

玻璃文物成分分析与鉴别的数学模型

张梦婷^a, 游慧玲^b, 张和涛^c, 全 然^a

(河南工业大学 a. 理学院; b. 化学化工学院; c. 管理学院, 郑州 450001)

摘要: 针对 2022 年全国大学生数学建模竞赛 C 题“玻璃文物的成分分析与鉴别问题”中的 4 个问题分别进行建模分析, 根据玻璃文物的化学成分含量、纹饰及颜色等数据, 基于单因素方差分析, 探究玻璃文物的表面风化与玻璃类型、纹饰及颜色的关系, 建立 ARIMA 模型, 预测风化前玻璃文物的化学成分含量; 通过主成分分析提取玻璃文物的分类特征, 结合聚类算法进行亚类划分; 通过基于欧氏距离的聚类算法鉴别未知玻璃文物的类别, 分析玻璃文物分类结果的敏感性; 通过 K-Means + 聚类算法分析不同类别玻璃文物的关联关系, 对不同类别玻璃文物关联关系的差异性进行判断, 给出结论。

关键词: 玻璃文物; 单因素方差分析; 聚类分析; 主成分分析; ARIMA 模型; K-Means + 聚类

中图分类号: O29 文献标志码: A 文章编号: 1674-8646(2024)04-0044-07

Mathematical Model of Composition Analysis and Identification of Glass Relics

Zhang Mengting^a, You Huiling^b, Zhang Hetao^c, Quan Ran^a

(a. School of Science; b. School of Chemistry and Chemical Engineering; c. School of Management, Henan University of Technology, Zhengzhou 450001, China)

Abstract: According to the four problems in the question C “Composition Analysis and Identification of Glass Relics” in 2022 China Undergraduate Mathematical Contest in Modeling, the corresponding mathematical models are built and analyzed respectively. Based on the data of chemical composition content, decoration and color of the glass relics, and based on the one-way ANOVA, the study establishes ARIMA model to explore the relationship between the surface weathering of glass relics and the type, decoration and color of glass, and to predict the chemical content of glass relics before weathering. The classification characteristics of the glass relics are extracted by principal component analysis, and the subcategory division is given by cluster algorithm. The unknown categories of the glass relics are identified by Euclidean distance cluster algorithm and the sensitivity of the classification results of the glass relics is analyzed. The K-Means + cluster algorithm is used to analyze the relationship between different categories of the glass relics and to judge the difference of relationship between different categories of the glass relics. Finally, the conclusion is given.

Key words: Glass relics; One-way ANOVA; Cluster analysis; Principal component analysis; ARIMA model; K-Means + cluster

0 引言

在玻璃文物成分分析与鉴别研究中, 大部分文献资料是通过某些特定指标的含量直接鉴别玻璃文物所属类型^[1-4]。但基于数学模型鉴别玻璃文物类型的研究较少, 只有极少数学者进行研究。如崔剑锋等^[5]采

用多元统计分析方法进行探索。本研究提出并建立多种有效的数学模型对玻璃文物的化学成分进行分析及预测, 鉴别未知玻璃文物类型, 这对玻璃文物的鉴别及保护具有重要的现实意义^[6]。

现有一批我国古代玻璃文物, 考古工作者依据这些文物样品的化学成分及其他检测手段已将其分为高钾玻璃和铅钡玻璃两种类型。基于这批文物样品的相关数据构建多种数学模型以解决以下 4 个问题。

问题一: 分析玻璃文物的表面风化与玻璃类型、纹饰及颜色的关系, 探究玻璃文物表面风化前后化学成分含量的统计规律, 预测玻璃文物风化前的化学成分含量。

问题二: 分析高钾玻璃和铅钡玻璃的分类规律, 对

收稿日期: 2023-11-01

基金项目: “常微分方程”课程的教学改革探索

作者简介: 张梦婷(2002-), 女, 本科生;

游慧玲(2001-), 女, 本科生;

张和涛(2001-), 男, 本科生。

通讯作者: 全 然(1977-), 男, 博士, 副教授。研究方向: 最优化理论及其应用。Email: quanran@haut.edu.cn。

玻璃文物进行亚类划分。

问题三:鉴别未知类别玻璃文物所属类型,对分类结果的敏感性进行分析。

问题四:分析不同类别玻璃文物化学成分之间的关联关系,对不同类别玻璃文物关联关系的差异性进行判断。

1 问题分析

对于问题一,将玻璃文物的相关数据进行预处理后,通过方差分析探究玻璃文物的表面风化与玻璃类型、纹饰及颜色的关系。构建化学成分模型,分类讨论玻璃文物风化前后的化学成分含量统计规律。构建差分整合移动平均自回归模型^[7](Autoregressive integrated moving average model, ARIMA),对玻璃文物风化前的化学成分含量进行预测。

对于问题二,在问题一数据处理结果的基础上基于主成分分析探究玻璃文物的分类规律构建线性分类规律模型,利用系统聚类模型对玻璃文物进行亚类划分,给出具体的划分方法及结果。

对于问题三,将问题二中的聚类中心作为特征指标,使用欧式距离衡量待检测成分与特征指标的匹配

度,鉴别未知类别玻璃文物所属类型,分析结果的敏感性及其稳定性。

对于问题四,基于 K-Means + 聚类模型选出初始聚类中心,基于卡林斯基-哈拉巴斯指数(Calinski-Harabazindex, CHI)和轮廓系数确定最优聚类个数,通过 CHI、戴维森堡丁指数(Davies-Bouldin index, DBI)和轮廓系数 3 种评估指标横向对比聚类效果,通过分类标签的形式得到不同类别玻璃文物化学成分关联关系的差异性。

2 模型建立与求解

2.1 玻璃文物的表面风化与玻璃类型、纹饰及颜色的关系

2.1.1 数据分析与处理

针对附件表单 1,删掉颜色数据缺失部分所在的行。针对附件表单 2,将化学成分占比数据缺失的位置填充数字 0,表示未检测到该成分。由于检测结果中各化学成分比例之和介于 85% ~ 105% 之外的数据视为无效数据,故删掉 15 号和 17 号数据。

为讨论方便,对玻璃的纹饰 x_1 、类型 x_2 及颜色 x_3 进行赋值量化,具体如表 1 所示。

表 1 文本数据赋值量化结果

Tab. 1 Results of text data assignment quantifies

文本数据类型	赋值量化结果
x_1	1(纹饰 A), 2(纹饰 B), 3(纹饰 C)
x_2	0(高钾玻璃), 1(铅钡玻璃)
x_3	0.5(浅绿色), 1(绿色), 1.5(深绿色), 2(浅蓝色), 2.5(蓝绿色), 3(深蓝色), 3.5(紫色), 4(黑色)

2.1.2 基于单因素方差分析的关联性分析

方差分析^[8]是通过对实验数据进行分析,检验方

差相等的多个正态总体均值是否相等,进而判断各因素对实验指标的影响是否显著的一种方法。

表 2 单因素方差分析结果

Tab. 2 Results of one-way ANOVA

	差异源	SS	df	MS	F	P-value	F-crit
玻璃纹饰 x_1	组间	1.419	2	0.710	3.037	0.057	3.179
	组内	11.914	51	0.234			
	总计	13.333	53				
玻璃类型 x_2	组间	1.333	1	1.333	5.778	0.020	4.027
	组内	12	52	0.231			
	总计	13.333	53				
玻璃颜色 x_3	组间	0.971	6	0.162	0.615	0.717	2.299
	组内	12.363	47	0.263			
	总计	13.333	53				

分别建立玻璃纹饰 x_1 、类型 x_2 及颜色 x_3 对玻璃文物表面风化的单因素方差分析表,计算玻璃纹饰

x_1 、类型 x_2 及颜色 x_3 这 3 个因素的组内平均值及总均值,计算离差平方和^[9],用其表示检测结果间的差异

性,再进行 F 检验。结果如表 2 所示。其中,SS 为离差平方和,df 为自由度,MS 为均方,F 为组间均方与组内均方之比,P-value 为影响因素对玻璃文物表面风化无显著影响的概率,F-crit 为显著性水平为 0.05 时的 F 临界值。

当 $F > F\text{-crit}$ 时,表示该因素对玻璃文物表面风化有显著影响。由表 2 可见,对于玻璃纹饰 x_1 ,由于 $F < F\text{-crit}$,故玻璃纹饰对玻璃文物的表面风化无显著影响。对于玻璃类型 x_2 ,由于 $F > F\text{-crit}$,故玻璃类型对玻璃文物的表面风化有显著影响。对于玻璃颜色 x_3 ,由于 $F < F\text{-crit}$,故玻璃颜色对玻璃文物的表面风化无显著影响。

2.1.3 风化前后玻璃文物化学成分的总体统计规律及可视化分析

将不同类型的玻璃文物按照风化前和风化后分成高钾风化、高钾未风化、铅钡风化、铅钡未风化 4 类,筛选数据分别统计讨论得出:高钾玻璃文物风化后,化学成分 K_2O 、 CaO 、 MgO 、 Al_2O_3 、 Fe_2O_3 、 CuO 明显减少;铅钡玻璃文物风化后,化学成分 PbO 、 BaO 明显减少。

为使结果更加精确,通过加权平均构造玻璃文物风化前后各个化学成分含量变化的数学模型,并进行详细讨论。记 a_{ij} 为风化前不同类型玻璃文物第 i 次检验到第 j 种化学成分的含量, b_{ij} 为风化后不同类型玻璃文物第 i 次检验到第 j 种化学成分的含量, m 表示化学成分个数, C_j 为第 j 种化学成分风化前后变化的加权平均百分数结果。建立如下风化前后各个化学成分变化的数学模型:

$$C_j = \frac{\frac{1}{n} \sum_{i=1}^n a_{ij} - \frac{1}{n} \sum_{i=1}^n b_{ij}}{\frac{1}{n} \sum_{i=1}^n b_{ij}} \times 100\%,$$

$$(i = 1, 2, \dots, m; j = 1, 2, \dots, m) \quad (1)$$

由式(1)可计算出高钾玻璃和铅钡玻璃在风化前后各个化学成分的变化百分比,具体如图 1 和图 2 所示。可见,在高钾玻璃文物中,化学成分含量变化率较大的是 SO_2 、 CuO 、 BaO (按变化率从高到低的次序排

列,下同);在铅钡玻璃文物中,化学成分含量变化率较大的是 P_2O_5 、 MgO 、 CaO 、 Al_2O_3 、 Na_2O ,这与前述通过筛选统计得到的结果一致。

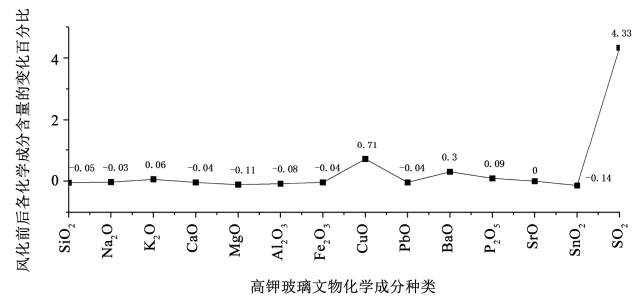


图 1 高钾玻璃文物风化前后各化学成分含量的变化百分比

Fig. 1 Percentage change of chemical constituents before and after weathering of high potassium glass relics

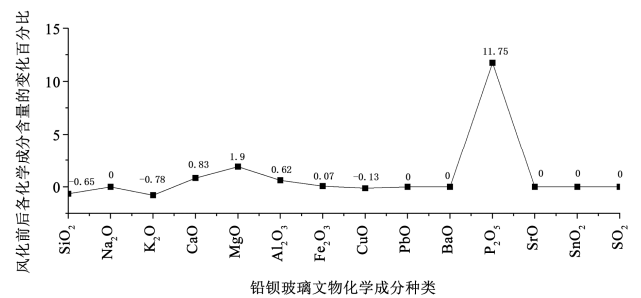


图 2 铅钡玻璃文物风化前后化学成分含量的变化百分比

Fig. 2 Percentage change of chemical composition before and after weathering of lead-barium glass relics

2.1.4 基于 ARIMA 模型预测风化前玻璃文物的化学成分含量

为预测风化前玻璃文物的化学成分含量,结合上述分析及图 1 和图 2 的数据变化趋势,识别化学成分变化为非平稳性,构建 ARIMA 模型进行分析。

通过做一阶差分得到平稳时间序列,通过计算并分析自相关函数 ACF 和偏自相关函数 PACF^[10],分别选取最佳的 ARIMA 模型阶数和阶数,构建模型为:

$$X_t = c + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (2)$$

通过 Matlab 运算^[11],预测出风化后的玻璃文物在风化前的化学成分,部分化学成分预测值如表 3 所示。

表 3 部分化学成分预测值

Tab. 3 Predicted values of some chemical components

文物采样点	SiO ₂	Na ₂ O	K ₂ O	CaO	MgO	Al ₂ O ₃
7	75.58804849	0	0	7.033396248	0	7.514010879
9	73.34773195	0	11.6430	3.855166961	0	4.738607686
10	74.47608177	0	18.1012	1.301892762	0	2.899120442
12	67.45640223	0	18.4723	4.149252592	0	4.857530708
22	59.92126007	0	12.2749	8.676248143	4.920708523	10.56129667
27	73.81449451	0	0	6.02803714	5.094078295	9.292811094

2.2 玻璃文物的亚类划分

2.2.1 基于 PCA 的玻璃文物分类特征提取

由于玻璃文物的化学成分种类多,各化学成分对分类结果的影响不尽相同,为了排除非主要成分对分类结果的影响,采用主成分分析(Principal component analysis, PCA)对数据进行降维处理^[12]。

针对附件表单 2 中 14 种化学成分的检测数据,令主成分信息保留率 $T = 0.70$,特征值大于 1 为提取标准,共得到 5 个主成分指标。取 5 个主成分指标的贡献率^[13]为权重,构建玻璃文物的主成分分类规律模型:

$$F = 0.30028z_1 + 0.17211z_2 + 0.11837z_3 + 0.08031z_4 + 0.07607z_5 \quad (3)$$

表 4 14 种化学成分的主成分系数及 5 个主成分的贡献率

Tab. 4 Principal component coefficients of 14 chemical components and contribution rates of 5 principal components

主成分	SiO ₂	Na ₂ O	K ₂ O	CaO	MgO	Al ₂ O ₃	Fe ₂ O ₃	CuO	PbO	BaO	P ₂ O ₅	SrO	SnO ₂	SO ₂	贡献率
z_1	-0.298	0.373	0.035	0.357	0.190	0.177	0.286	0.241	-0.110	-0.395	-0.331	-0.203	-0.326	0.102	0.30028
z_2	0.002	-0.339	-0.213	0.150	0.422	0.405	0.236	0.365	0.016	0.199	-0.044	0.437	0.229	0.030	0.17211
z_3	0.274	0.108	0.035	-0.290	-0.253	0.278	0.134	-0.098	-0.641	0.181	-0.381	0.063	0.031	0.257	0.11837
z_4	-0.032	0.156	-0.806	-0.080	-0.046	-0.203	-0.355	0.124	-0.053	-0.004	-0.174	0.195	-0.251	-0.012	0.08031
z_5	0.242	0.046	0.198	-0.066	0.223	-0.130	-0.027	0.110	-0.150	0.081	-0.292	0.053	-0.145	-0.823	0.07607

2.2.2 基于聚类算法的玻璃文物亚类划分

由高钾玻璃文物和铅钡玻璃文物相关数据的散点图可以看出,该问题适合采用系统聚类算法^[10]。沿用 2.1 节中的数据及均值公式,分别求出风化前后高钾

其中,14 种化学成分的主成分系数及 5 个主成分的贡献率如表 4 所示。化学成分的主成分系数越大表明该化学成分越能代表主成分的大小;主成分的贡献率越大表明该主成分保留的数据信息越多。由表 4 可以看出,主成分 z_1 的主要化学成分有 BaO、Na₂O、CaO 等,贡献率为 30.0278%,其保留的数据信息最多;主成分 z_2 的主要化学成分有 SrO、MgO、Al₂O₃ 等,贡献率为 17.211%;主成分 z_3 的主要化学成分有 Al₂O₃ 等,贡献率为 11.837%;主成分 z_4 的主要化学成分有 K₂O、Fe₂O₃ 等,贡献率为 8.0308%;主成分 z_5 的主要化学成分有 SiO₂ 等,贡献率为 7.607%。

玻璃文物和铅钡玻璃文物化学成分的均值;通过 Matlab 循环计算轮廓系数及聚类数^[7],绘制风化高钾玻璃文物、未风化高钾玻璃文物、风化铅钡玻璃文物及未风化玻璃文物的树状图,如图 3 所示。

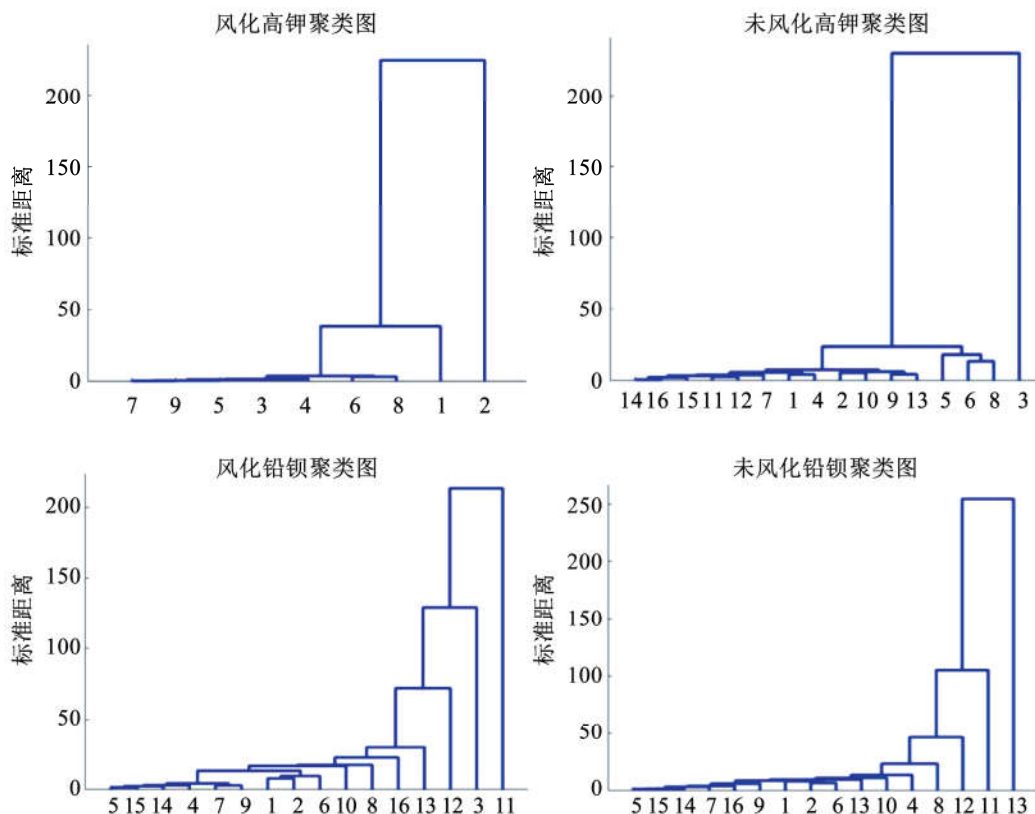


图 3 树状聚类图

Fig. 3 Tree cluster graph

根据图 3 得到聚类分析结果,对玻璃文物进行亚类划分,结果如表 5 所示。

表 5 亚类划分结果汇总

Tab.5 Summary of subclass division results

名称	类别	聚类中心	范围
风化高钾玻璃文物	第一类	2.32	SiO ₂
	第二类	1.66	除第一类其他物质
未风化高钾玻璃文物	第一类	3.41	SiO ₂ 、Al ₂ O ₃ 、CaO、K ₂ O
	第二类	1.93	除第一类其他物质
风化铅钡玻璃文物	第一类	3.18	PbO
	第二类	3.54	SiO ₂
	第三类	1.90	除第一、二类其他物质
未风化铅钡玻璃文物	第一类	3.21	SiO ₂
	第二类	2.67	PbO
	第三类	2.03	除第一、二类其他物质

2.3 基于聚类分析的化学成分分析与鉴别
针对附件表单 3 中未知类别玻璃文物的化学成分,将表 5 中亚类聚类中心的结果作为特征指标,使用

欧氏距离^[9]衡量待检测成分与特征指标的匹配度,从而实现未知玻璃文物的分类鉴定,结果如表 6 所示。

表 6 未知文物的类别鉴定结果

Tab.6 Results of the classification of unknown artifacts

文物编号	A1	A2	A3	A4	A5	A6	A7	A8
表面风化	无风化	风化	无风化	无风化	风化	风化	风化	无风化
匹配度	0.97	0.72	0.85	0.87	0.78	0.76	0.93	0.79
推断分类	高钾	铅钡	铅钡	铅钡	铅钡	高钾	高钾	铅钡
推断亚类	2	2	2	2	1	1	1	1

对分类结果的敏感性进行分析。由于数据量适中,采取描述性分析方法进行检验,结果如表 7 所示。可见,数据波动幅度较大,所建模型敏感性较强。

为测试模型结果的稳定性,对原始未分类玻璃文物的各项化学成分进行 1%、2%、5%、10%、20%、25% 的扰动,结果准确,故稳定性良好。

表 7 敏感性分析

Tab.7 Sensitivity analysis

文物	A1	A2	A3	A4	A5	A6	A7	A8
平均值	1.89046154	5.70615385	2.81846154	2.81846153	5.41942307	0.52846153	0.52846153	3.95890109
最小值	0	0	0.01642857	0.01642857	0	0	0	0
方差	4.8577968	236.77379	24.0732445	24.0732444	170.917949	0.73528076	0.73528076	75.6358976

2.4 基于 K-Means++ 聚类算法的玻璃文物化学成分关联度分析

2.4.1 基于 K-Means++ 的初始聚类中心选取

K-Means 算法中的 k 个初始聚类中心的选择对最终聚类结果及运行时间有极大的影响。传统 K-Means 算法的聚类效果依赖于聚类中心的初始化,而 K-Means++ 聚类算法^[14]以“使初始聚类中心间的相互距离尽可能的远”为基本原则,对初始聚类中心的选择方法进行优化,从而提高聚类精度,加快收敛速度。

给出 K-Means++ 聚类算法选取初始聚类中心的算法步骤:

步骤 1:在数据集 X 中随机令一个样本点作为初

始聚类中心 C_1 。

步骤 2:计算每个样品点与初始聚类中心 C_1 的距离 $D(x_m)$ 。

步骤 3:从数据集 X 中随机选择下一个聚类中心 C_j ,每个样品点被选择的概率为:

$$P(x) = \frac{D(x_m)^2}{\sum_{j=1}^k M(x_j)^2} \quad (4)$$

其中, $M(x_j)$ 为 C_j 与距样本点最近的选择点之间的距离。令 $P(x)$ 最大时对应的样本点为新的初始聚类中心。

步骤 4:重复步骤 2 和步骤 3,直至选出 k 个初始聚类中心。

2.4.2 最优聚类数的确定

采用 CHI 和轮廓系数确定最优聚类数。CHI 的本质为类间距离与类内距离的比值。将容量为 N 的数据集 X 聚成 K 类,用各个类中心点与数据集 X 中心点的距离平方和来度量类间距离 B_k ,用类内各点与类中心点的距离平方和来度量类内距离 W_k 。CHI 的计算公式为:

$$s = \frac{B_k}{W_k} = \frac{\text{tr}[\sum_{q=1}^k n_q (c_q - c_e)(c_q - c_e)^T](N-K)}{\text{tr}[\sum_{q=1}^k \sum_{x \in Q_q} (x - c_q)(x - c_q)^T](K-1)} \quad (5)$$

其中, c_q 为类 q 的中心点, c_e 为数据集 X 的中心点, n_q 为某一类 q 中的数据量, Q_q 为类 q 的数据集。

轮廓系数可用于评价聚类效果,其范围在 $[-1, 1]$ 。同类样本距离越近,不同类样本距离越远轮廓系数越大,说明聚类效果越合理。令某个样本的轮廓系数为 S ,聚类总的轮廓系数为 SC ,具体如下:

$$s = \frac{b-a}{\text{sum}(a,b)}, SC = \frac{\sum_{i=1}^N s_i}{N} \quad (6)$$

其中, a 表示一个样本与其所在簇内其他样本的平均距离, b 代表一个样本与其他簇样本的平均距离。

运用 Matlab 计算并制成簇计算标准值,如图 4 所示。图中分类所代表的柱最高时聚类的个数最优。得出:高钾未风化玻璃文物在 $k=3$ 时最优,高钾风化玻

璃文物在 $k=4$ 时最优,铅钡未风化玻璃文物在 $k=5$ 时最优,铅钡风化玻璃文物在 $k=4$ 时最优。

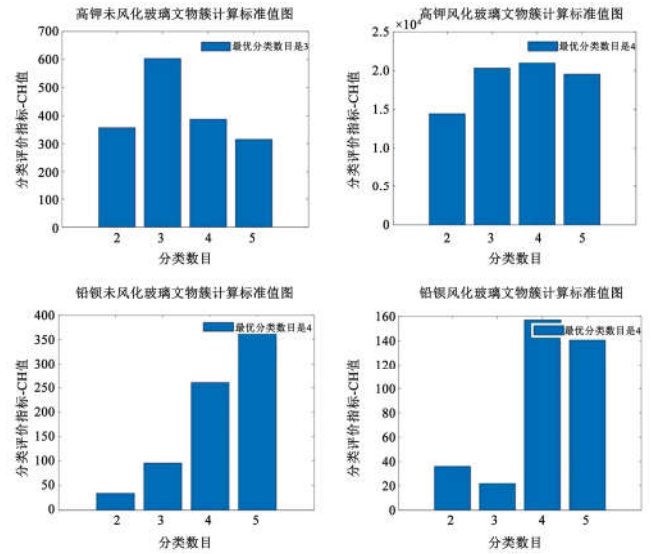


图 4 簇计算标准值

Fig. 4 Cluster computation standard value

2.4.3 基于 K-Means++ 聚类算法的玻璃文物化学成分关联度分析模型

为对数据进行可视化分析,构建了玻璃文物化学成分关联度分析模型。运用 Matlab 制图,结果如图 5 所示。图 5 中每一行、每一列的聚类均较为分散,故聚类效果较好。

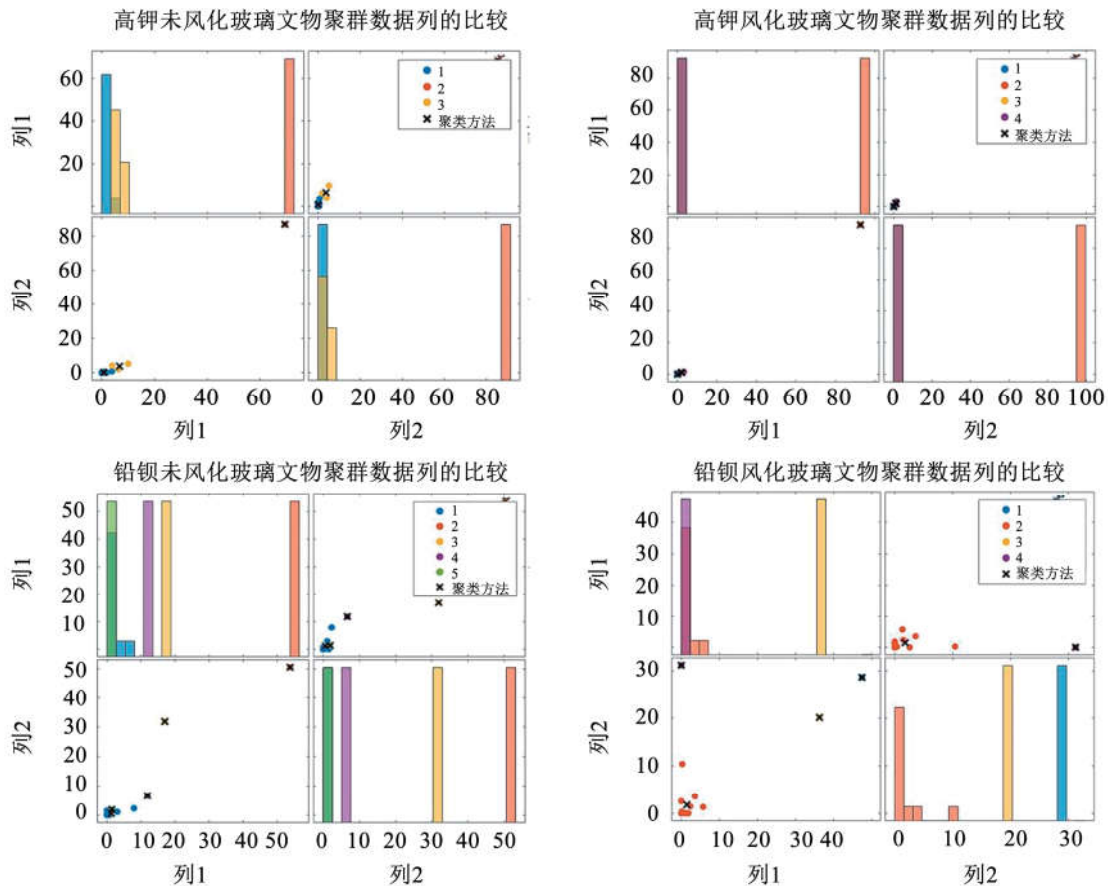


图 5 最优聚类可视化分析

Fig. 5 Visual analysis of optimal clustering

通过表 8 对比可得:高钾玻璃文物风化前后的分类只有 K_2O 、 CaO 、 CuO 发生改变,而铅钡玻璃文物全部化学成分都发生了改变。

表 8 分类标签
Tab. 8 Category label

类别	SiO_2	Na_2O	K_2O	CaO	MgO	Al_2O_3	Fe_2O_3
高钾未风化	2	1	3	3	1	3	1
高钾风化	2	1	1	4	1	3	1
铅钡未风化	2	1	1	1	1	5	1
铅钡风化	3	2	2	2	2	2	2

类别	CuO	PbO	BaO	P_2O_5	SrO	SnO_2	SO_2
高钾未风化	1	1	1	1	1	1	1
高钾风化	4	1	1	1	1	1	1
铅钡未风化	1	3	4	1	1	1	1
铅钡风化	2	1	4	2	2	2	2

采用 CHI、DBI、轮廓系数^[15] 3 个评价指标分析聚类效果,如表 9 所示。发现 4 个分类下的轮廓系数均接近 1,说明簇内距离小,簇间距离大,CHI 同理。各项 DBI 都较小,说明其类内距离较小,类间距离较大,即不同类别之间的相似度较小,说明聚类结果的质量较好,各类别之间差异较明显。

表 9 聚类效果分析
Tab. 9 Clustering effect analysis

玻璃类型	最优分类数目	CHI	DBI	轮廓系数
高钾未风化	3	603.9374	0.41187	0.83888
高钾风化	4	20991.0155	0.46621	0.78836
铅钡未风化	5	387.4223	0.13902	0.88246
铅钡风化	4	156.951	0.11074	0.93844

2.4.4 不同类别之间化学成分关联关系的差异性判断

为了得到更加准确的聚类结果,利用 Matlab 运行算法 4 次后,将不同类别玻璃文物风化前后化学成分的 4 种聚类结果以分类标签的形式汇总,结果如表 10 所示。

通过表 10 可以看出,只有极少次数的聚类会使聚类结果发生明显变化。其中,铅钡未风化玻璃文物聚类数为 5 与高钾风化玻璃文物聚类数为 4 时对应的聚类结果相较其他结果差异性较大,而其他大多数聚类结果都围绕着众数上下波动或保持不变,但一般会高于正常值,说明不同距离造成的影响具有一定的同质性,但弱于正常值。

表 10 聚类结果的分类标签形式汇总

Tab. 10 Summary of category label forms of clustering results

高钾未风化	2	1	3	3	1	3	1	1	1	1	1	1	1	1
高钾风化	2	1	1	4	1	3	1	4	1	1	1	1	1	1
铅钡未风化	2	1	1	1	1	5	1	1	3	4	1	1	1	1
铅钡风化	3	2	2	2	2	2	2	2	1	4	2	2	2	2

3 结论

针对玻璃文物表面风化问题,通过单因素方差分析研究发现:玻璃类型显著影响玻璃文物的表面风化

程度,而玻璃纹饰和颜色影响不显著。建化学成分模型分析得到:风化前后,高钾玻璃文物 SO_2 、 CuO 、 BaO 等化学成分含量变化率较大,铅钡玻璃文物 P_2O_5 、 MgO 、 CaO 、 Al_2O_3 、 Na_2O 等化学成分含量变化率较大。通过 ARIMA 模型预测得到风化前多个玻璃文物的化学成分含量。

针对玻璃文物的分类问题,通过主成分分析确定化学成分的分类规律,通过系统聚类算法对 4 种类型的玻璃文物进行了亚类划分,具有良好的合理性及稳定性。

针对未知类别玻璃文物的化学成分分析与类别鉴定问题,基于亚类划分结果构建化学成分识别模型,确定了 8 种未知类别文物的类别,对匹配度进行描述统计并对化学成分数据进行 5% ~ 30% 的扰动,结果表明,该模型具有较好的稳定性及敏感性。

针对不同类别玻璃文物化学成分的相关性及差异性,通过 K - Means + 聚类算法描述玻璃文物化学成分间的相关性,通过多次聚类揭示 4 类玻璃文物间的差异。

通过以上分析可知,所建数学模型合理,能实现对玻璃文物成分的有效分析,可为其他玻璃文物成分研究提供参考。

参考文献:

- [1] 温睿,赵志强,马健,等.新疆巴里坤石人子沟遗址群出土玻璃珠的成分分析[J].光谱学与光谱分析,2016(09):2901-2965.
- [2] 干福熹,赵虹霞,李青会,等.湖北省出土战国玻璃制品的科技分析与研究[J].江汉考古,2010,115(02):108-116,151.
- [3] 李青会,黄教珍,李飞,等.中国出土的一批战国古玻璃样品化学成分的检测[J].文物保护与考古科学,2006(02):8-13.
- [4] 胡志中,李佩,蒋璐蔓,等.古代玻璃材料 LA-ICP-MS 组分分析及产源研究[J].岩矿测试,2020,39(04):505-514.
- [5] 崔剑锋,吴小红,谭远辉,等.湖南沅水流域战国时期楚墓出土古代玻璃器的成分分析[J].硅酸盐学报,2009,37(11):1909-1913,1918.
- [6] 赵华玲.淄博地区古代玻璃历史发展的研究[D].苏州:苏州大学,2008.
- [7] 邢航.回归分析中建立数学模型的方法及其应用[J].职大学报,2008,16(04):68-70.
- [8] 林明海.对主成分分析法运用中的十个问题的解析[J].统计与决策,2007(16):16-18.
- [9] 姜启源.数学模型(第二版)[M].北京:高等教育出版社,1993:6-71.
- [10] 阎慈琳.关于用主成分分析做综合评价的若干问题[J].数理统计与管理,1998,17(02):22-25.
- [11] 王正林,刘明.精通 MATLAB7 [M].北京:电子工业出版社,2008.
- [12] 楼建华.数学建模与数学实验[J].黑龙江高教研究,2003(02):126-127.
- [13] 毕玉珊.基于聚类分析与决策树算法的社交网络客户细分研究[D].北京,北京工业大学,2019.
- [14] 司守奎.数学建模算法与应用[M].北京:国防工业出版社,2011:595-601.
- [15] Saroj Kavita. Review: study on simple kmean and modified K mean clustering technique[J]. International Journal of Computer Science Engineering and Technology,2016,6(07):279-281.