

文章编号: 1671-251X(2025)03-0113-09

DOI: 10.13272/j.issn.1671-251x.2025020046

# 基于半监督学习的煤层钻孔预抽瓦斯状态评价方法

晏立<sup>1,2</sup>, 文虎<sup>1,3</sup>, 王振平<sup>1,3</sup>, 金永飞<sup>1,2</sup>

(1. 西安科技大学 安全科学与工程学院, 陕西 西安 710054; 2. 西安天河矿业科技有限责任公司, 陕西 西安 710054; 3. 西安科技大学 西部矿井开采及灾害防治教育部重点实验室, 陕西 西安 710054)

**摘要:** 目前单一钻孔抽采状态评价方法通常依赖于瓦斯抽采浓度, 而忽视了煤层瓦斯赋存的多样性。监督学习模型依赖于样本的特征标记, 在样本量较大时, 人工标注的成本较高; 无监督学习模型缺乏样本标记, 无法实现定性评价。针对上述问题, 提出一种基于半监督学习的煤层钻孔预抽瓦斯状态评价方法。构建了包含甲烷浓度、抽采负压、环境温度等 8 项指标的多维度评价体系, 采用层次分析法(AHP)与模糊评价法(FEM)结合的权重赋值方法, 建立抽采效果等级划分标准。在此基础上, 提出基于高斯混合模型(GMM)与 K-Means 算法的半监督学习模型(SSGMM/SSK-Means), 通过融合少量人工标注样本与大量未标注数据, 实现单一钻孔抽采状态的动态分类。SSGMM 聚集度更好, SSK-Means 效率更高, 形成“精度-效率”的互补关系。在陕西黄陵二号煤矿 215 工作面的应用结果表明: SSGMM 和 SSK-Means 的最大聚集度(MVCR)和修正 Rand 指数(ARI)分别达 82.64% 和 85.83%, 显著优于传统聚类方法; 通过动态反馈机制优化后, 原等级为“差”的钻孔抽采效率提升 5.26%~5.80%, 补差率达 100%。

**关键词:** 煤层瓦斯; 抽采效果评价; 半监督学习; 层次分析法; 模糊评价法; 高斯混合模型; K-Means 算法  
中图分类号: TD712 文献标志码: A

Evaluation method for gas pre-extraction status in coal seam boreholes based on semi-supervised learning

YAN Li<sup>1,2</sup>, WEN Hu<sup>1,3</sup>, WANG Zhenping<sup>1,3</sup>, JIN Yongfei<sup>1,2</sup>

(1. College of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; 2. Xi'an Tianhe Mining Technology Co., Ltd., Xi'an 710054, China; 3. Key Laboratory of Western Mine Exploitation and Hazard Prevention, Ministry of Education, Xi'an University of Science and Technology, Xi'an 710054, China)

**Abstract:** Current evaluation methods for single-borehole gas extraction status typically rely on gas concentration, while overlooking the diversity of coal seam gas occurrence. Supervised learning models depend on labeled sample features, but manual labeling becomes costly when the sample size is large. Unsupervised learning models lack sample labeling, making qualitative evaluation infeasible. To address these issues, an evaluation method based on semi-supervised learning was proposed for the gas pre-extraction status evaluation of coal seam boreholes. A multi-dimensional evaluation system was established, incorporating eight indicators such as methane concentration, extraction negative pressure, and ambient temperature. The weighting method combining the analytic hierarchy process (AHP) and fuzzy evaluation method (FEM) was used to establish classification standards for extraction performance. Building on this, a semi-supervised learning model based on the Gaussian mixture model (GMM) and K-Means algorithm (SSGMM/SSK-Means) was developed. By integrating a small number of manually labeled samples and a large quantity of unlabeled data, the model enabled

收稿日期: 2025-02-21; 修回日期: 2025-03-22; 责任编辑: 胡娟。

基金项目: 国家自然科学基金项目(52274227)。

作者简介: 晏立(1997—), 男, 重庆人, 讲师, 博士, 从事应急技术与管理工作, E-mail: yanli@xust.edu.cn。

引用格式: 晏立, 文虎, 王振平, 等. 基于半监督学习的煤层钻孔预抽瓦斯状态评价方法[J]. 工矿自动化, 2025, 51(3): 113-121.

YAN Li, WEN Hu, WANG Zhenping, et al. Evaluation method for gas pre-extraction status in coal seam boreholes based on semi-supervised learning[J]. Journal of Mine Automation, 2025, 51(3): 113-121.



扫码移动阅读

dynamic classification of single-borehole extraction status. The SSGMM demonstrated better clustering rate, while the SSK-Means achieved higher efficiency, developing a complementary "accuracy-efficiency" relationship. The application results from the 215 working face of the Huangling No. 2 Coal Mine in Shaanxi Province showed that the maximum validity clustering rate (MVCR) and adjusted rand index (ARI) of SSGMM and SSK-Means reached 82.64% and 85.83%, respectively, significantly outperforming conventional clustering methods. After optimization through a dynamic feedback mechanism, boreholes initially classified as "poor" showed an improvement of 5.26% to 5.80% in extraction efficiency, achieving a 100% remediation rate.

**Key words:** coal seam gas; extraction performance evaluation; semi-supervised learning; Analytic Hierarchy Process; Fuzzy Evaluation Method; Gaussian Mixture Model; K-Means algorithm

## 0 引言

在煤与瓦斯突出矿井及高瓦斯赋存矿井的开采中,煤层瓦斯事故频发<sup>[1-2]</sup>,提升瓦斯钻孔预抽效率已成为保障煤矿安全生产的核心议题之一<sup>[3-4]</sup>。预先抽采煤层瓦斯是减少开采过程中瓦斯灾害与风险的关键措施<sup>[5-6]</sup>。在此背景下,结合先进监测技术和综合评价体系,强化煤层瓦斯安全管理体系,是提升煤矿安全生产水平的必然要求,也是可持续发展的重要保障<sup>[7-8]</sup>。

在预抽过程中,通常采用人工现场取样检测方法评价抽采效果<sup>[9]</sup>。该方法虽然操作简便直观,但存在测定时间难以精确控制的问题,无法实现对预抽效果的实时反馈,长期而言可能潜藏着导致采掘失衡的重大隐患。针对以上难题,Xia Tongqiang 等<sup>[10]</sup>通过瓦斯-空气组分模型计算各参数对抽采效率的贡献度,综合判断抽采质量。Jin Kan 等<sup>[11]</sup>建立了一种远距离下保护层开采煤矿瓦斯治理效果评价方法。Xue Yi 等<sup>[12]</sup>构建了基于煤体变形与气-空气组分流动耦合模型的煤层预抽瓦斯质量评价方法,深入解析了预抽瓦斯过程中出现的低质量现象。Ni Xiaoming 等<sup>[13]</sup>通过层次分析法(Analytic Hierarchy Process, AHP)确定各指标权重,计算综合评价值并划分产能等级,适用于区域瓦斯产能评价。Guo Jianhong 等<sup>[14]</sup>和 Song Hongqing 等<sup>[15]</sup>利用监督学习算法,通过模型输出的预测值进行分级,但该过程依赖大量人工标注样本,模型泛化能力受样本量限制。

当前研究主要聚焦于整体煤层瓦斯抽采效果的评价,对单一钻孔的抽采状态分析较为欠缺<sup>[16-17]</sup>,评价结果缺乏精细化,难以实现对单一钻孔的精准调控<sup>[18-19]</sup>。另外目前的钻孔抽采状态评价方法过于单一,通常依赖于瓦斯抽采浓度,而忽视了煤层瓦斯赋存的多样性。由于不同区域的瓦斯赋存规律存在显著差异,高瓦斯区域的预抽钻孔浓度自然高于低瓦斯区域<sup>[20]</sup>。因此,需综合考虑钻孔抽采负压、环境

压力、温度等多种参数,同时结合历史数据,以全面反映单一钻孔的抽采状态,从而实现更加科学的评价与调控<sup>[21-22]</sup>。

目前,随机森林<sup>[23]</sup>、支持向量机<sup>[24]</sup>和线性鉴别分析<sup>[25]</sup>等监督学习模型能够较好地对样本进行分类并得出评价结果。然而,这类模型依赖于样本的特征标记,在样本量较大时,人工标注的成本较高<sup>[26-27]</sup>。无监督学习模型如 K-Means 能够对数据样本进行聚类,但由于缺乏样本标记,训练目标不明确,只能得出聚类结果,而无法实现定性评价<sup>[28-29]</sup>。

针对上述问题,本文提出了一种基于半监督学习的煤层钻孔预抽瓦斯状态评价方法。结合 AHP 与模糊评价法(Fuzzy Evaluation Method, FEM)对评价指标赋予权重,对少量样本进行特征标记,利用高斯混合模型(Gaussian Mixture Model, GMM)和 K-Means 算法对混合样本进行聚类,自动提取特征参数间的关联关系,自适应学习评价指标及当前数据与历史数据的依赖关系,从而实现对单一钻孔抽采状态的精细化评价。

## 1 数据采集和处理

### 1.1 数据来源

以陕西黄陵二号煤矿 215 工作面为研究对象,采用煤层工作面瓦斯预抽钻孔精细化布控系统和人工采集结合的方式进行数据采集。该系统由瓦斯抽放管路智能管控系统、测量监测系统及数据传输处理系统组成,能够实现瓦斯抽放参数实时监测和控制,并选择性地提取目标管路数据,系统结构如图 1 所示。

### 1.2 数据预处理

利用独热(One-hot)编码(图 2)将分类变量的每个取值转换为二进制形式,为每个类别创建独立的列,消除分类变量间的序关系。对分类变量进行标记,生成对应的二进制表示。每个分类变量的 One-hot 编码列数等于其可能取值的个数。将生成

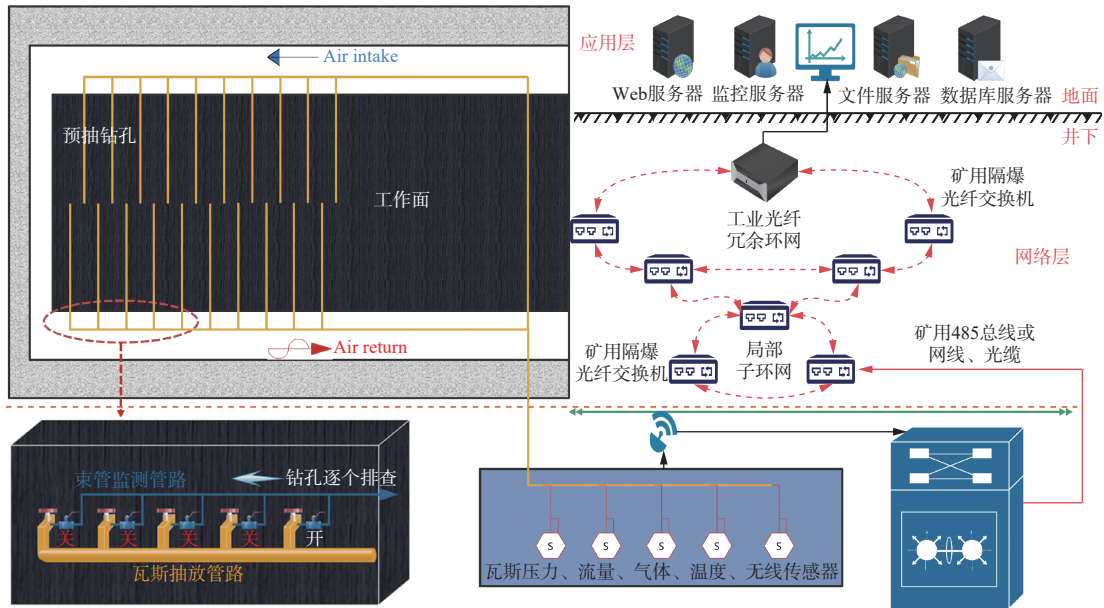


图 1 煤层工作面瓦斯预抽钻孔精细化布控系统结构

Fig. 1 Structure of refined control system for gas pre-extraction boreholes in coal seam working face

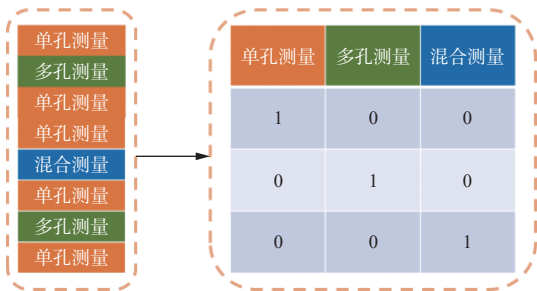


图 2 One-hot 编码

Fig. 2 One-hot encoding

的特征矩阵与原始数据合并, 形成消除类别序关系的标准化输入结构。

采用 MinMaxScaler 归一化方法对数据进行预处理, 即用每个数据值减去最小值, 再除以最大值和最小值的差, 从而将数据映射到 [0,1] 范围内。

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

式中:  $X_{scaled}$  为映射后的值;  $X$  为原始数据;  $X_{max}$  和  $X_{min}$  分别为  $X$  的最大值和最小值。

## 2 半监督学习模型

### 2.1 模型架构

传统 GMM 与 K-Means 模型因缺少标记样本, 存在训练目标不明、聚类结果需人为判断特征空间分布等问题。因此, 本文构建了一种混合半监督学习模型, 融合 AHP 与 FEM, 对少量样本进行特征标记, 再将其与未标记样本混合。利用主成分分析 (Principal Component Analysis, PCA) 优化算法进行特征降维, 提高计算效率。借助 GMM 和 K-Means 聚类器对混合样本聚类, 把已标记样本信息映射到未标记样本, 生成 SSGMM 和 SSK-Means 分类器, 以实现混合样本分类。利用 SSGMM 分类器评价 AHP-FEM 筛选的样本质量与数量, 经自动化学习过程获取最优输入样本及高精度分类器。半监督学习模型架构如图 3 所示。

### 2.2 AHP-FEM 评价体系

#### 2.2.1 评价指标选取

考虑监测数据的类型及其固有的时序性特征,

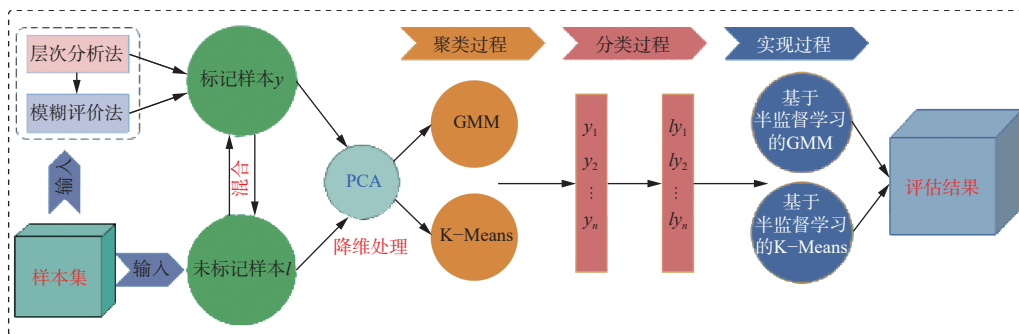


图 3 半监督学习模型架构

Fig. 3 Architecture of semi-supervised learning model

采用现场调研与理论分析相结合的方法,充分利用监测数据特性,确保所选评价指标具有可操作性,并与常规监测实践相一致。钻孔抽采效果评价指标见表 1。

表 1 钻孔抽采效果评价指标

Table 1 Evaluation indicators for borehole and extraction performance

因素名称	因素标号	单位	因素名称	因素标号	单位
甲烷浓度	F1	%	抽采负压	F2	kPa
环境温度	F3	℃	抽采差压	F4	kPa
环境压力	F5	kPa	工况流量	F6	m <sup>3</sup> /min
瓦斯纯流量	F7	m <sup>3</sup> /min	前一天的瓦斯纯流量	F8	m <sup>3</sup> /min

2.2.2 权重计算

AHP 通过构建多级递阶层次结构模型实现决策目标分解。该模型包含目标层、准则层、子准则层及方案层。其核心步骤:通过两两比较法建立判断矩阵;采用特征向量法计算各层元素权重;引入一致性检验,确保判断逻辑合理性。邀请业内专家依据 9 标度法对比较结果进行打分,打分规则见表 2。如果指标 1 相对于指标 2 的重要性打分为  $x$ ,那么指标 2 相对于指标 1 的重要性打分为  $1/x$ 。

表 2 AHP 打分规则

Table 2 AHP scoring rules

$x$	意义	$x$	意义	$x$	意义
1	相等重要	4	介于 3, 5 中间	7	非常重要
2	介于 1, 3 中间	5	较为重要	8	介于 7, 9 中间
3	略微重要	6	介于 5, 7 中间	9	最重要

AHP 充分考虑了决策者对问题的主观感受,并通过一致性检验保证专家打分的合理性。对判断矩阵  $A$  的每一列向量进行归一化处理,得

$$\bar{m}_{ij} = \frac{m_{ij}}{\sum_{i=1}^I m_{ij}} \quad i, j = 1, 2, \dots, I \quad (2)$$

式中:  $m_{ij}$  为判断矩阵中的元素;  $i, j$  分别为判断矩阵的行号和列号;  $I$  为判断矩阵的阶数。

将归一化判断矩阵的各行相加,得

$$w_i = \sum_{j=1}^I \bar{m}_{ij} \quad (3)$$

对  $w_i$  进行归一化处理,得到权重向量:

$$\bar{w}_i = \frac{w_i}{\sum_{i=1}^I w_i} \quad (4)$$

计算判断矩阵的最大特征值和对应的特征向

量。最大特征值计算公式为

$$\lambda_{\max} = \frac{1}{I} \sum_{i=1}^I \frac{(Aw)_i}{w_i} \quad (5)$$

式中:  $w$  为归一化后的权重向量;  $(Aw)_i$  为  $Aw$  向量的第  $i$  个元素。

计算判断矩阵的一致性指标 CI:

$$CI = \frac{\lambda_{\max} - I}{I - 1} \quad (6)$$

判断矩阵的一致性指标 CI 越接近 0, 表示判断矩阵的一致性越好。随机一致性指标 RI 根据随机生成的判断矩阵的平均一致性指标计算,且与判断矩阵的阶数相关。判断矩阵的一致性比例越小,判断矩阵的一致性越好。通常,当  $CR < 0.1$  时,认为判断矩阵的一致性是可接受的,否则需要对判断矩阵进行修正。CR 计算公式为

$$CR = \frac{CI}{RI} \quad (7)$$

利用 AHP 对煤层预抽钻孔的抽采效果评判指标进行权重计算,结果见表 3 和表 4。可看出判断矩阵的最大特征值  $\lambda_{\max} = 8.433$ , 一致性指标  $CI = 0.061$ , 查表可得  $RI = 1.41$ , 最终计算可得  $CR = 0.043 < 0.1$ , 说明该判断矩阵的一致性较好。

表 3 评价指标的判断矩阵、权重及一致性检验

Table 3 Judgment matrix, weights, and consistency test for evaluation indicators

因素	F1	F2	F3	F4	F5	F6	F7	F8	$w_i$	$\lambda_{\max}$	CR
F1	1	3	8	4	5	2	1	1/3	0.191	8.433	0.043
F2	1/3	1	5	2	3	1/3	1/3	1/5	0.057		
F3	1/8	1/5	1	1/4	1/3	1/6	1/8	1/9	0.008		
F4	1/4	1/2	4	1	2	1/3	1/5	1/7	0.043		
F5	1/5	1/3	3	1/2	1	1/4	1/5	1/6	0.028		
F6	1/2	3	6	3	4	1	1/3	1/5	0.131		
F7	1	3	8	5	5	3	1	1/2	0.192		
F8	3	5	9	7	6	5	2	1	0.251		

表 4 随机一致性指标 RI 的取值

Table 4 Value of random consistency index (RI)

序号	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

2.2.3 评价标签体系建立

为提升 AHP 评价指标权重确定的灵活性与可靠性,引入半梯形分布函数,降低专家判断的绝对主观性。鉴于 AHP 评价指标相对关系常呈递增或递减趋势,采用半梯形分布函数模拟该趋势,助力构建更具解释力的评价体系。

降半梯形模型的隶属度为

$$L(X)_i = \begin{cases} 1 & X_i \leq X_{i\min} \\ -\frac{X_i - X_{i\max}}{X_{i\max} - X_{i\min}} & X_{i\min} < X_i < X_{i\max} \\ 0 & X_i \geq X_{i\max} \end{cases} \quad (8)$$

升半梯形模型的隶属度为

$$L(X)_i = \begin{cases} 1 & X_i \leq X_{i\min} \\ \frac{X_i - X_{i\max}}{X_{i\max} - X_{i\min}} & X_{i\min} < X_i < X_{i\max} \\ 0 & X_i \geq X_{i\max} \end{cases} \quad (9)$$

对于降半梯形模型, 隶属度越大, 对满意度的贡献率越小; 对于升半梯形模型, 情况则相反。

煤层瓦斯预抽钻孔抽采效果评价满意值为

$$P = \sum_{i=1}^8 w_i L(X_i) \quad (10)$$

将满意值的评语集分为优、良、中、差 4 类。

优:  $P \in (0.75, 1]$ , 煤层预抽钻孔抽采效果优, 预抽效率高, 钻孔间距合理, 封孔状态良好, 无漏气情况, 抽采管路状态稳定, 负压设置及管道布置合理, 抽采浓度高, 无需调整。

良:  $P \in (0.5, 0.75]$ , 煤层预抽钻孔抽采效果良好, 预抽效率较高, 各项指标符合国家要求, 封孔状态良好, 负压设置合理, 抽采浓度达标。建议检查钻孔是否存在漏气, 根据现场情况适当调节抽采负压。

中:  $P \in (0.25, 0.5]$ , 煤层预抽钻孔抽采效果一般, 预抽效率偏低, 存在漏气情况, 负压设置需要调整, 抽采浓度偏低, 钻孔布置间距不合理, 需要优化相关参数。

差:  $P \in [0, 0.25]$ , 煤层预抽钻孔抽采效果差, 预抽效率极低, 抽采浓度严重不足, 亟需对钻孔及抽采参数进行全面调整。

### 2.3 GMM

GMM 是一种基于概率的聚类算法, 假设每个簇都服从 1 个高斯分布, 而数据集则是由这些高斯分布的线性组合所生成的。

假设数据集中的每个样本都是由 1 个高斯分布随机生成的, 每个高斯分布对应 1 个簇, 其参数为均值  $\mu_k$ 、方差  $\Sigma_k$  和权重  $\alpha_k$ , 满足  $\sum_{k=1}^K \alpha_k = 1$ 。因此, GMM 可以表示为

$$p(\mathbf{x}) = \sum_{k=1}^K \alpha_k G(\mathbf{x}|\mu_k, \Sigma_k) \quad (11)$$

式中:  $p(\mathbf{x})$  为概率密度;  $\mathbf{x}$  为样本数据点;  $K$  为簇的个数;  $G(\mathbf{x}|\mu_k, \Sigma_k)$  为高斯分布的概率密度函数。

$$G(\mathbf{x}|\mu_k, \Sigma_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \mu_k)^T \Sigma_k^{-1}(\mathbf{x} - \mu_k)\right)}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} \quad (12)$$

式中:  $d$  为样本维度;  $|\Sigma_k|$  为方差矩阵的行列式。

GMM 的一个重要特征是不仅给出了每个样本属于哪个簇的最大似然估计, 还给出了每个样本属于每个簇的后验概率。后验概率也称为隐变量, 记为  $\gamma_{nk}$ , 表示第  $n$  个样本属于第  $k$  个簇的概率,  $n=1, 2, \dots, N, N$  为样本总数。

$$p(k|\mathbf{x}) = \frac{p(\mathbf{x}|k)p(k)}{p(\mathbf{x})} = \frac{\alpha_k G(\mathbf{x}|\mu_k, \Sigma_k)}{\sum_{j=1}^K \alpha_j G(\mathbf{x}|\mu_j, \Sigma_j)} \quad (13)$$

式中:  $p(k|\mathbf{x})$  为样本给定簇的条件概率;  $p(k)$  为簇的先验概率;  $p(\mathbf{x})$  为样本的边缘概率。

### 2.4 K-Means 聚类算法

K-Means 是一种基于划分的聚类算法, 假设每个簇由 1 个中心点和其周围的样本组成, 而且每个样本只属于 1 个簇。K-Means 的目标是找到一组簇中心, 使得每个样本与其所属簇中心的距离之和最小, 即最小化以下目标函数:

$$J = \sum_k \sum_{\mathbf{x} \in C_k} \|\mathbf{x} - \mu_k\|^2 \quad (14)$$

式中:  $C_k$  为第  $k$  个簇;  $\|\mathbf{x} - \mu_k\|$  为样本  $\mathbf{x}$  与簇中心  $\mu_k$  的欧氏距离。

### 2.5 PCA 优化算法

煤层钻孔预抽数据共有 8 个评价指标, 然而在高维空间中, 数据点之间的距离变得更加稀疏, 导致聚类算法的性能下降。通过 PCA 优化算法进行降维处理, 以减少计算复杂度, 提高算法效率。

1) 假设原始数据是一个  $N \times d$  矩阵  $\mathbf{X}$ ,  $d$  为维度。

2) 将每个特征的均值和方差分别设为 0 和 1, 使得数据在不同特征上具有相同的尺度。

$$X_{ij} = \frac{X_{ij} - \bar{X}_j}{\sigma_j} \quad (15)$$

式中:  $X_{ij}$  为数据矩阵  $\mathbf{X}$  中第  $i$  行第  $j$  列的元素;  $\bar{X}_j$  为第  $j$  列的均值;  $\sigma_j$  为第  $j$  列的标准差。

3) 计算数据的协方差矩阵:

$$\mathbf{S} = \frac{1}{N} \mathbf{X}^T \mathbf{X} \quad (16)$$

4) 计算协方差矩阵的特征值  $\lambda$  和特征向量  $\mathbf{v}$ :

$$\mathbf{S} \mathbf{v} = \lambda \mathbf{v} \quad (17)$$

5) 选择前  $m$  个最大特征值对应的特征向量作为数据的新坐标轴, 构成一个  $d \times m$  的投影矩阵  $\mathbf{W}$ 。

$$\mathbf{W} = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_m] \quad (18)$$

6) 将数据从原始坐标系转换到新坐标系, 即将数据乘以投影矩阵, 得到一个  $N \times m$  的降维后的数据矩阵  $\mathbf{Y}$ 。

$$Y = XW \quad (19)$$

### 3 钻孔抽采评价结果

利用 AHP-FEM 对 10 个瓦斯抽采样本进行特征标记, 并与剩余 91 个未标记的样本混合, 采用 PCA 进行优化处理后, 分别利用 SSGMM 和 SSK-Means 算法进行聚类分析, 最终聚类结果如图 4 和图 5 所示。

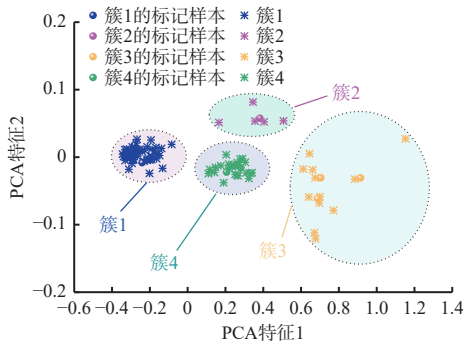


图 4 SSGMM 聚类结果

Fig. 4 SSGMM clustering results

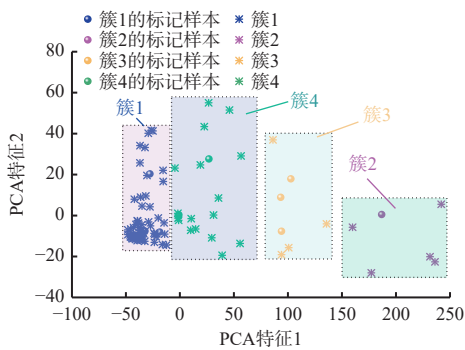


图 5 SSK-Means 聚类结果

Fig. 5 SSK-Means clustering results

SSGMM 和 SSK-Means 聚类算法将钻孔抽采数据分为 4 个聚类。相较于 SSK-Means, SSGMM 的聚类结果阴影面积更小, 聚类密度更大。其中, 聚类 1(蓝色)最密集, 聚类 2(粉色)最分散, 可能受到噪声或异常值的影响。

采用无监督学习中广泛应用的 $\text{max}$ 最大聚集度 (MVCR) 和修正 Rand 指数 (ARI)<sup>[30]</sup> 验证训练结果。

$$\text{MVCR} = \frac{\max\{n_{pk}, k \notin \Theta_p\}}{\sum_k n_{pk}} \times 100\% \quad (20)$$

$$\text{ARI} = \frac{\sum_{pk} \left[ \binom{n_{pk}}{2} - \frac{2ab}{N(N-1)} \right]}{\frac{1}{2}(a+b) - \frac{2ab}{N(N-1)}} \times 100\% \quad (21)$$

式中:  $n_{pk}$  为第  $p$  类样本被聚集到第  $k$  个分量中的样本个数;  $\Theta_p$  为已经被其他类别占据的聚类分量集合;

$a$  为在真实标签和聚类结果中都被分配在同一类簇的样本对个数;  $b$  为在真实标签和聚类结果中都被分配在不同类簇的样本对个数。

聚类算法的评价指标及分类个数见表 5。

表 5 聚类算法的评价指标及分类数

Table 5 Evaluation indicators and classification numbers of clustering algorithms

算法	MVCR/%	ARI/%	分类个数			
			簇1	簇2	簇3	簇4
SSGMM	82.64	85.83	54	6	15	26
SSK-Means	59.10	71.27	69	6	7	19

从表 5 可看出, SSGMM 和 SSK-Means 的聚类结果存在一定差异。GMM 在 MVCR 和 ARI 指数上表现优于 SSK-Means, 分别达到 82.64% 和 85.83%。这是由于 SSGMM 能够针对不同类别的数据采用高斯分布建模, 更能适应实际数据的分布特征, 从而获得更加稳定和有效的聚类结果。在聚类数量方面, 第 3 簇的差异最为显著, SSGMM 聚类为 15 个, SSK-Means 为 7 个, 然而两者总体差异并不显著。针对 325—425 号钻孔 2024 年 3 月 2 日抽采数据, 2 种算法的评价分类结果如图 6 和图 7 所示。

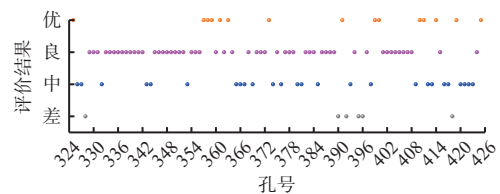


图 6 SSGMM 算法的钻孔评价结果

Fig. 6 Evaluation results of boreholes using SSGMM algorithm

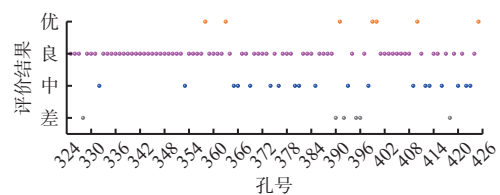


图 7 SSK-Means 算法的钻孔评价结果

Fig. 7 Evaluation results of boreholes using SSK-Means algorithm

结果表明, 2 种算法评价结果高度一致, 说明聚类方法能有效区分钻孔抽采效果。经计算, 试验钻孔抽采效果优良率方面, SSGMM 算法为 59.41%, SSK-Means 算法为 74.26%, 总体抽采效果良好。针对优、良、中、差评语集的相似率分析显示, “优”评语集相似率最低 (46.67%), “差”评语集相似率最高 (100%)。这表明 2 种算法在区分抽采效果较好与较差钻孔时准确性高, 但区分优秀抽采效果时存在差异, 或因聚类结果不同所致, 这可能是由于 SSGMM 需计算每个样本对所有簇的概率密度, 而 SSK-Means

需计算样本与簇中心的欧氏距离，无复杂矩阵运算。由此可见，SSGMM 与 SSK-Means 在瓦斯抽采评价中形成互补关系：SSGMM 以复杂计算换取高精度，适合精准分析及离线分析场景；SSK-Means 以简

单快速满足工程实时性需求。实际应用中需结合数据特性、硬件条件、场景需求灵活选择，或通过混合模型实现平衡。

各钻孔具体评价结果见表 6。

表 6 煤层瓦斯预抽钻孔评价结果

Table 6 Evaluation results of coal seam gas pre-extraction boreholes

模型	钻孔编号			
	优	良	中	差
SSGMM	425, 419, 414, 411, 410, 400, 399, 391, 373, 363, 361, 358, 359, 357, 325	424, 415, 408, 407, 406, 405, 404, 403, 402, 401, 397, 394, 389, 388, 387, 386, 384, 383, 382, 379, 378, 377, 375, 372, 371, 370, 368, 364, 362, 360, 356, 355, 354, 352, 351, 350, 349, 348, 347, 346, 345, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 331, 330, 329	423, 422, 421, 420, 417, 416, 413, 412, 409, 398, 393, 385, 381, 380, 376, 374, 369, 367, 366, 365, 353, 344, 343, 332, 327, 326	418, 396, 395, 392, 390, 328
SSK-Means	425, 410, 400, 399, 391, 363, 358	424, 421, 419, 417, 415, 414, 411, 408, 407, 406, 405, 404, 403, 402, 401, 397, 394, 389, 388, 387, 386, 384, 383, 382, 379, 378, 377, 375, 373, 372, 371, 370, 368, 367, 364, 362, 361, 360, 359, 357, 356, 355, 354, 352, 351, 350, 349, 348, 347, 346, 345, 344, 343, 342, 341, 340, 339, 338, 337, 336, 335, 334, 333, 331, 330, 329, 327, 326, 325	423, 422, 420, 416, 413, 412, 409, 398, 393, 385, 381, 380, 376, 369, 366, 365, 353, 332	418, 396, 395, 392, 390, 328

### 4 应用建议

针对评价结果，对结果为“优”的钻孔布控参数进行了详细分析，并将这些参数设置经验应用于结果为“差”的钻孔。同时，仔细检查了“差”钻孔的抽采状态，结果见表 7。

表 7 状态为“差”的钻孔缺陷

Table 7 Borehole defects with the "poor" status

孔号	堵孔	负压不合理	漏气	封孔不佳
418		√	√	√
396		√		√
395			√	
392	√	√	√	√
390		√	√	
328			√	√

对这些钻孔进行改进，施工现场如图 8 所示。以 4 月 1 日的数据为蓝本，于 4 月 2 日对数据进行了重新评价，评价结果如图 9 和 10 所示。4 月 2 日的评价结果表明，进行优化调整后，原先评价为“差”的钻孔抽采效率明显提升，利用 SSGMM 与 SSK-Means 进行再评价，优良率分别提高了 5.80% 和 5.26%，补差率达 100%。这一优化显著缩短了预抽时间，降低了瓦斯灾害和井下采掘失衡的风险。

通过本文的评价模型，评价为“优”的钻孔可总结其成功经验，指导未来的钻孔布置和抽采策略；对于评价为“良”的钻孔，进行小幅优化以进一步提升抽采效率；对于评价为“中”的钻孔，需分析存在的问题，制定改进方案，如调整抽采参数或优化布孔位置；对于评价为“差”的钻孔，进行全面原因分析，可能需要重新布孔或采用其他抽采方法，以确保瓦斯治理效果。



(a) 钻孔封孔

(b) 钻机打孔



(c) 负压调整

(d) 抽采记录

图 8 试验现场

Fig. 8 Test site

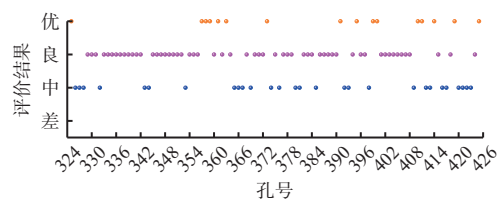


图 9 SSGMM 算法钻孔再评价结果

Fig. 9 Re-evaluation results of boreholes using SSGMM algorithm

### 5 结论

1) 提出了一种基于半监督学习的煤层钻孔预抽瓦斯状态评价方法，融合 AHP 与 FEM 对少量样本进



- 2022/3424367.
- [15] SONG Hongqing, DU Shuyi, YANG Jiaosheng, et al. Evaluation of hydraulic fracturing effect on coalbed methane reservoir based on deep learning method considering physical constraints[J]. *Journal of Petroleum Science and Engineering*, 2022, 212. DOI: 10.1016/j.petrol.2022.110360.
- [16] SOLEIMANI F, SI Guangyao, ROSHAN H, et al. Numerical modelling of gas outburst from coal: a review from control parameters to the initiation process[J]. *International Journal of Coal Science & Technology*, 2023, 10. DOI: 10.1007/s40789-023-00657-7.
- [17] LI Yunzhuo, SU Hetao, JI Huaijun, et al. Numerical simulation to determine the gas explosion risk in longwall goaf areas: a case study of Xutuan Colliery[J]. *International Journal of Mining Science and Technology*, 2020, 30(6): 875-882.
- [18] LIU Peng, FAN Jinyang, JIANG Deyi, et al. Evaluation of underground coal gas drainage performance: mine site measurements and parametric sensitivity analysis[J]. *Process Safety and Environmental Protection*, 2021, 148: 711-723.
- [19] LU Y Y, ZHANG H D, ZHOU Z, et al. Current status and effective suggestions for efficient exploitation of coalbed methane in China: a review[J]. *Energy & Fuels*, 2021, 35(11): 9102-9123.
- [20] ZHOU Aitao, LI Jingwen, GONG Weili, et al. Theoretical and numerical study on the contribution of multi-hole arrangement to coalbed methane extraction[J]. *Energy*, 2023, 284. DOI: 10.1016/j.energy.2023.128598.
- [21] DU Feng, CUI Weilong, WANG Kai. Study on gas migration mechanism and multi-borehole spacing optimization in coal under negative pressure extraction[J]. *Processes*, 2023, 11(1). DOI: 10.3390/pr11010259.
- [22] LIU Jun, ZHANG Luwei, WANG Lin, et al. Negative pressure distribution of variable mass flow in coal mine drainage boreholes[J]. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 2025, 47(1): 1650-1667.
- [23] EL ISLEM KARABADJI N, AMARA KORBA A, ASSI A, et al. Accuracy and diversity-aware multi-objective approach for random forest construction[J]. *Expert Systems with Applications*, 2023, 225. DOI: 10.1016/j.eswa.2023.120138.
- [24] HUANG Qiuhaio, WANG Chao, YE Ye, et al. Recognition of EEG based on improved black widow algorithm optimized SVM[J]. *Biomedical Signal Processing and Control*, 2023, 81. DOI: 10.1016/j.bspc.2022.104454.
- [25] NIE Feiping, CHEN Hong, XIANG Shiming, et al. On the equivalence of linear discriminant analysis and least squares regression[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(4) : 5710-5720.
- [26] 马莲净, 王颂, 赵宝峰, 等. 基于 PCA-SSA-BPNN 模型的矿井突水水源识别方法 [J/OL]. *采矿与安全工程学报*: 1-10[2024-06-18]. <http://kns.cnki.net/kcms/detail/32.1760.TD.20231221.1014.002.html>. MA Lianjing, WANG Song, ZHAO Baofeng, et al. A method for identifying the water inrush sources in mines based on PCA-SSA-BPNN model[J/OL]. *Journal of Mining & Safety Engineering*: 1-10[2024-06-18]. <http://kns.cnki.net/kcms/detail/32.1760.TD.20231221.1014.002.html>.
- [27] AN Yiyao, ZHANG Ke, CHAI Yi, et al. Gaussian mixture variational-based transformer domain adaptation fault diagnosis method and its application in bearing fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2024, 20(1): 615-625.
- [28] NI Xiaoming, ZHAO Zheng, WANG Yanbin, et al. Optimisation and application of well types for ground development of coalbed methane from No. 3 coal seam in Shizhuang south block in Qinshui Basin, Shanxi Province, China[J]. *Journal of Petroleum Science and Engineering*, 2020, 193. DOI: 10.1016/j.petrol.2020.107453.
- [29] LOU Zhen, WANG Kai, KANG M, et al. Plugging methods for underground gas extraction boreholes in coal seams: a review of processes, challenges and strategies[J]. *Gas Science and Engineering*, 2024, 122. DOI: 10.1016/j.jgsce.2024.205225.
- [30] YU Yang, ZENG Jun, ZHONG Lin, et al. Multi-views contrastive learning for dense text retrieval[J]. *Knowledge-Based Systems*, 2023, 274. DOI: 10.1016/j.knsys.2023.110624.