

潘竞涛, 邱晨, 赵丹, 等. 基于模糊 C 均值聚类算法的采空区自燃“三带”划分[J]. 安全与环境学报, 2026, 26(6): 2116–2127.

PAN J T, QIU C, ZHAO D, et al. Division of spontaneous combustion “three zones” in goaf based on fuzzy C -means clustering algorithm[J]. Journal of Safety and Environment, 2026, 26(6): 2116–2127.

基于模糊 C 均值聚类算法的采空区自燃“三带”划分*

潘竞涛¹, 邱晨¹, 赵丹², 刘立仁³, 袁强³

(1 辽宁工程技术大学矿业学院, 辽宁阜新 123000; 2 辽宁工程技术大学安全科学与工程学院, 辽宁葫芦岛 125105;

3 国家能源集团国神三道沟煤矿, 陕西榆林 719407)

摘要: 针对采空区煤自燃危险区域边界模糊、传统划分方法难以有效处理不确定性的问题, 基于模糊集理论, 构建了一套完整的模糊 C 均值(Fuzzy C -Means, FCM)聚类分析框架, 通过引入模糊隶属度机制, 构建融合温度、氧气浓度等多维指标的评价体系, 建立从模糊相似矩阵到等价矩阵的完整数学推导流程, 并借助 MATLAB 实现算法编程与可视化分析。设计基于 λ -截集的动态聚类机制。在聚类优化方面, 提出结合 F 统计量的自适应粒度控制方法, 实现对过渡区域边界的最优识别。结果表明, 调节聚类参数 λ 可有效控制分类粒度, 近工作面区域测点聚类结果对 λ 变化敏感, 反映其空间差异性较大; 远工作面区域测点分类稳定, 体现环境均质特征。进一步利用 F 统计量确定最优 λ 值, 获得的“三带”划分结果: 进风侧氧化带范围为 72~96 m、回风侧为 36~72 m, 与现场“灾害综合智能管控平台”基于氧气阈值的划分结果进风侧 72~103 m、回风侧 36~66 m 相比, 该方法在进风侧识别出 7 m“次氧化区”, 在回风侧延伸出 6 m“潜伏风险区”, 两侧偏差均小于 10%, 验证了划分结果的可靠性。研究不仅实现了与现场数据的一致性, 更通过多指标融合揭示了单一氧气阈值法无法捕捉的风险细节, 为采空区煤自燃危险区域识别提供了一种更精细、更可靠的数据驱动划分手段。

关键词: 安全工程; 模糊聚类; 采空区; 自燃“三带”; 数据分类

中图分类号: X936; TP391

文献标志码: A

文章编号: 1009-6094(2026)06-2116-12

DOI: 10.13637/j.issn.1009-6094.2025.1766

0 引言

煤炭资源的高效开采与安全生产始终是我国能源战略的压舱石^[1]。然而, 在综采过程中形成的采空区内, 由于遗煤分布不均、顶板垮落特征复杂、漏风流场不稳定等因素^[2-4], 煤自燃危险性在空间上呈现高度异质性分布。传统采空区自燃“三带”划分方法多依赖于定点监测与经验推断^[5-7], 对非线性、模糊性较强的监测数据适应性较差, 尤其难以准确界定氧化带与窒息带之间的过渡边界, 从而制约了矿井防灭火工作的精准性与有效性。

随着计算机仿真技术的飞速发展, 计算流体力学(Computational Fluid Dynamics, CFD)技术在研究复杂多物理场问题中展现出强大优势。如 Fluent、COMSOL Multiphysics 等高性能数值模拟软件的问世, 为采空区煤自燃“三带”的精确划分与研究开辟了全新的技术路径^[8-10]。然而, 其分析结果的可靠性受限于模型假设的合理性且强烈依赖于精确的边界条件与物理参数^[11], 而这些参数在复杂的地质与

开采环境中往往难以准确获取, 致使模拟结果可能存在一定不确定性, 同时高昂的计算成本限制了其现场适用性与时效性^[12-14]。聚类算法源于统计学与计算机科学等学科的理论推进, 在处理高维、非线性数据集时展现出卓越的划分能力。2025 年, Zhang 等^[15]首次结合 t 分布随机邻近嵌入(t -Distributed Stochastic Neighbor Embedding, t -SNE)和 K -均值(K -means)聚类方法对采空区自燃数据进行了精确降维和聚类, 细化了传统采空区自燃发火区域, 将预警等级划分为安全、低风险、中风险、高风险、重度风险和极严重风险 6 类。2024 年, 宋冉等^[16]基于 K -means 算法引入三种聚类算法和六个聚类评价指标得到采空区“三带”划分模型。2020 年, 郭松等^[17]提出了一种基于主成分分析(Principal Components Analysis, PCA)降维的改进模糊聚类算法, 对模糊 C 均值聚类模型初始类中心和隶属度参数进行动态优化, 用于评价稳定性影响因素复杂的采空区场地。与聚焦于降低维度以提升收敛性的思路不同, 本文以三道沟煤矿 45207 工作面采空区为

* 收稿日期: 2025-10-10

作者简介: 潘竞涛, 教授, 博士, 从事矿井智能通风与防灭火技术等研究, 163tao@163.com。

基金项目: 国家自然科学基金项目(52374202)

研究对象,针对其复杂环境参数(温度、氧气浓度等)中存在的确定性,完整保留原始数据的维度信息,引入模糊 C 均值(Fuzzy C -Means, FCM)聚类算法进行自燃“三带”的划分。与传统的 K -means 聚类方法相比,FCM 通过引入隶属度函数描述样本点对不同类别的归属程度,更能反映采空区参数渐变的物理特性,尤其适用于处理边界模糊、属性重叠的复杂数据集,该方法可实现自燃“三带”的连续渐变划分,并构建一套可迁移于同类复杂系统的模糊聚类理论模式,从而为多源异构数据驱动的地质灾害识别提供新的方法论支持。

1 现场测点布置

三道沟 45207 综采工作面位于 5⁻²煤层四盘区,地面标高介于 1 148.6 ~ 1 262.2 m,底板标高在 1 073.56 ~ 1 084.41 m 之间,可采长度为 2 338.0 m,煤层倾角为 1° ~ 3°,倾向长度为 276.0 m,设计平均采高为 3.05 m。采空区气体浓度测点及温度测点布置如图 1 所示,在采空区进风侧、回风侧埋设取样探头,设置测点在采空区内沿工作面推进方向均匀分布,随着支架的推进,探头会逐渐深入采空区,地面束管监测系统连续监测采空区束管测点的 O_2 等各种气体含量。同时设置隔爆兼本安型分布式光纤测温主机,采用光纤传感技术对各测点进行温度监测。工作面每个测点束管铺设预埋管时用 DN50 钢管作为保护套管,每根束管负责抽取一个测点的气样。为了防止采空区积水或粉尘堵塞束管,每个探头抬高 1.5 m 以上(根据采空区积水情况适当调整高度),并将取气端头方向朝下放置。

2 模糊聚类模型构建

2.1 基本原理

模糊聚类核心源于模糊集理论,用以处理现实

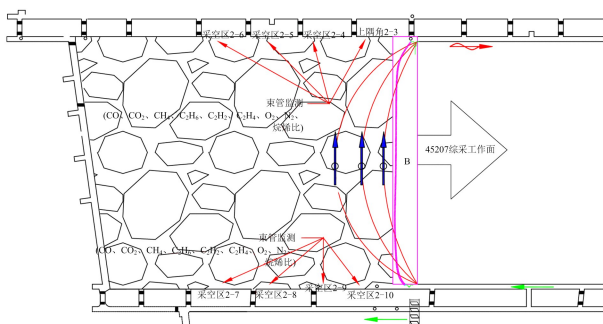


图 1 采空区束管监测点布置示意图

Fig. 1 Schematic diagram of bundled-tube monitoring point arrangement in goaf

世界普遍存在、类别界限不分明数据的划分问题。常用的模糊聚类方法有:(1)基于模糊等价矩阵聚类方法(传递闭包法、Bool 矩阵法);(2)基于模糊相似矩阵的直接聚类法(直接聚类法、最大树法、编网法);(3)基于目标函数的模糊聚类分析法^[18-21]。与传统硬聚类(如 K -means)要求每个数据点仅能归属于一个类别不同,模糊聚类的根本原则是引入“隶属度”(membership degree)的概念,通过迭代优化目标函数得到每个数据点对所有类别的隶属度,根据当前各个数据点到各个聚类中心的距离计算得出模糊矩阵^[22-23]。距离聚类中心越近的点对该类的隶属度越高。其目标函数 J_m 为

$$J_m = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m d_{ij}^2 \quad (1)$$

式中 n 为样本总数; c 为预设聚类数目; u_{ij} 表示第 i 个样本隶属于第 j 个聚类的隶属度,且满足 $\sum_{j=1}^c u_{ij} = 1, u_{ij} \in [0, 1]$; m 为模糊加权指数,用于控制聚类结果的模糊程度; $d_{ij} = \|\mathbf{x}_i - \mathbf{v}_j\|$ 为第 i 个样本 $\mathbf{x}_i$ 到第 j 个聚类中心 $\mathbf{v}_j$ 的欧式距离。通过对目标函数应用拉格朗日乘子法进行优化,可推导隶属度 u_{ij} 与聚类中心 $\mathbf{v}_j$ 的迭代更新公式为

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}}} \quad (2)$$

$$\mathbf{v}_j = \frac{\sum_{i=1}^n (u_{ij})^m \mathbf{x}_i}{\sum_{i=1}^n (u_{ij})^m} \quad (3)$$

式中 k 为迭代步数。通过交替更新隶属度矩阵 $\mathbf{U} = (u_{ij})_{n \times c}$ 和聚类中心 $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\}$ 直至收敛,最终得到每个样本点对于各个“三带”类别的隶属程度,从而实现模糊划分。上述目标函数与迭代公式构成了标准 FCM 算法的理论基础,基于模糊 C 均值聚类采空区自燃“三带”划分流程如图 2 所示。

2.2 评价指标选取

为科学划分采空区煤自燃危险区域,综合选取多项关键评价指标^[24]。核心指标包括:用于界定氧化带的氧气体积分数、作为早期预警信号的一氧化碳体积分数、表征煤温变化阶段的烃类气体体积分数比值(烷烃比)、直接反映积热状态的采空区温度。此外,随工作面推进的束管埋设深度(L)也作为动态观测变量纳入考虑。鉴于现场束管监测数据

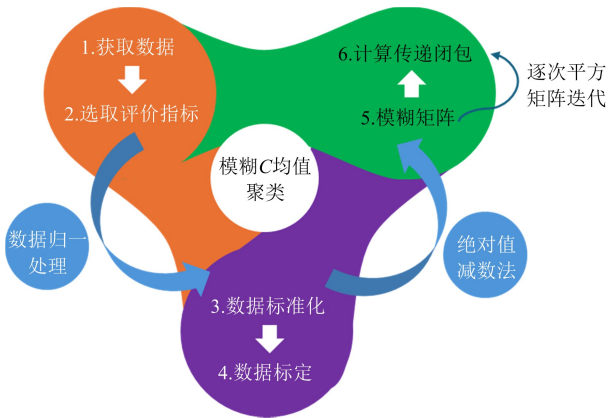


图2 模糊C均值算法流程图

Fig. 2 Flowchart of the fuzzy C-means algorithm

表明,乙烷(C_2H_6)体积分数长期为0,不具备有效的指示意义,此处不予选取。为验证本文所构建方法的稳定性与普适性,在完成核心测点的详细分析后,进一步对采空区内其他位置的测点数据进行了初步的聚类验证,其他测点所展现出的“三带”演化规律与文中所呈现的结论基本一致。由于测点及数据较多且分析过程类似,此处选取45207进风侧测点2-10、回风侧测点2-3的数据进行分析,具体数据见表1。

得到评价矩阵 R_{2-10} 、 R_{2-3} ,即

$$R_{2-10} = \begin{pmatrix} 18.21 & 25 & 0.02 & 14.3 & 10.2 \\ 17.59 & 30 & 0.11 & 14.6 & 17.5 \\ 17.24 & 31 & 0.14 & 15.1 & 23.7 \\ 16.89 & 34 & 0.15 & 15.9 & 29.3 \\ 15.22 & 34 & 0.20 & 16.4 & 36.6 \\ 14.61 & 35 & 0.31 & 16.5 & 41.7 \\ 14.26 & 36 & 0.42 & 16.9 & 45.9 \\ 13.9 & 38 & 0.51 & 17.1 & 52.4 \\ 12.9 & 39 & 0.74 & 17.6 & 60.1 \\ 12.1 & 39 & 0.93 & 18.0 & 67.5 \\ 12 & 39 & 1.08 & 18.4 & 72.4 \\ 11.4 & 41 & 1.25 & 19.1 & 78.9 \\ 11.1 & 41 & 1.39 & 19.2 & 82.3 \\ 10.7 & 42 & 1.56 & 19.7 & 89.6 \\ 10 & 43 & 1.64 & 20.1 & 95.4 \\ 9.2 & 43 & 1.68 & 20.5 & 100.3 \\ 8.1 & 45 & 1.71 & 20.9 & 105.4 \\ 7.2 & 46 & 1.84 & 21.3 & 110.1 \\ 6.3 & 47 & 1.87 & 21.7 & 114.5 \\ 5.1 & 47 & 1.91 & 21.9 & 120.3 \\ 4.1 & 48 & 1.95 & 22.1 & 124.7 \end{pmatrix}$$

$$R_{2-3} = \begin{pmatrix} 17.24 & 15 & 0.01 & 14.3 & 10.7 \\ 17.33 & 20 & 0.1 & 14.5 & 19.5 \\ 18.21 & 21 & 0.13 & 15.1 & 24.1 \\ 16.8 & 24 & 0.14 & 15.8 & 29.7 \\ 16.1 & 24 & 0.19 & 16.3 & 38.6 \\ 14.79 & 35 & 0.3 & 16.5 & 43.7 \\ 14.35 & 36 & 0.41 & 16.8 & 46.9 \\ 13.9 & 37 & 0.51 & 17.1 & 53.4 \\ 12.7 & 39 & 0.74 & 17.6 & 58.6 \\ 12.2 & 39 & 0.93 & 18.0 & 67.5 \\ 12 & 42 & 1.08 & 18.4 & 73.4 \\ 11.6 & 41 & 1.25 & 19.1 & 79.9 \\ 11.2 & 39 & 1.39 & 19.3 & 83.3 \\ 10.9 & 42 & 1.55 & 19.7 & 90.6 \\ 9.9 & 45 & 1.64 & 20.4 & 95.4 \\ 9.2 & 40 & 1.68 & 20.7 & 101.3 \\ 9.1 & 45 & 1.71 & 21.1 & 106.4 \\ 8.8 & 45 & 1.76 & 21.5 & 111.1 \\ 8.5 & 47 & 1.8 & 21.9 & 115.5 \\ 8.4 & 47 & 1.85 & 21.9 & 122.3 \\ 7.8 & 48 & 1.9 & 22.1 & 126.7 \end{pmatrix}$$

2.3 数据标准化

由于划分“三带”的评价指标量纲各不相同,需要对数据进行标准和归一化处理,计算式分别为

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad (4)$$

$$x''_{ij} = \frac{x'_{ij} - x'_{j\min}}{x'_{j\max} - x'_{j\min}} \quad (5)$$

式中 $\bar{x}_j$ 为 f 个样本中指标 j 的平均值, S_j 为其标准差,计算式分别见式(6)和(7); $x'_{j\min}$ 和 $x'_{j\max}$ 分别为 $x'_{1j}, x'_{2j}, \dots, x'_{fj}$ 中最小值和最大值; x'_{ij} 为标准化后的指标值; x''_{ij} 为归一化后的指标值。

$$\bar{x}_j = \frac{1}{f} \sum_{i=1}^f x_{ij} \quad (6)$$

$$S_j = \sqrt{\sum_{i=1}^f (x_{ij} - \bar{x}_j)^2} \quad (7)$$

2.4 数据标定

为便于后续聚类分析识别,需对原始数据进行标定以消除量纲影响,构建元素间相似关系。数据标定即在论域 $X = \{x_1, x_2, \dots, x_n\}$ 上定义一个相似关系矩阵 R ,其中的每个元素 r_{ij} 称为相似系数,代表 x_i 与 x_j 之间的相似程度,且满足 $0 \leq r_{ij} \leq 1$ 。特别地,当 $i = j$ 时,有 $r_{ii} = 1$ 表示自身比较或完全相似,数值越小则相似度越低。本次标定采用绝对值减数法,通过计算特征空间中的曼哈顿距离来反推相似

表 1 45207 进风侧 2-10 测点与回风侧 2-3 测点气体体积分数、温度及埋管深度 L Table 1 Gas volume fraction, temperature, and buried pipe depth L at monitoring points 2-10 on the intake side and 2-3 on the return side of 45207 working face

日期	$\varphi(\text{O}_2)/\%$		$\varphi(\text{CO})/10^{-6}$		$\varphi(\text{CH}_4)/\%$		$\theta/^\circ\text{C}$		L/m	
	2-10	2-3	2-10	2-3	2-10	2-3	2-10	2-3	2-10	2-3
4 月 15 日	18.21	17.24	25	15	0.02	0.01	14.3	14.3	10.2	10.7
4 月 16 日	17.59	17.33	30	20	0.11	0.1	14.6	14.5	17.5	19.5
4 月 17 日	17.24	18.21	31	21	0.14	0.13	15.1	15.1	23.7	24.1
4 月 18 日	16.89	16.8	34	24	0.15	0.14	15.9	15.8	29.3	29.7
4 月 19 日	15.22	16.1	34	24	0.20	0.19	16.4	16.3	36.6	38.6
4 月 20 日	14.61	14.79	35	35	0.31	0.3	16.5	16.5	41.7	43.7
4 月 21 日	14.26	14.35	36	36	0.42	0.41	16.9	16.8	45.9	46.9
4 月 22 日	13.9	13.9	38	37	0.51	0.51	17.1	17.1	52.4	53.4
4 月 23 日	12.9	12.7	39	39	0.74	0.74	17.6	17.6	60.1	58.6
4 月 24 日	12.1	12.2	39	39	0.93	0.93	18.0	18.0	67.5	67.5
4 月 25 日	12	12	39	42	1.08	1.08	18.4	18.4	72.4	73.4
4 月 26 日	11.4	11.6	41	41	1.25	1.25	19.1	19.1	78.9	79.9
4 月 27 日	11.1	11.2	41	39	1.39	1.39	19.2	19.3	82.3	83.3
4 月 28 日	10.7	10.9	42	42	1.56	1.55	19.7	19.7	89.6	90.6
4 月 29 日	10	9.9	43	45	1.64	1.64	20.1	20.4	95.4	95.4
4 月 30 日	9.2	9.2	43	40	1.68	1.68	20.5	20.7	100.3	101.3
5 月 1 日	8.1	9.1	45	45	1.71	1.71	20.9	21.1	105.4	106.4
5 月 2 日	7.2	8.8	46	45	1.84	1.76	21.3	21.5	110.1	111.1
5 月 3 日	6.3	8.5	47	47	1.87	1.8	21.7	21.9	114.5	115.5
5 月 4 日	5.1	8.4	47	47	1.91	1.85	21.9	21.9	120.3	122.3
5 月 5 日	4.1	7.8	48	48	1.95	1.9	22.1	22.1	124.7	126.7

度,公式为

$$r_{ij} = 1 - \alpha d_{ij} \quad (8)$$

式中 d_{ij} 为样本 i 与 j 之间的绝对值距离, $d_{ij} = \sum_{k=1}^{m^*} |x_{ik} - x_{jk}|$; x_{ik} 、 x_{jk} 分别为第 i 个样本和第 j 个样本关于第 k 个评价指标的标准化观测值; m^* 为评价指标的总数; α 为尺度因子,根据所选的评价指标参数取 0.2,将距离信息 d_{ij} 线性映射到标准的相似度空间,使得相似系数的分布覆盖有效范围。

数据的标准化计算量较大,数据标定借助 MATLAB 实现,得到 21 阶相似模糊矩阵 $\mathbf{R}'_{2-10}$ 、 $\mathbf{R}'_{2-3}$,模糊相似矩阵可视化图见图 3。

2.5 传递闭包 $t(R)$ 计算

采用逐次平方法分别计算 $\mathbf{R}_{2-10}$ 、 $\mathbf{R}_{2-3}$ 的传递闭

包 $t(R)$,当 $\mathbf{R}_{2-10}^{2(n+1)} = \mathbf{R}_{2-10}^{2n}$ 时停止迭代,此时传递闭包矩阵不再发生变化,满足收敛条件。迭代过程采用 MATLAB 实现,经计算 $\mathbf{R}_{2-10}$ 、 $\mathbf{R}_{2-3}$ 均收敛于第 4 次迭代,得到模糊等价矩阵 $\mathbf{R}_{2-10}^*$ 、 $\mathbf{R}_{2-3}^*$,等价模糊矩阵可视化图见图 4。

2.6 不同 λ 下 X 的聚类分析

将表 1 数据按行进行编号,在传递闭包的基础上,结合 λ -截集进行聚类分析,可将模糊相似关系转换为清晰的等价关系,从而实现样本的划分,具体步骤如图 5 所示。

当 λ 过大时,取样点数据单独成类;当 λ 过小时,所有取样点数据归为一类。因此, λ 需在一定范围内取值。通过 MATLAB 算法遍历 λ 并记录聚类数目,来确定 λ 合适取值。 λ 对聚类数目的影响见

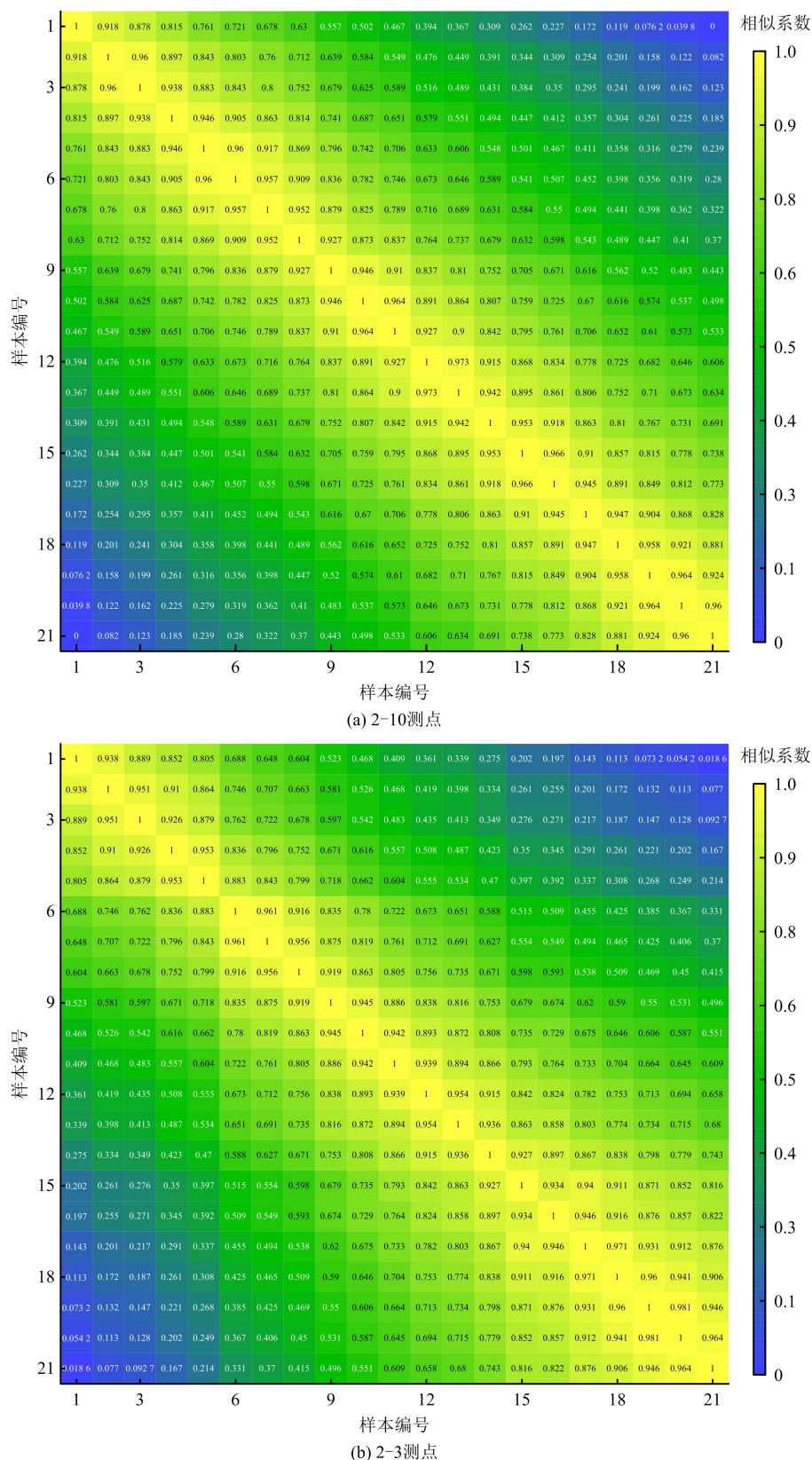


图3 束管监测点模糊相似矩阵分布

Fig. 3 Distribution of fuzzy similarity matrices for bundled-tube monitoring points

图6。

由图6可知, λ_{2-10} 在 0.92 ~ 0.96 范围内、 λ_{2-3}

在 0.89 ~ 0.96 范围内够有效地将取样点进行分类。

取 λ_{2-10} 为 0.85、0.86、0.87、0.88、0.89、0.90、0.91、

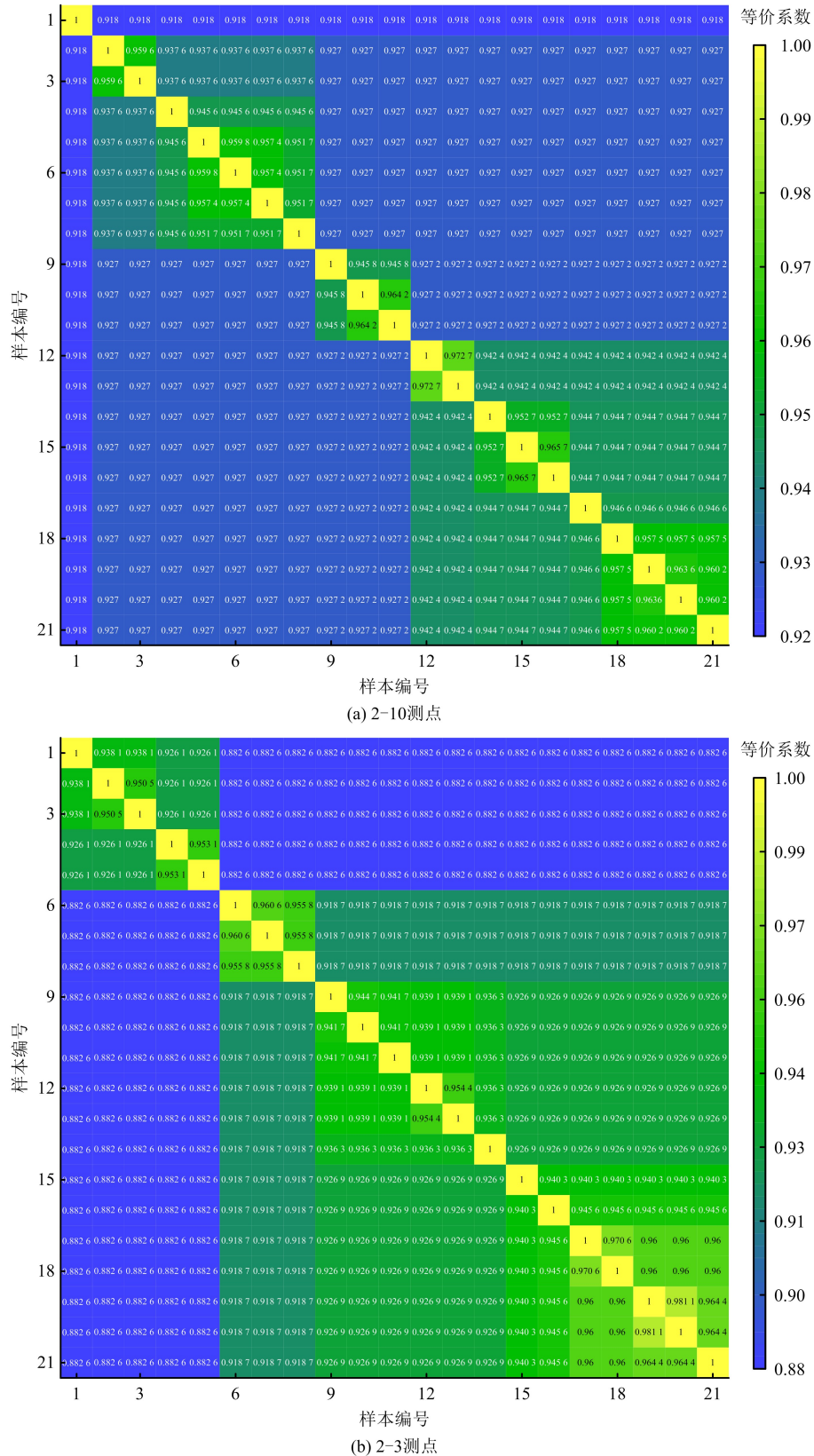


图 4 束管监测点等价模糊矩阵分布

Fig. 4 Distribution of fuzzy equivalence matrices for bundled-tube monitoring points

0.92、0.93、0.94、0.95, λ_{2-3} 为 0.83、0.84、0.85、0.86、0.87、0.88、0.89、0.90、0.91、0.92, 得到不同

截距阵,按照不同截距阵将 X 分成不同类别,见表 2 和 3,其截距阵分布见图 7。

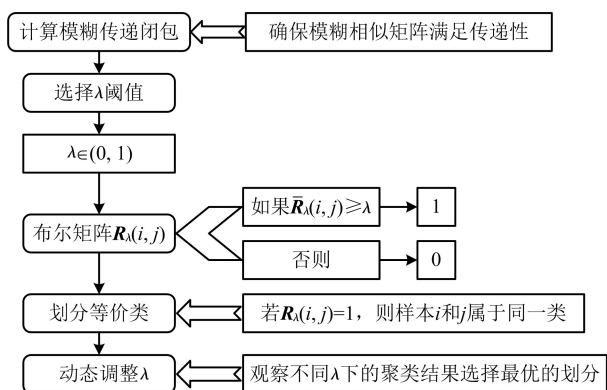


图5 λ-截集聚类分析流程图

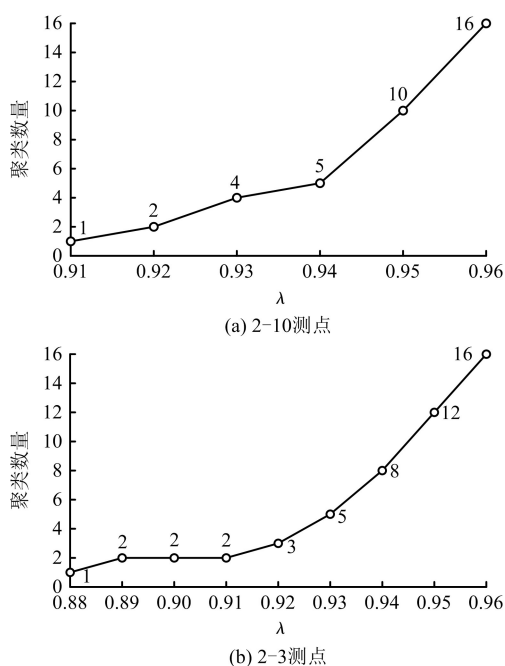
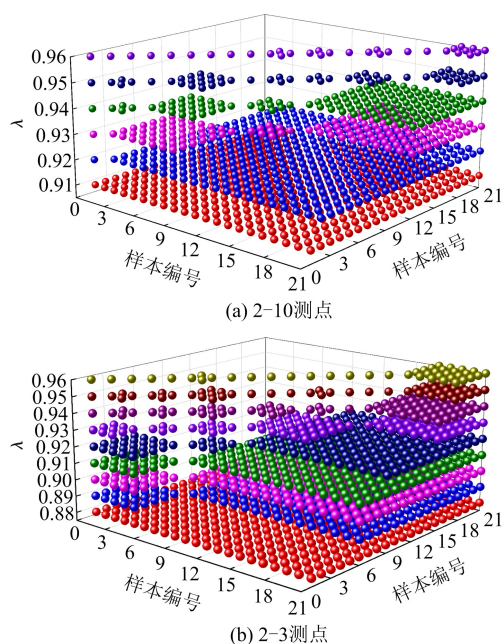
Fig. 5 Flowchart of λ -cut clustering analysis图6 不同测点 λ -截集聚类分析Fig. 6 λ -cut clustering analysis for different monitoring points

图7 λ-截集矩阵散点分布图

Fig. 7 Scatter plots of λ -cut matrices

综上所述,靠近采空区的中、后部组别性质稳定,难以划分不同的子类,如2-10测点的{4, 5, 6, 7, 8} {15, 16} {19, 20, 21}、2-3测点的{6, 7, 8} {17, 18, 19, 20, 21},而靠近工作面区域的采空区组别性质差异较大,易受不同 λ 影响划为较小的子类。

2.7 F 统计量检验计算

F 统计量法又称 F 检验,该方法主要用于两个样本的比较,得出两个样本之间的差异与各自样本的标准差之间的比值。F 统计量越大,两个样本之间的差异越明显。设分类的全体样本集为 $X = \{x_1, x_2, \dots, x_n\}$,其中每个样本 x_j 是一个 m 维特征向量,给定参数 λ 可得到类数 k ,第 i 类($i = 1, 2, 3, \dots$,

表2 2-10 测点数据分类

Table 2 Data classification for monitoring points 2-10

编号	λ	组数	具体分类
1	$\lambda \leq 0.91$	1	X
2	0.92	2	{1} {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21}
3	0.93	4	{1} {2, 3, 4, 5, 6, 7, 8} {9, 10, 11} {12, 13, 14, 15, 16, 17, 18, 19, 20, 21}
4	0.94	5	{1} {2, 3} {4, 5, 6, 7, 8} {9, 10, 11} {12, 13, 14, 15, 16, 17, 18, 19, 20, 21}
5	0.95	10	{1} {2, 3} {4, 5, 6, 7, 8} {9} {10, 11} {12, 13} {14, 15, 16} {17} {18, 19, 20, 21}
6	0.96	16	{1} {2} {3} {4} {5} {6} {7} {8} {9} {10, 11} {12, 13} {14} {15, 16} {17} {18} {19, 20, 21}

表 3 2-3 测点数据分类

Table 3 Data classification for monitoring points 2-3

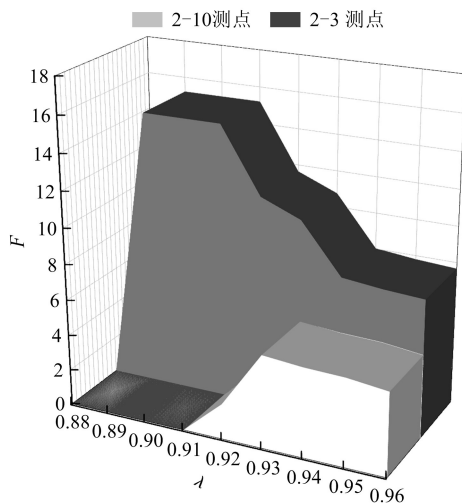
编号	λ	组数	具体分类
1	$\lambda \leq 0.88$	1	X
2	0.89, 0.90, 0.91	2	$\{1, 2, 3, 4, 5\} \{6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$
3	0.92	3	$\{1, 2, 3, 4, 5\} \{6, 7, 8\} \{9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$
4	0.93	5	$\{1, 2, 3\} \{4, 5\} \{6, 7, 8\} \{9, 10, 11, 12, 13, 14\} \{15, 16, 17, 18, 19, 20, 21\}$
5	0.94	8	$\{1\} \{2, 3\} \{4, 5\} \{6, 7, 8\} \{9, 10, 11\} \{12, 13\} \{14\} \{15, 16, 17, 18, 19, 20, 21\}$
6	0.95	12	$\{1\} \{2, 3\} \{4, 5\} \{6, 7, 8\} \{9\} \{10\} \{11\} \{12, 13\} \{14\} \{15\} \{16\} \{17, 18, 19, 20, 21\}$
7	0.96	16	$\{1\} \{2\} \{3\} \{4\} \{5\} \{6, 7\} \{8\} \{9\} \{10\} \{11\} \{12\} \{13\} \{14\} \{15\} \{16\} \{17, 18, 19, 20, 21\}$

k) 包含 n_i 个样本, 记作 $X^i = \{x_1^i, x_2^i, \dots, x_{n_i}^i\}$, 其公式为

$$F = \frac{M_b}{M_w} = \frac{\frac{S_b}{k-1}}{\frac{S_w}{n-k}} = \frac{\sum_{i=1}^k \frac{n_i \|v_i - \bar{x}\|^2}{k-1}}{\sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\|x_j - v_i\|^2}{n-k}} \quad (9)$$

式中 M_b 、 M_w 分别为组间、组内均方; S_b 、 S_w 分别为组间、组内平方和; v_i 为第 i 类的聚类中心, $\bar{x}$ 为全体元素的总体中心; 组间方差自由度为 $k-1$, 组内方差自由度为 $n-k$ 。

F 统计量越大, 表明聚类效果越优。因此, 使 F 达到最大值的 λ 所对应的分类数 k 即为最佳分类数, 不同 λ 对应的 F 如图 8 所示。

图 8 不同 λ 值下的 F 统计值Fig. 8 F -statistic values at different λ values

上述 F 统计量计算为单个测点的聚类效果提供了量化评估。为对整个采空区“三带”的统一划分, 必须从一个统一的 λ 值出发。然而, 不同测点

的 F 统计量峰值往往对应着不同的 λ 区间。为从中遴选出适用于整个工作面的统一最优 λ 值, 可将每个测点视为一个独立的优化目标, 其目标是使该测点数据在聚类划分下的 F 统计量最大化。确定每个测点的 F 统计量峰值或高位值所对应的 λ 区间, 这些区间共同构成了一个帕累托最优解集。在此基础上, 通过计算并最大化所有测点 F 统计量的几何平均值 $F_G(\lambda)$, 从而在权衡各测点聚类效果的基础上, 客观地遴选出适用于整个工作面的统一最优 λ 值。

$$F_G(\lambda) = \sqrt[p]{\prod_{p=1}^p F_p(\lambda)} \quad (10)$$

式中 $F_p(\lambda)$ 为第 p 个测点在给定 λ 下的 F 统计量。

当 $\lambda_{2-10} = 0.93$ 时, F_{2-10} 取得极大值 5.19; 当 $\lambda_{2-3} = \{0.89, 0.90, 0.91\}$ 时, F_{2-3} 取得极大值 15.21, 在上述区间内不存在 λ 使得 F_{2-10} 与 F_{2-3} 同时取得最大值。因此, 考虑几何平均最优点, 选取最接近理想点 $\lambda = 0.93$, 此时测点 2-10、2-3 的截距矩阵及聚类分布情况见图 9, 2-10 测点数据分类为 4 类, 2-3 测点数据分类为 5 类。过于细致的分类使得应用过于烦琐, 分类失去应用价值, 此时将最靠近工作面、中间过渡区域、最远离工作面的测点合并分类, 2-10 测点分类结果为 $\{1, 2, 3, 4, 5, 6, 7, 8\} \{9, 10, 11\} \{12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$, 2-3 测点分类结果为 $\{1, 2, 3, 4, 5\} \{6, 7, 8\} \{9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$ 。

根据上述分类结果, 45207 采空区自燃“三带”范围如图 10 粗实线所示, 第 I 类散热带对应聚类结果靠近工作面的类别。该区域漏风强度较大, 其氧气浓度较高, 气流将煤氧化产生的热量迅速带走, 无

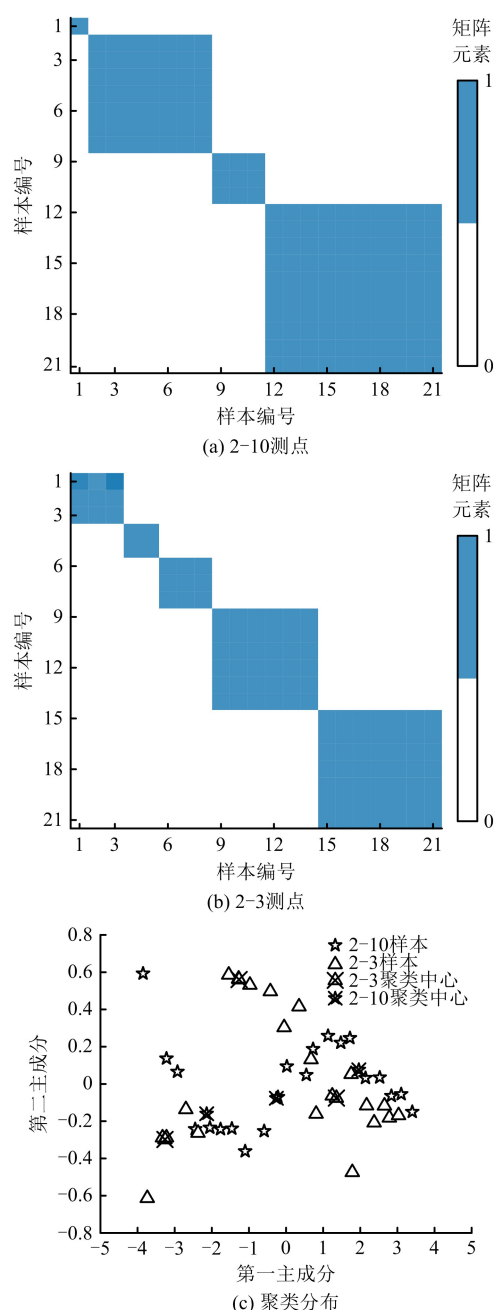


图9 $\lambda = 0.93$ 的 λ -截集矩阵分布及聚类情况

Fig. 9 Distribution of λ -cut matrices and clustering status at $\lambda = 0.93$

法积聚,结合表1煤体温度与周围环境温度相近,为14~17℃,不具备自燃条件。聚类分析通过较低的样本间相似性准确捕捉到了该区域因漏风导致的空间异质性。第Ⅱ类氧化带对应聚类结果中间过渡类别,此区域是煤自燃发生的关键区域,该区域的氧气浓度虽低于散热带,但足以支持剧烈的煤氧复合反应,同时此区域漏风强度减弱,对流散热小于煤氧化产热,导致热量在此持续积聚,温度从17℃稳步升至19℃,CO等指标性气体浓度显著增加。聚类算法通过稳定的隶属关系,将具有相似热-流-化学耦合特征的样本聚合,圈定了自燃风险最高的空间范围。第Ⅲ类窒息带对应聚类结果距工作面最远的类别,该区域由于深埋采空区内部,漏风微弱甚至消失,氧气浓度持续下降至煤氧化反应所需浓度以下,煤氧化过程被抑制。环境趋于均一稳定,因此该区域样本在聚类分析表现出极高的相似性,分类结果对于 λ 值变化不敏感,形成一个稳定的大类。本文方法划分的“三带区域”与基于氧气浓度阈值经典判别方法的灾害综合智能管控平台“三带”划分范围对比见表4。

对比结果表明,本文方法与平台在氧化带起点判定上完全一致,但在终点判定上存在系统性优化。在进风侧,本方法将氧化带边界精准收缩7 m,有效识别并剔除了氧气浓度达标但氧化反应微弱的“次氧化区”;在回风侧,则将边界延伸6 m,识别出因热惯性而存在的“潜伏风险区”。通过精确的划分,凸显了本文方法融合多维指标、实现风险精细化识别的核心优势,为矿井防灭火的精准决策提供了更科学的依据。

3 结论

(1) 本文采用模糊C均值聚类分析方法考虑多组分气体浓度、温度、埋管距离多个指标,通过计算模糊相似、等价矩阵并进行F统计量检验,为采空区“三带”划分提供了一种数学统计的方法,得到进风

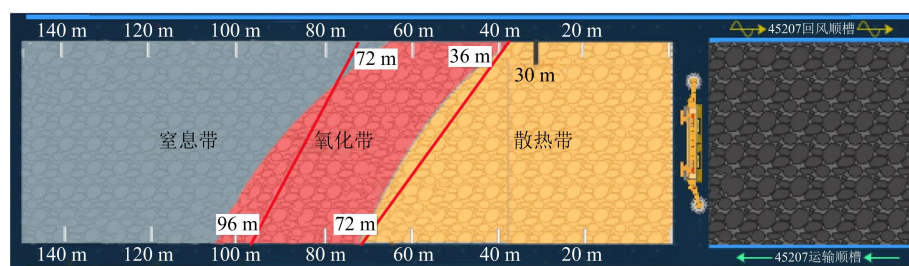


图10 基于模糊C均值聚类的采空区自燃“三带”划分

Fig. 10 Division of spontaneous combustion "three zones" in goaf based on fuzzy C-means clustering

表 4 FCM 方法与平台氧化带边界对比

Table 4 Comparison of oxidation zone boundaries between FCM method and platform

侧边	边界	本文方法/m	平台结果/m	绝对偏差/m	相对偏差/%
进风侧	起始点	72	72	0	0
	终止点	96	103	7	6.8
	范围长度	24	31	7	22.6
回风侧	起始点	36	36	0	0
	终止点	72	66	6	9.1
	范围长度	36	30	6	20.0

侧氧化带边界 72 ~ 96 m、回风侧氧化带边界 36 ~ 72 m,在进风侧终点判定上缩短 6.8% 绝对偏差,避免了将防灭火资源(如注氮、灌浆)浪费在实际氧化反应微弱、风险较低的“次氧化区”,实现了资源的精准配置,在回风侧通过识别“潜伏风险区”实现了安全边界的拓展。

(2)模糊 C 均值聚类分析优势在于能够处理类别间的模糊过渡。随着 λ 取值提高,初始聚类逐渐分裂为更精细的子类别。此类变化主要出现在采空区内靠近当前开采工作面的区段,而在距离工作面较远的监测点,其数据分类结果则保持相对恒定,未进一步产生细分。因此,相比于采空区深部,靠近工作面的区域在环境属性上表现出更大的异质性,反映出该区域受到更显著的工作面漏风影响。

(3)当 λ 变化时可以得到多种不同的聚类结果,通过 F 统计量确定最佳的聚类阈值。文中选取了 5 种评价指标,为更为细致地分析采空区性质,可增加其他辅助性评价指标并给予相应权重考虑,同时进一步增加测点密度,构建更为精细的采空区自燃危险区域可视化模型。

参考文献 (References):

[1] 葛世荣,樊静丽,刘淑琴,等. 低碳化现代煤基能源技术体系及开发战略[J]. 煤炭学报, 2024, 49(1): 203 - 223.

GE S R, FAN J L, LIU S Q, et al. Low carbon modern coal-based energy technology system and development strategy[J]. Journal of China Coal Society, 2024, 49(1): 203 - 223.

[2] 王正帅. 复杂通风综采面采空区“三带”分布及防灭火参数研究[J]. 矿业研究与开发, 2023, 43(8): 73 - 77.

WANG Z S. Study on the distribution of “three zones” and fire prevention and extinguishing parameters in goaf of

complex ventilation fully mechanized mining face[J]. Mining Research and Development, 2023, 43(8): 73 - 77.

[3] 张晓旭,罗伙根. 寸草塔二矿注氮防灭火数值模拟研究[J]. 中国安全科学学报, 2022, 32(增刊2): 131 - 135.

ZHANG X X, LUO H G. Numerical simulation research on nitrogen injection for fire prevention in Cuncaota No. 2 mine[J]. China Safety Science Journal, 2022, 32(Suppl. 2): 131 - 135.

[4] 王经玺. 大采高综采工作面沿空留巷防灭火实践[J]. 煤炭科学技术, 2021, 49(增刊2): 168 - 173.

WANG J X. Practice of fire prevention for gob-side entry driving in fully-mechanized mining faces with large mining heights[J]. Coal Science and Technology, 2021, 49(Suppl. 2): 168 - 173.

[5] 谭刚,宋国栋,刘宇,等. 基于 AHP - FCE 的煤矿火灾风险评价研究[J]. 工矿自动化, 2025, 51(7): 52 - 58.

TAN G, SONG G D, LIU Y, et al. Coal mine fire risk assessment based on AHP - FCE[J]. Journal of Mine Automation, 2025, 51(7): 52 - 58.

[6] 胡海峰,陈明浩,李晨光,等. 沿空留巷采空区注氮防灭火参数优化[J/OL]. 安全与环境学报: 1 - 14 [2025 - 09 - 21]. <https://doi.org/10.13637/j.issn.1009-6094.2025.0776>.

HU H F, CHEN M H, LI C G, et al. Optimization of nitrogen injection parameters for fire prevention and extinction in the gob-side entry retaining goaf[J/OL]. Journal of Safety and Environment: 1 - 14 [2025 - 09 - 21]. <https://doi.org/10.13637/j.issn.1009-6094.2025.0776>.

[7] 岳宁芳,金彦,陈桂林,等. 基于多参量的采空区煤自燃危险区域划分方法研究[J]. 矿业安全与环保, 2021, 48(4): 1 - 5, 11.

YUE N F, JIN Y, CHEN G L, et al. Research on the

- division method of the dangerous area of coal spontaneous combustion in goaf based on multi-parameter[J]. Mining Safety & Environmental Protection, 2021, 48(4): 1 – 5, 11.
- [8] WANG C P, CHEN L J, BAI Z J, et al. Study on the dynamic evolution law of spontaneous coal combustion in high-temperature regions[J]. Fuel, 2022, 314: 123036.
- [9] ZHANG P Y, MA L, SUN J X, et al. Numerical investigation of air flow field evolution and leakage patterns in large-scale goaf areas[J]. Physics of Fluids, 2025, 37(2): 027151.
- [10] ZAHNG Y S, NIU K, DU W Z, et al. A method to identify coal spontaneous combustion-prone regions based on goaf flow field under dynamic porosity[J]. Fuel, 2021, 288: 119690.
- [11] WEI W, ZHAO Q, JIANG Q H, et al. Three new boundary conditions for the seismic response analysis of geomechanics problems using the numerical manifold method[J]. International Journal of Rock Mechanics and Mining Sciences, 2018, 105: 110 – 122.
- [12] 徐世乾, 郭建春, 曾凡辉, 等. 代理优化算法驱动的致密油压裂施工参数优化方法[J/OL]. 控制理论与应用: 1 – 10[2025 – 12 – 11]. <https://link.cnki.net/urlid/44.1240.TP.20250903.1534.002>.
- XU S Q, GUO J C, ZENG F H, et al. Optimization method of fracturing treatment parameters driven by surrogate optimization algorithm in tight oil[J/OL]. Control Theory & Applications: 1 – 10[2025 – 12 – 11]. <https://link.cnki.net/urlid/44.1240.TP.20250903.1534.002>.
- [13] 韩省思, 吴文昌, 陈涛, 等. 航空动力系统热流体高精度混合湍流模拟方法若干进展[J/OL]. 推进技术: 1 – 34[2025 – 12 – 11]. <https://link.cnki.net/urlid/11.1813.V.20250610.1034.001>.
- HAN X S, WU W C, CHEN T, et al. Some progresses in high-fidelity hybrid turbulence simulation methods for thermofluids in aviation propulsion[J/OL]. Journal of Propulsion Technology: 1 – 34[2025 – 12 – 11]. <https://link.cnki.net/urlid/11.1813.V.20250610.1034.001>.
- [14] 冯国庆, 常海铃, 王苛宇, 等. 基于ANN算法的微纳米水气分散体系驱产油量预测方法[J/OL]. 油气藏评价与开发: 1 – 9[2025 – 12 – 11]. <https://link.cnki.net/urlid/32.1825.TE.20250530.1507.002>.
- FENG G Q, CHANG H L, WANG K Y, et al. Prediction method of oil production driven by micro-nano water and gas dispersion system based on ANN algorithm[J/OL]. Petroleum Reservoir Evaluation and Development: 1 – 9[2025 – 12 – 11]. <https://link.cnki.net/urlid/32.1825.TE.20250530.1507.002>.
- [15] ZHANG P, CHEN X. An unsupervised learning approach for coal spontaneous combustion warning level classification using t -SNE and k -means clustering[J]. Applied Sciences, 2025, 15(7): 3756.
- [16] 宋冉, 王彤. 基于 K -means 算法的采空区“三带”划分[J]. 新疆钢铁, 2024(1): 83 – 85.
- SONG R, WANG T. “Three Zones” division of goaf based on K -means algorithm[J]. Xinjiang Iron and Steel, 2024(1): 83 – 85.
- [17] 郭松, 郭广礼, 李怀展, 等. 基于降维模糊 C 均值聚类算法的采动场地稳定性评价[J]. 煤炭科学技术, 2020, 48(10): 137 – 142.
- GUO S, GUO G L, LI H Z, et al. Stability evaluation of mining-induced goaf site based on dimensionality reduction fuzzy C -means clustering algorithm[J]. Coal Science and Technology, 2020, 48(10): 137 – 142.
- [18] 魏文军, 张轩铭, 杨立本. 基于模糊聚类和改进Densenet网络的小样本轴承故障诊断[J]. 哈尔滨工业大学学报, 2024, 56(3): 154 – 163.
- WEI W J, ZHANG X M, YANG L B. Fault diagnosis of small sample bearings based on fuzzy clustering and improved Densenet network[J]. Journal of Harbin Institute of Technology, 2024, 56(3): 154 – 163.
- [19] 岳光华, 滕丽, 周文波, 等. 模糊聚类分析在粗集料分类选择中的应用[J]. 江苏大学学报(自然科学版), 2022, 43(4): 453 – 457.
- YUE G H, TENG L, ZHOU W B, et al. Application of fuzzy clustering analysis in classification and selection of coarse aggregate[J]. Journal of Jiangsu University (Natural Science Edition), 2022, 43(4): 453 – 457.
- [20] 王敏, 嵇绍春. 基于模糊聚类和模糊模式识别的数字图书馆个性化推荐研究[J]. 现代情报, 2016, 36(4): 52 – 56.
- WANG M, JI S C. Design of the digital library personalized recommendation system based on the fuzzy clustering and fuzzy pattern recognition[J]. Journal of Modern Information, 2016, 36(4): 52 – 56.
- [21] 罗眉, 田涛, 方林波, 等. 基于模糊聚类的复杂系统资源调度[J]. 北京工业大学学报, 2013, 39(11): 1688 – 1691, 1699.
- LUO M, TIAN T, FANG L B, et al. Resource scheduling of complex system based on fuzzy clustering[J]. Journal of Beijing University of Technology, 2013, 39(11): 1688 – 1691, 1699.
- [22] JIANG C M, LI Z C, YAO J T. A shadowed set-based three-way clustering ensemble approach[J]. International Journal of Machine Learning and Cybernetics, 2022, 13(9): 2545 – 2558.

- [23] 李顺勇, 曹利娜, 刘坤, 等. 动态视图融合矩阵驱动的多视图矩阵分解聚类算法[J]. 计算机工程与应用, 2025, 61(17): 171–184.
- LI S Y, CAO L N, LIU K, et al. multi-view matrix factorization clustering algorithm based on dynamic view fusion matrix [J]. Computer Engineering and Applications, 2025, 61(17): 171–184.
- [24] 邓军, 杨成, 任立峰, 等. 基于机器学习的煤自燃预测研究进展及展望[J]. 煤炭学报, 2025, 50(增刊 1): 336–360.
- DENG J, YANG C, REN L F, et al. A review on coal spontaneous combustion prediction based on machine learning[J]. Journal of China Coal Society, 2025, 50 (Supl. 1): 336–360.

Division of spontaneous combustion “three zones” in goaf based on fuzzy C -means clustering algorithm

PAN Jingtao¹, QIU Chen¹, ZHAO Dan², LIU Liren³, YUAN Qiang³

(1 School of Mining, Liaoning Technical University, Fuxin 123000, Liaoning, China;

2 College of Safety Science & Engineering, Liaoning Technical University, Huludao 125105, Liaoning, China;

3 Sandaogou Coal Mine of CHN Energy Guoshen Group, Yulin 719407, Shaanxi, China)

Abstract: To address the challenges posed by fuzzy boundaries and the limitations of traditional methods in managing uncertainty when delineating the spontaneous combustion “three zones” in goafs, this study proposes an enhanced division method based on the Fuzzy C -Means (FCM) clustering algorithm. A multi-dimensional evaluation index system that integrates temperature, oxygen concentration, carbon monoxide levels, and burial depth was established to capture the complex environmental characteristics. The study develops a comprehensive mathematical model encompassing data standardization and calibration using Manhattan distance, the construction of fuzzy similarity matrices, and their transformation into equivalence matrices via transitive closure. Utilizing MATLAB, a dynamic clustering mechanism based on λ -cut sets was implemented. Notably, to determine the optimal classification granularity, an adaptive threshold selection method for λ was developed by maximizing the geometric mean of F -statistics across multiple spatial monitoring points. Results indicate that clustering sensitivity exhibits spatial variation: monitoring points near the working face respond sensitively to changes in λ , reflecting high environmental heterogeneity caused by air leakage, while deeper goaf areas demonstrate strong stability. Using the optimal λ , the goaf was distinctly categorized into heat dissipation, oxidation, and suffocation zones. The identified ranges for the oxidation zone are 72–96 m on the intake side and 36–72 m on the return side. In comparison with the on-site “Disaster Comprehensive Intelligent Management Platform”, which is based on single oxygen thresholds (intake: 72–103 m; return: 36–66 m), this method effectively identified a 7 m “sub-oxidation zone” on the intake side where oxidation remains weak despite high oxygen levels, and it extended the risk boundary by 6 m on the return side, revealing a “latent risk area” attributed to thermal inertia. The relative deviations of the boundary endpoints are 6.8% and 9.1%, respectively. This data-driven approach uncovers risk details overlooked by single-threshold methods, providing a more scientific foundation for precise fire prevention.

Key words: safety engineering; fuzzy clustering; goaf; spontaneous combustion “three zones”; data classification