content | 990.97 |
| $x_{13}$ | Tin oxide content | 62.119 |
| $x_5$ | Magnesium oxide content | 107.48 |
| $x_4$ | Calcium oxide content | 165.28 |
| $x_{14}$ | Sulphur dioxide content | 255.25 |
| $x_1$ | Silicon dioxide content | 363.17 |

The 14 chemical components of the first 28 groups of glass were selected as independent variables and their classification results as dependent variables, and the estimates of the coefficients in the generalized linear regression model GM for glass weathering were calculated using the fitglm function in Matlab, as shown in Table 3.

*Table 3: Estimates of coefficients when glass has weathered*

| | **Coefficient** | **Standard error** | **T-test value** | **P-value** |
|---|---|---|---|---|
| **(intercept)** | -56.863 | 1.3571e+08 | -4.19e-07 | 1 |
| $x_1$ | 204.98 | 1.3864e+08 | 1.4785e-06 | 1 |
| $x_2$ | 253.64 | 1.5288e+09 | 1.65916e-07 | 1 |
| $x_3$ | 741.11 | 6.1271e+08 | 1.2096e-06 | 1 |
| $x_4$ | -960.28 | 5.5091e+08 | -1.7431e-06 | 1 |
| $x_5$ | -1746.3 | 1.8666e+09 | -9.3553e-07 | 1 |
| $x_6$ | -1427.1 | 8.4582e+08 | -1.6872e-06 | 1 |
| $x_7$ | 1959.7 | 1.0891e+09 | 1.7993e-06 | 1 |
| $x_8$ | 1279.9 | 5.41e+08 | 2.3658e-06 | 1 |
| $x_9$ | -43.074 | 8.5243e+07 | -5.0531e-07 | 1 |
| $x_{10}$ | -586.03 | 2.226+08 | -2.6327e-06 | 1 |
| $x_{11}$ | 662.45 | 2.673e+08 | 2.4783e-06 | 1 |
| $x_{12}$ | -5677.5 | 5.3923e+09 | -1.0529e-06 | 1 |
| $x_{13}$ | 11457 | 6.7992e+09 | 1.685e-06 | 1 |
| $x_{14}$ | 1332.3 | 4.1855e+08 | 3.1833e-06 | 1 |

At this point, the overall $p$ value is 0.000367, a very small value and the model is significant.

For testing the logistic regression model when the glass was weathered, again the last three sets of data were left as the test group and the rest of the data were used as known data to train the logistic model, resulting in predicted and actual values for the last three sets of weathered glass classification results. The logistic model was able to achieve an accuracy of 83.3%, which is better.

## 4. Conclusions

Ancient glass production in China was made locally by absorbing techniques from West Asia and Egypt, so it is similar in appearance to foreign glass products, but different in chemical composition. In this paper, a logistic model is constructed to classify the composition data of ancient glass products, and the results show that the accuracy of the classification model reaches 83.3%, which can effectively identify the different product categories in the product sample data, and is of great significance for industrial production and heritage identification.

ISSN 2616-5880 Vol. 4, Issue 3: 47-50, DOI: 10.25236/AJMC.2023.040307

**References**

*[1] Cheng Qiyun, Sun Caixin, Zhang Xiaoxing, et al. Short-Term load forecasting model and method for power system based on complementation of neural network and fuzzy logic [J]. Transactions of China Electrotechnical Society, 2004, 19(10): 53-58.*
*[2] Fang Fang. Research on power load forecasting based on Improved BP neural network [D]. Harbin Institute of Technology, 2011.*
*[3] Amjady N. Short-term hourly load forecasting using time series modeling with peak load estimation capability [J]. IEEE Transactions on Power Systems, 2001, 16(4): 798-805.*
*[4] Ma Kunlong. Short term distributed load forecasting method based on big data [D]. Changsha: Hunan University, 2014.*
*[5] Wang C, Zhang Z, Chen Y, et al. Comparing different smoothing methods to detect double-cropping rice phenology based on LAI products - a case study in the Hunan province of China[J]. International journal of remote sensing, 2018, 39(19-20):6405-6428.*