ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

Identification of Weathered Glass Artifacts Components Based on NSGA-II Algorithm with Entropy TOPSIS Method

Ying He*

South China Normal University, Guangzhou, 510631, China
**Corresponding author: yinghe0202@163.com*

Abstract: *Glass has a long history and was one of the first man-made materials invented by man. Since glass artifacts are highly susceptible to weathering, the surface composition of artifacts is easily altered. This situation affects the judgement of the type of artefact greatly. In this regard, this paper investigates the data of lead-barium glass and high potassium glass before and after weathering. Firstly, in order to demonstrate that the weathering of high potassium and lead-barium glass is related to the chemical composition of its surface, the relationship between the weathering of glass and the proportion of the chemical composition content of the glass surface needed to be investigated. In addition, it was decided to carry out an analysis using an independent sample rank sum test to obtain a pattern between weathered glass and the proportion of chemical composition content, as most of the data did not conform to a normal distribution. Finally, using the obtained information, a mechanism analysis model was designed based on the chemical kinetic theory, and a multi-objective optimization model was established using the NSGA-II Algorithm with Entropy Topsis Method. The resulting predictive equations for the chemical composition of lead-barium glass and high potassium glass before and after weathering were acquired. The results show that the discrimination between two types of glass based on the percentage of lead oxide content and the percentage of barium oxide content can procure reliable results in general.*

Keywords: *Weathered Glass Artifacts, Entropy Topsis Method, NSGA-II Algorithm*

1. Introduction

Since its inception, the Silk Road has greatly accelerated the material and cultural exchanges between China and the West. As one of the early man-made materials, glass was brought from the West to the East via the Silk Road, enriching raw materials for making things all over the world. The common types of glass found in existing artefacts are high potassium glass and lead-barium glass. The susceptibility of excavated glass artefacts to weathering makes it difficult to determine the original composition of glass even after modern technology has tested the chemical composition of glass artefacts. The analysis of the original composition of glass is of great importance to the development of archaeology, as it can help archaeologists to identify the type of glass and thus the origin and dynasty of ancient glass.[1]

Based on the data from CUMCM in 2022, the tasks to be addressed in this paper are as follows. Analysis of artefacts for patterns of chemical composition content regarding glass weathering and identify Weathered Glass artifacts components.

For these tasks, the solutions are as follows. This paper will apply statistical knowledge to the link between weathered glass and its chemical composition content. In addition, Based on NSGA-II Algorithm with Entropy Topsis Method, a model that can derive pre-weathering data from the chemical composition content of glass artefacts at the point of weathering, needs to be constructed to help determine the original composition of the glass artefacts.

2. Materials and methods

2.1 Data description and pre-processing

The blanks in the test data indicate that the chemical content of the corresponding part is extremely low, so assign a value of 0 to the blanks, which are characterised by their compositional nature, i.e. the sum of the component ratios should be 100%, but the sum of the component ratios may not be equal to

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

100% due to testing methods etc. Therefore, the summation of the component proportions is carried out in Python3.7 and any data with a summation of less than 85% or more than 105% of the component proportions is considered an outlier and is deleted.

In order to distinguish the degree of weathering in more detail, the degree of weathering of each cultural object is analysed as a whole and in part, one being weathering as a whole and one being weathering in part. In the case of 'unweathered sites', this is recorded as '0'. In the case of 'overall weathering', this is recorded as '1'. In the case of 'severely weathered sites', this is recorded as '2'.

2.2 Statistical patterns of chemical content of weathered artefacts

The proportion of each chemical component content in the two types of weathered artefacts was first tested for normality in order to subsequently select an appropriate treatment. Subsequently, independent sample rank sum tests were carried out using IBM SPSS Statistics 26 on samples with and without weathering of different chemical compositions separately to obtain the significance of the content of different chemical compositions and the rank sum means of weathered and unweathered. At last, the proportional changes in chemical composition content before and after weathering were derived for lead-barium glass and high-potassium glass.

2.3 Mechanistic analysis model based on chemical kinetic theory

x_i and $\tilde{x}_i$ denote the proportion of the composition of the ith chemical component in the glass product before and after weathering respectively. The chemistry of the weathering process is influenced by a number of factors, including environmental factors such as temperature, humidity and time in the environment in which the glass is buried, and internal factors such as the chemical composition of the glass itself. Assuming that the environmental factors are the same and that ionic interactions are not considered, the ith chemical composition m_{x_i} of glass before weathering can only be derived from the chemical composition $\tilde{m}_{x_i}$ after weathering.

$$m_i = f(\tilde{m}_i). \tag{1}$$

The following is an attempt to analyse these mechanistic relationships, using chemical kinetic theory as a base.[2]

It is assumed that the chemical reactions involved in glass weathering are all chemical reactions with a simple series, the basic characteristic of which is that the rate of reaction is proportional to the series power of the concentration of the substance. In the absence of significant changes in surface area and volume, the equivalence is proportional to the order of the mass of the chemical components. The rate of change of the ith chemical component at any point in time in a zero-level reaction satisfies the following equation.

$$v = k_i. \tag{2}$$

k_i represents some constant, influenced by the nature of the chemical reaction itself and the environment. When $k_i > 0$, it means that the mass of the ith chemical component increases during weathering. When $k_i < 0$, it means that the mass of the chemical component decreases with weathering. In reactions where the mass of the chemical component increases, the chemical component of the glass is considered to be the product. The reactants originate from the natural environment and their concentration can be regarded as constant. Therefore consider all reactions in which the chemical composition of the glass increases in mass after weathering as zero-level reactions with a constant reaction rate, and only investigate the possibility that reactions in which the chemical composition of the glass decreases in mass have a multi-level reaction.

Let the weathering process time be t. In a zero-level reaction, the masses of the chemical components before and after weathering satisfy the following relationship.

$$m_{x_i} = \tilde{m}_{x_i} - k_i t. \tag{3}$$

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

For non-zero order reactions, the following relationship equation is available.

$$\frac{m_{x_i}}{\tilde{m}_{x_i}} = e^{k_i t}, \tag{4}$$

$$\frac{1}{\tilde{m}_{x_i}^{n_i-1}} - \frac{1}{m_{x_i}^{n_i-1}} = (n_i - 1)k_i t. \tag{5}$$

Eq. (4) shows the relationship between the mass of the chemical composition before and after weathering when the number of reaction steps is $n_i = 1$. Eq. (5) then represents the situation at reaction level $n_i > 1$. Based on mathematical knowledge, Eqs. (4) and (5) can be deduced from each to the following equations.

$$m_{x_i} = \tilde{m}_{x_i} e^{k_i t}, \tag{6}$$

$$m_{x_i} = \frac{\tilde{m}_{x_i}}{\sqrt[n_i-1]{1 - \tilde{m}_{x_i}^{n-1}(n-1)k_i t}}. \tag{7}$$

In addition, it is assumed that M and $\tilde{M}$ are the masses of the sampling points before and after weathering, respectively. For all artefact sampling points of glasswork, we take one unit of $\tilde{M}$ for the study. This leads to:

$$\tilde{m}_{x_i} = \tilde{M}\tilde{x}_i = \tilde{x}_i. \tag{8}$$

In the zero order reaction, assuming that $\hat{k}_i = k_i t$, equation (9) can be deduced from equation (3).

$$m_{x_i} = \tilde{x}_i - \hat{k}_i. \tag{9}$$

In the first order reaction, assuming that $\hat{k}_i = e^{k_i t}$, equation (10) can be deduced from equation (4).

$$m_{x_i} = \hat{k}_i \tilde{x}_i. \tag{10}$$

In a multistage reaction of $n_i > 1$, assuming that $\hat{k}_i = (n_i - 1)k_i t$, equation (11) can be deduced from equation (7).

$$m_{x_i} = \frac{\tilde{x}_i}{\sqrt[n_i-1]{1 - \hat{k}_i \tilde{x}_i^{n_i-1}}}. \tag{11}$$

Let $\hat{k}_i$ at any level be the reaction parameter. Since the weathering process time is assumed to be the same, $\hat{k}_i$ can be considered a constant if the reaction level is determined. If the reaction level and reaction parameter are known, the ith chemical composition mass before weathering per unit mass of glass product, m_i, can be predicted from the ith chemical composition proportion after weathering, $\tilde{x}_i$. Finally, we organise the model developed in this section into the following equation.

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

$$m_i = f(\tilde{x}_i) = \begin{cases} \tilde{x}_i - \hat{k}_i & n_i = 0 \\ \hat{k}_i \tilde{x}_i & n_i = 1 \\ \dfrac{\tilde{x}_i}{\sqrt[n-1]{1 - \hat{k}_i \tilde{x}_i^{n-1}}} & n_i > 1 \end{cases} . \tag{12}$$

$$x_i = \frac{m_i}{M} = \frac{m_i}{\sum_{i=1}^{14} m_i}$$

For the sake of keeping the expressions in the latter as simple as possible, the model for $\tilde{x}_i$ predicting x_i is written as:

$$\hat{x}_i = g\left(\tilde{x}_i, \hat{k}_i, n_i\right). \tag{13}$$

The above formula is not rigorous as it involves a proportional calculation and a single input of $\tilde{x}_i$ does not provide an accurate prediction of x_i.

2.4 Model parameter estimation based on NSGA-II with multi-attribute decision making

2.4.1 Multi-objective optimization modeling

The process of parameter estimation can be identified as a multi-objective optimisation problem. Unlike single-objective optimisation, in multi-objective optimisation the objective functions conflict with each other, resulting in a situation where there may not necessarily be an optimal solution on all objective functions. For this reason, the results obtained for multi-objective optimisation problems are usually Pareto optimal solutions, and the more conventional solutions include the NSGA algorithm and the NSGA-II algorithm which introduces an elite strategy.[3]

The parameter estimation of the regression model, represented by the gradient descent and least square methods, can be considered as a single objective optimisation problem. On the basis of the notational description, it is assumed that $x_{i,j}$ is the true value of the i th chemical component proportion of the j th item of data and $\hat{x}_{i,j}$ is the predicted value for it. The essence of parameter estimation is to minimise the sum of squares of the errors between the true and predicted values, i.e.

$$\min \sum_{j=1}^{m} (\hat{x}_{i,j} - x_{i,j})^2 . \tag{14}$$

However, this problem cannot be solved by conventional methods for the parameters, the reason being that $\hat{x}_{i,j}$ needs to be predicted from the input $\tilde{x}_{i,j}$. In the dataset given by the question there are no pairs of $\tilde{x}_{i,j}$ and $x_{i,j}$, i.e. there is almost no data given at the same time for unweathered and weathered sampling points of the same glass product.[4]

In this regard, we proceed from the overall distribution of the data. Expectation and variance are the optimal statistics for describing the distribution of the data, and so the investigation is focused on these two indicators. For the i th chemical proportion, the mean and variance are calculated as $E(x_i)$ and $var(x_i)$, respectively, based on the data from the unweathered sampling points given in the dataset. Assuming that the model parameters are determined, the model developed in the previous section is used to predict the pre-weathering chemical composition from the weathered sampling points given in the dataset, and the mean and variance of the predicted results are given as $E(\hat{x}_i)$ and $var(\hat{x}_i)$. If the model is accurate in predicting the pre-weathering chemical composition, the predicted results and the

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

data from the actual unweathered sampling points can be considered as two subsamples drawn from the same total. Their sample means and sample variances would theoretically be very close to each other.

Combining equation (13), the process of parameter estimation can be transformed into the following multi-objective optimisation problem.

$$\begin{aligned} & v-\min \quad h\left(\hat{k},n\right)=\left[h_1\left(\hat{k},n\right),h_2\left(\hat{k},n\right)\right] \\ & h_1\left(\hat{k},n\right)=\sum_{i=1}^{14}\frac{\left|E\left(g\left(\tilde{x}_i,\hat{k}_i,n_i\right)\right)-E(x_i)\right|}{E(x_i)} \\ & h_2\left(\hat{k},n\right)=\sum_{i=1}^{14}\frac{\left|var\left(g\left(\tilde{x}_i,\hat{k}_i,n_i\right)\right)-var(x_i)\right|}{var(x_i)} \\ & s.t. \quad n_i\in N \quad i=1,2,\cdots,14 \end{aligned} \tag{15}$$

This multi-optimisation problem estimates the parameters of $\hat{x}_i=g\left(\tilde{x}_i,\hat{k}_i,n_i\right)$ with the objective of minimising the differences between $E(\hat{x}_i)$ and $E(x_i)$, between $var(\hat{x}_i)$ and $var(x_i)$ for all chemical components. Constraints In addition to the fact that the number of reaction levels listed above are all natural numbers, the actual modelling process requires attention to the following details.

(1) The need to model using data of type lead-barium glass and high potassium glass respectively.

(2) Determine whether a chemical component undergoes a chemical reaction based on the results of a hypothesis test. If no chemical reaction occurs, $\hat{k}_i=0$ and $n_i=0$. If a reaction occurs in which the chemical composition increases in mass after weathering, $n_i=0$ and is considered a zero-level reaction only.

(3) When $n_i=1$, then $\hat{k}_i>1$, combined with equation (10), indicates that the mass of the chemical composition before weathering is greater than the mass of the chemical composition after weathering; when $n_i>1$, then $\hat{k}_i\in\left(0,\frac{1}{x_i^{n-1}}\right)$, otherwise equation (11) gives an imaginary or negative value, and also qualifies the relationship between the mass of the chemical composition before and after weathering.

2.4.2 Model solving based on NSGA-II Algorithm with multi-attribute decision making

For this multi-objective optimisation model, we use NSGA-II algorithm to implement the solution. As an improvement to NSGA, NSGA-II reduces the computational complexity of the algorithm by introducing a fast non-dominated sorting algorithm, Elite Strategy, and the use of congestion and congestion comparison operators, allowing individuals in the Pareto front to scale evenly across the Pareto domain while ensuring population diversity. NSGA-II algorithm has high performance in terms of both running time and optimisation results.

The implementation process of the NSGA-II algorithm is briefly summarised as follows.

(1) Firstly, an initial population of a certain plan is randomly generated, and after non-dominated sorting, the first generation of child populations is obtained by the three basic operations of the genetic algorithm: selection, crossover and mutation.

(2) Starting from the second generation, the parent population is merged with the child population and a fast non-dominated sort is performed, while the congestion degree is calculated for the individuals in each non-dominated layer. The appropriate individuals are selected to form a new parent population according to the non-dominated relationship and the congestion degree of the individuals.

(3) Finally, a new population of children is generated by the basic operations of the genetic algorithm; and so on until the conditions for the end of the program are met.

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

NSGA-II algorithm solves for a set of multiple Pareto optimal solutions. Our goal is to provide accurate estimates of the parameters of the mechanistic analysis model, so we need to select representative data from them as the final result. In this paper, the problem of finding optimal solutions from the Pareto frontier is converted into an evaluation class problem, with the Pareto front as the set of data to be evaluated.

TOPSIS Method is one of the common methods for multi-attribute decision making, and is also known as the ideal solution method. The relative closeness of each data item to be evaluated to the ideal solution is then calculated, i.e. how close to the positive ideal solution and how far away from the negative ideal solution it is, so that the data to be evaluated can be ranked. Finally the overall optimal solution is selected.[5]

TOPSIS Method usually requires the construction of a weighted canonical array $C=(c_{ij})_{m\times 2}$ before calculating the ideal solution and relative closeness. Let the weight vector for each attribute be $w=[w_1,w_2]$ and $B=(b_{ij})_{m\times 2}$ be the normalized decision matrix.

$$c_{ij}=w\cdot b_{ij}\quad i=1,2,\cdots,m\quad j=1,2\,. \tag{16}$$

The common TOPSIS Method usually consists of manually determining the weight vectors or assuming identical weights for each indicator. Entropy TOPSIS method is based on the entropy method of identifying weights. Entropy method is an objective weighting method that calculates the weights as shown in the following equation.

$$f_{ij}=\frac{b_{ij}}{\sum_{i=1}^{m}b_{ij}}\quad i=1,\cdots,m\quad j=1,2\,. \tag{17}$$

$$H_j=-\frac{1}{\ln m}\sum_{i=1}^{m}f_{ij}\cdot\ln f_{ij}\quad j=1,2\,. \tag{18}$$

$$\omega_j=\frac{1-H_j}{\sum_{j=1}^{2}(1-H_j)}\,. \tag{19}$$

By substituting the weights calculated by the entropy weighting method into equation (16), the construction of the weighted canonical array can be completed in the absence of manual decision making. The optimal solution is then derived through the process of the traditional TOPSIS Method.

3. Results

3.1 Characteristics of the ratio of the chemical content of different glasses

Based on the normality test, it was found that most of the data did not satisfy the normal distribution. Therefore, an independent sample rank sum test was used to analyse the statistical pattern of the proportion of chemical content in weathered lead-barium glass versus high potassium glass.

The obvious conclusions that can be drawn are as follows. For lead-barium glass, the proportion of silica, sodium oxide and aluminium oxide decreases after weathering; the proportion of calcium oxide, lead oxide and phosphorus pentoxide increases after weathering. For high potassium glass, the proportion of silica increases after weathering; the proportion of potassium oxide, magnesium oxide, aluminium oxide, iron oxide, lead oxide and phosphorus pentoxide decreases after weathering[6-7].

3.2 Parameter estimation of the model

The data from Pareto front were evaluated by the entropy-weighted TOPSIS method. In order to reduce the randomness of the results, the number of iterations of the NSGA-II algorithm was set to 1000, the population size was set to 500, and the experiment was cycled through 100 times. The optimal solutions

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

obtained for the lead-barium glass data and the high-potassium glass data were as follows.

$$\begin{cases} h_1\left(\hat{k},n\right)=4.10 \\ h_2\left(\hat{k},n\right)=3.09 \end{cases} \quad y=y_0\ . \tag{20}$$

$$\begin{cases} h_1\left(\hat{k},n\right)=9.15 \\ h_2\left(\hat{k},n\right)=10.22 \end{cases} \quad y=y_0\ . \tag{21}$$

In addition, the model's predictions of the pre-weathering composition of lead-barium glass are likely to be more accurate than those for high potassium glass.

3.3 The establishment of simulation model

Based on the results of the program run, the equations for the chemical composition of lead-barium glass and high-potassium glass before and after weathering were obtained separately.

$$\begin{cases} m_1=3.3461\tilde{m}_1 \\ m_2=7.6287\tilde{m}_2 \\ m_4=\tilde{m}_4-0.4131 \\ m_6=2.6144\tilde{m}_6 \\ m_9=\tilde{m}_9-11.1206 \\ m_{11}=\tilde{m}_{11}-3.9240 \end{cases} \quad y=y_0\ . \tag{22}$$

$$\begin{cases} m_1=\tilde{m}_1-5.0018 \\ m_3=9.0655\tilde{m}_3 \\ m_5=2.2368\tilde{m}_5 \\ m_6=2.6888\tilde{m}_6 \\ m_7=16.3645\tilde{m}_7 \\ m_{11}=6.4797\tilde{m}_{11} \end{cases} \quad y=y_1\ . \tag{23}$$

3.4 Analysis of experimental results

By looking at the above two sets of equations the following information can be learned.

(1) Reactions in which the chemical content of the glass increases with weathering mass are all zero-level reactions, while reactions in which the chemical composition decreases with weathering mass are all primary reactions. Excluding cases where the concentration of reactants is constant (glass chemistry as a product and reactants provided by the natural environment), it is very rare in nature for chemical reactions to be independent of the concentration of substances, usually surface-catalyzed reactions. The phenomenon is therefore consistent with the general rule.

(2) Common sense suggests that the proportion of lead oxide, barium oxide and potassium oxide is a key parameter in classifying lead-barium glass from high-potassium glass. According to the above formula, the percentage of barium oxide content does not change significantly during the weathering process of either lead-barium glass or high-potassium glass. During the weathering of lead-barium glass, the percentage of lead oxide increases significantly. During the weathering of high potassium glasses, the proportion of potassium oxide decreases significantly. The results show that the differentiation of high potassium glass from lead-barium glass based on the percentage of lead oxide and the percentage of barium oxide is not affected by weathering and generally gives a more reliable result than that based on the percentage of potassium oxide.

ISSN 2616-5880 Vol. 4, Issue 3: 51-58, DOI: 10.25236/AJMC.2023.040308

The model predictions for the weathered parts of the heritage site 50 were compared with the data from the unweathered parts and the results are shown in Table 1.

Table 1: Unweathered parts of sample point 50 compared to predicted values

Comparison	SiO_2	Na_2O	CaO	Al_2O_3	PbO	P_2O_5
Weathered parts	17.98	0	3.19	1.87	44	6.34
Unweathered parts	45.02	0	3.12	4.16	30.61	6.34
Predicted value	45.24	0	2.09	3.68	24.72	1.82

In this example, comparing the chemical composition of the weathered part before weathering with the current chemical composition of the unweathered part, the model has good predictions for the composition ratios of SiO_2, Na_2O, Al_2O_3, PbO However, the predictions for CaO and P_2O_5 are relatively poor. However, in practice, even if a certain two parts are both weathered or unweathered, their chemical compositions may differ significantly. Therefore the results of the above analysis are only of some reference value.

4. Conclusions

Since the amount of data does not allow the construction of regression models by traditional machine learning methods, this paper firstly analyses the weathering process mechanistically from the perspective of chemical kinetics and constructs a plain prediction model. Then the method of parameter estimation is innovated to convert the parameter estimation problem into a multi-objective optimisation problem by simply estimating the overall expectation and variance. At the same time, the NSGA-II algorithm and the entropy-weighted TOPSIS method are chosen to solve the optimal results exactly. The method can be applied to most of the cases where there is not enough information in the data but still need to construct a prediction model.

The robustness of the model developed in this paper is better, so that the model can be applied to any primitive reaction. The model is therefore capable of reducing the chemical composition of other post-weathering ancient artefacts, and can also model most problems involving chemical reactions.

References

[1] Qun Wu, Congrong Xu, Bo Xiang. Composition analysis and identification of ancient glass products [J]. Journal of Materials, Processing and Design, 2022, 6(2).
[2] Hammes G.G. Principles of chemical kinetics [J]. Methods of Biochemical Analysis, 2007,50.
[3] Reşat ACAR, Youssef Alioui. An Evaluation of A Constrained Multi-Objective Genetic Algorithm [J]. Journal of Scientific Perspectives, 2020, 4(2).
[4] Mesquita-Cunha Mariana, Figueira José Rui, Barbosa-Póvoa Ana Paula. New [formula omitted] constraint methods for multi-objective integer linear programming: A Pareto front representation approach [J]. European Journal of Operational Research, 2023,306(1).
[5] Ma Kunlong. Short term distributed load forecasting method based on big data [D]. Changsha: Hunan University, 2018.
[6] Li Yan, Liu Yifan, Li Shasha, Qi Leijie, Xie Jun, Xie Qing. A Novel Multi-Objective Optimal Design Method for Dry Iron Core Reactor by Incorporating NSGA-II, TOPSIS and Entropy Weight Method [J]. Energies, 2022, 15(19).
[7] Fang Fang. Research on power load forecasting based on Improved BP neural network [D]. Harbin Institute of Technology, 2021.