

DOI:10.16031/j.cnki.issn.1003-8035.2020.06.15

基于主成分层次聚类模型的采空塌陷场地稳定性评价

郭松^{1,2}, 郭广礼^{1,2}, 李怀展^{1,2}, 杨向升^{1,2}

(1. 中国矿业大学江苏省资源环境信息工程重点实验室, 江苏 徐州 221116;

2. 中国矿业大学环境与测绘学院, 江苏 徐州 221116)

摘要: 对采空塌陷场地进行稳定性评价是保证后续工程建设安全的重要前提。针对某磷矿急倾斜矿层采空塌陷场地稳定性评价影响因素多、地质采矿条件复杂的问题,提出了基于主成分层次聚类模型的稳定性判别方法。在确定采空塌陷场地稳定性评价范围的基础上,筛选了表征采空塌陷场地稳定性的 8 个主要指标作为学习样本进行训练,经主成分降维后,建立采空塌陷场地稳定性评价的 AGNES (AGglomerative NESTing) 层次聚类模型并应用于该磷矿采空塌陷场地稳定性评价中。实验结果表明,采空塌陷场地稳定性影响因素的前 4 项主成分的累计贡献率为 81.8%,较好地表征原始样本指标所包含的信息,采空塌陷场地稳定性总体能够与此区域城市规划不同用地性质的土地承载力相适应,并与其他手段评价结果相比较证明了主成分层次聚类模型应用于采空塌陷场地稳定性评价中的可行性和有效性。

关键词: 采空塌陷;层次聚类;主成分分析;AGNES 算法;稳定性评价

中图分类号: P642

文献标识码: A

文章编号: 1003-8035(2020)06-0116-06

Goaf-collapse sites stability evaluation based on principal component hierarchical clustering model

GUO Song^{1,2}, GUO Guangli^{1,2}, LI Huaizhan^{1,2}, YANG Xiangsheng^{1,2}

(1. Jiangsu Key Laboratory of Resources and Environment Information Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China; 2. School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China)

Abstract: Stability evaluation of goaf-collapse sites is the primary problem to be solved in the subsequent engineering construction. In this paper, a principal component hierarchical clustering analysis method for a goaf-collapse site stability evaluation has been proposed to solve the problem caused by multiple influencing factors and complicated geological and mining conditions of a steeply pitching phosphate orebody. On the basis of determining the stability evaluation range of the goaf-collapse site, 8 major indicators representing stability of goaf-collapse sites were selected after principal component analysis (PCA) as learning samples for training. AGNES (AGglomerative NESTing) hierarchical clustering analysis model for evaluating stability of goaf-collapse sites was established. After dimensionality reduction, first four principal components of cumulative contribution rate were 81.8%. The results show that the goaf-collapse site can adapt to different land bearing capacity of urban planning in the study area, the discriminant result is consistent with other methods, it indicates that the hierarchical cluster analysis model has a good discriminant ability. The proposed approach

收稿日期: 2020-08-28; 修订日期: 2020-09-25

基金项目: 国家自然科学基金 (51974292); 中央高校基本科研业务费重大项目培育专项 (2018ZDPY05)

第一作者: 郭松 (1990-), 男, 大地测量学与测量工程专业, 博士研究生, 主要从事采动地层灾害分析与稳定性评价方面的研究工作。E-mail: guosong@cumt.edu.cn

通讯作者: 郭广礼 (1965-), 男, 大地测量学与测量工程专业, 博士生导师, 教授, 主要从事开采沉陷与岩层控制等方面的教学与科研。E-mail: glguo1965@163.com

demonstrates the feasibility and effectiveness in the field of stability assessment of goaf-collapse sites.

Keywords: goaf-collapse; hierarchical clustering; principal component analysis; AGNES algorithm; stability evaluation

0 引言

采空塌陷是典型的由地下资源开采引起的地质灾害问题,其产生的根源是地下资源开采改变原岩应力状态后反映在地表的采动损害^[1]。随着城市建设用地日益紧张,采空塌陷场地不仅限制了矿区土地资源利用,也阻碍了城市建设用地的拓展,为处理好地下开采与地面建设的关系,定性分析与定量评价采空塌陷场地稳定性是践行资源环境承载能力和国土空间适宜性评价的重要前提,对实现资源型城市的可持续发展具有重要的现实意义。

采空塌陷场地稳定性主要是以采空区场地与拟建工程相互影响程度为主要判别依据,综合考虑拟采取的抗采动影响技术措施的难易程度及工程造价等因素进行采空塌陷场地稳定性划分^[2]。采空塌陷场地稳定性评价方法主要有:①开采条件判别法:以采空区终采时间为主要因素,结合地表移动变形特征、顶板岩性及松散层厚度等因素进行判别;②地表移动变形判别法:按地表移动变形判别准则对采空区场地稳定性进行评估;③煤(岩)柱稳定性分析法:以采空区内残留煤(岩)柱安全稳定性系数确定场地稳定性等级。实际采空塌陷场地稳定性是地质采矿因素相互影响的时空动态过程,上述采空塌陷场地稳定性评价方法没有明确的综合判别准则,使用难度较高,存在较大的局限性。

目前,围绕采空区场地稳定性评价的研究形成了以理论分析及综合评价为主的分析体系。GUO等^[3]提出了固体充填开采的沉陷控制方法,较可靠地预测了固体充填后地表移动变形情况;谭志祥等^[4]建立了采动区建筑地基、基础和上部结构协同作用的力学模型,分析了不同影响因素下建筑物附加应力的变化规律;张永波^[5]提出了建筑物荷载大小、相对位置等对建筑物地基稳定性的影响及其变形破坏规律。随着地下资源开采规模不断扩大,地表塌陷面积持续增加,采空塌陷场地的开发与治理需求日趋迫切,唐山南湖公园与徐州潘安湖小镇是采空塌陷场地综合治理与规划的典型案例^[6-7],通过将采煤塌陷区整治成生态或建设用地,突破了传统的农业复垦方式,实现矿产资源开发与城市建设协同发展的有效统一^[8]。张曦沐等^[9]针对国外典型资源型城市,如德国鲁尔区、法国洛林地区、英国煤炭城市盖茨黑德、

日本产煤地域等采空塌陷区开展了人居环境建设研究。在采空塌陷场地稳定性综合评价方面,机器学习理论与方法已逐渐得到应用并取得了许多研究成果:聚类分析通过对无标记训练样本的学习揭示数据内在性质及规律,为采空塌陷场地的数据分析提供了基础^[10],聚类方法从 K 均值(k -means)聚类将采空塌陷场地稳定性评价的样本集按最小化平方误差聚类迭代优化来近似得到簇划分,发展到多层次空间聚类、模糊 C 均值聚类(Fuzzy C-means, FCM)等“软聚类”(soft clustering)算法评价体系,允许每个样本以不同程度同时属于多个原型^[11-12]。随着样本数据量大、影响因素多,针对聚类不客观的缺陷,开展以主成分降维层次聚类分析的算法模型研究将成为解决多因素影响下的采空塌陷场地综合评价问题的有效途径。

1 采空塌陷场地地质采矿条件概况

研究区域位于连云港市区北郊(图1)所示。该磷矿共5个可采矿体,均进行了不同程度的开采,采矿方式为地下开采,采用竖井、盲竖井开拓。采空区主要有底板埋深-440 m以上的7P采空区、13P采空区、11P采空区和底板埋深-120 m以上的8P采空区。在受到外力(如拟建建(构)筑物基底扩散压力、地下水位升降或地震等)的影响后,采空塌陷区域内破裂岩体失稳的可能性很大,是诱发采空区“活化”的主要区域。



图1 研究区域卫星影像图

Fig.1 Satellite images of the study area

矿界范围内采空塌陷场地作为拟规划建设场地,地表塌陷的破坏程度不仅有连续移动变形,还存在以突发性、隐藏性为特点的非连续移动变形并形成了较大面积的采空区。开采历史过程中已发生3次较大塌陷事故,第3次地面塌陷发生在-120 m中段,形成长轴约130 m,短轴约100 m的倒立圆台塌陷坑,塌陷面积约9 400 m²。同时,地裂缝发育较明显的区域集中在矿区北侧道路,该段北部裂缝走向147°~162°,南部走向50°~75°,缝宽1~2 cm,沿走向呈羽状展布(图2)。



图2 研究区域采空塌陷现状

Fig. 2 Present situation of goaf-collapse in the study area

2 主成分层次聚类原理与模型检验

2.1 主成分层次聚类的基本原理

层次聚类(Hierarchical Clustering Method)是非监督机器学习聚类算法的一种,是基于簇间相似度在不同层次对数据集进行划分所形成的树形聚类结构^[13]。AGNES (AGglomerative NESting)算法作为自底向上聚合策略的层次聚类聚合算法,其基本原理是:以指标个数为维度的空间数据集,把数据集中的每个数据看作一个簇,采用最短欧式距离^[14]将距离相近的两簇合并为一个新簇;再计算新簇与其他各簇间的距离,继续合并距离最近的两类为一个新类,循环以上过程,直至达到预设的聚类簇个数。对于给定的聚类簇中心 C_i 和 C_j 通过式(1)计算聚类簇之间的最小距离:

$$d_{\min}(C_i, C_j) = \min_{x \in C_i, z \in C_j} \text{dist}(x, z) \quad (1)$$

式中, $\text{dist}(x, z)$ 为两样本之间的距离, $d_{\min}(C_i, C_j)$ 对应于簇 C_i 与簇 C_j 最近的间距。AGNES算法的聚类簇个数即对应的《采空塌陷勘查规范》(以下简称为《规范》)中采空塌陷场地稳定性评价分级:“稳定”“基本稳定”和“不稳定”。

针对采空塌陷场地稳定性评价影响因素多、数据维度高的问题,在进行层次聚类前,引入主成分分析方法对高维度数据进行降维后再参与模型训练,以实现训练效率的提升。主成分分析的基本步骤如下:

(1)假设采空塌陷场地稳定性评价样本数据集为 $D = \{X_1, X_2, \dots, X_n\}$,其中 $X_i = (x_{i1}, x_{i2}, \dots, x_{im})^T$,即数据集中的每个样本均为一个 m 维列向量,第 i 个样本数据 X_i 的第 j 维分量为 x_{ij} ,构造标准化样本值 Z_{ij} :

$$z_{ij} = \frac{x_{ij} - u_j}{s_j} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, m) \quad (2)$$

式中, u_j 和 s_j 分别是样本数据集 D 中第 j 维分量的均值和标准差,即

$$u_j = \sum_{i=1}^n x_{ij}/n, s_j^2 = \sum_{i=1}^n (x_{ij} - u_j)^2/n - 1$$

(2)构造标准化样本矩阵 Z_{ij} 并计算协方差矩阵

$$C = \frac{1}{m} Z^T Z;$$

(3)求出协方差矩阵 C 的全部特征根并按从大到小次序排列,将前 k 个特征值对应的特征向量按行排列构成变换矩阵 W ;

(4)保留变换后的样本数据的前 k 个分量,则得到相应累计方差贡献率 Ω :

$$\Omega = \sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i \quad (3)$$

式中, $\lambda_1, \lambda_2, \dots, \lambda_n$ 为协方差矩阵 C 的特征值由大到小顺序排列。

2.2 聚类要素的模型检验

通过引入外部性能度量指标Rand指数(Rand Index)对构建的主成分层次聚类模型的聚类精度进行检验。Rand指数将模型聚类结果与采用《规范》中的评价方法进行比较来评估其优劣。具体步骤是:对数据集 $D = \{x_1, x_2, \dots, x_m\}$,通过聚类给出的簇划分为 $C = \{C_1, C_2, \dots, C_k\}$,《规范》中采空塌陷场地稳定性评价方法给出的簇划分为 $C^* = \{C_1^*, C_2^*, \dots, C_s^*\}$ 。令 λ 与 λ^* 分别表示与 C 和 C^* 对应的簇标记向量。将样本两两配对,分别定义:

$$a = |SS|, SS = \{(x_i, x_j) \mid \lambda_i = \lambda_j, \lambda_i^* = \lambda_j^*, i < j\} \quad (4)$$

$$b = |SD|, SD = \{(x_i, x_j) \mid \lambda_i = \lambda_j, \lambda_i^* \neq \lambda_j^*, i < j\} \quad (5)$$

$$c = |DS|, DS = \{(x_i, x_j) \mid \lambda_i \neq \lambda_j, \lambda_i^* = \lambda_j^*, i < j\} \quad (6)$$

$$d = |DD|, DD = \{(x_i, x_j) \mid \lambda_i \neq \lambda_j, \lambda_i^* \neq \lambda_j^*, i < j\} \quad (7)$$

其中,集合 SS 表示包含了在 C 中隶属于相同簇且在 C^* 中也隶属于相同簇的样本对,集合 SD 包含了在 C 中隶属于相同簇但在 C^* 中隶属于不同簇的样本对,由

于每个样本对 $(x_i, x_j) (i < j)$ 仅能出现在一个集合中, 因此有 $a + b + c + d = m(m - 1)/2$ 成立, 评价聚类性能度量的外部指标 RI 即可表示为:

$$RI = \frac{2(a + d)}{m(m - 1)} \tag{8}$$

显然, RI 的取值范围为 $[0, 1]$, 其值越大表征聚类分析的准确率越高。

3 主成分层次聚类模型的效果分析

3.1 急倾斜矿层采空塌陷影响范围的确定

根据拟建设场地初步规划, 采空塌陷范围内的用地性质主要有公园绿地, 城市道路用地、商业用地等。结合拟建建筑结构特点与采空塌陷影响宽度基岩移动影响角的取值, 综合考虑研究区域基岩移动角为 60° , 松散层移动角为 45° , 对该急倾斜矿层采空塌陷场地圈定稳定性评价范围(图 3)。评估范围内涉及的主要采空区有 7P、8P、11P 和 13P 采空塌陷区。

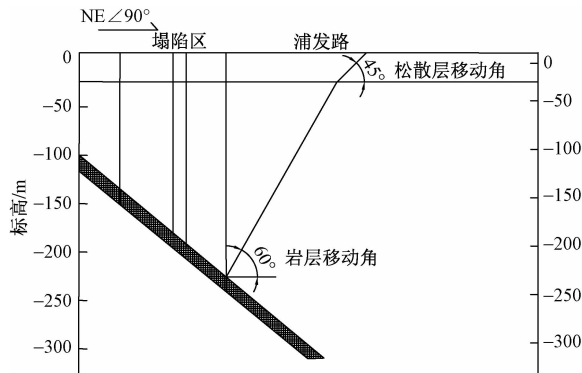


图 3 岩层移动角和松散层移动角推测示意图
Fig.3 Schematic diagram of strata movement and soil layer angle

3.2 采空塌陷影响因素的降维层次聚类

采空塌陷场地的覆岩破坏和地表塌陷主要受开采深度、采厚、矿层倾角、采矿方式、地质构造、覆岩物理力学性质及顶板管理方式等因素综合影响。对于不同地质采矿条件下的采空塌陷场地稳定性评价, 通过上述因素的不同组合作为定性定量评判的主要依据。

结合研究区域的覆岩破坏和地表塌陷特征, 选取采空塌陷场地稳定性的 8 个主要因素进行主成分层次聚类分析建模, 分别为: X_1 采厚, m ; X_2 采深, m ; X_3 围护带宽度, m ; X_4 走向工作面面长, m ; X_5 倾向工作面面长, m ; X_6 矿层倾角($^\circ$); X_7 停采年限, a ; 和 X_8 顶板管理方法。评估范围内的采空塌陷场地稳定性的主要影响因素如表 1 所示。

表 1 层次聚类的主要影响因素

Table 1 Main influencing factors of hierarchical clustering

编号	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
7P	7.28	440	20	711	813	55	4.5	1
8P	3.10	120	20	200	170	56	3.8	1
11P	2.42	440	20	430	623	58	5.2	1
13P	3.49	440	20	288	558	54	7.1	1

对初始变量进行极差标准化处理, 采用最短欧氏距离定义样本数据之间的距离, 标准化后的样本数据如表 2 所示。

表 2 标准化后的初始变量

Table 2 Standardized initial variables

编号	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
7P	1.696 2	0.577 4	0	1.568 8	1.163 4	0	1.617 1	-1.414 2
8P	-0.514 3	-1.732 1	0	-1.070 4	-1.586 9	0	-0.420 9	0
11P	-0.873 9	0.577 4	0	0.117 5	0.350 7	0	-1.094 3	0
13P	-0.308 0	0.577 4	0	-0.615 9	0.072 7	0	-0.101 9	1.414 2

综合考虑主成分分析对数据空间维数与损失信息的相关性, 经主成分分析降维后的各因素贡献率如图 4 所示。前 4 项主成分的累计贡献率为 81.8%。

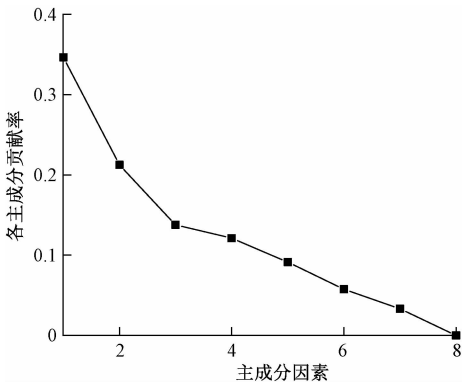


图 4 各主成分贡献率

Fig.4 Contribution rate of principal components

以前四项主成分作为输入参数建立 AGNES 层次聚类的采空塌陷场地稳定性判别模型, 图 5 所示即为采空塌陷场地稳定性的聚类过程。

在聚类图的特定层次上进行分割, 得到相应的簇划分结果, 结合采空塌陷场地实际稳定性情况, 经主成分层次聚类后的采空塌陷场地稳定性分级如表 3 所示。

表 3 采空塌陷场地稳定性分级

Table 3 Stability classification of goaf-collapse sites

稳定性类型	场地编号
稳定	11P/13P
基本稳定	8P
不稳定	7P

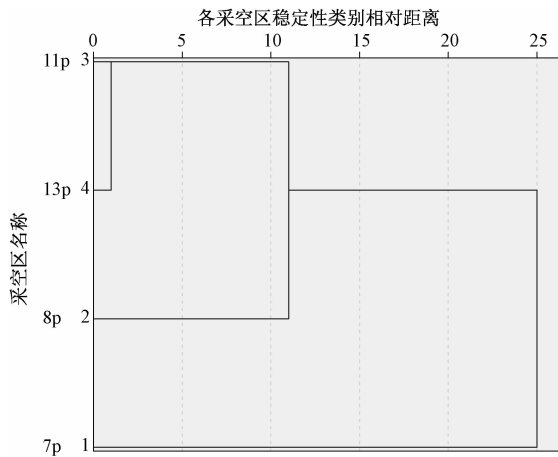


图5 采空塌陷场地稳定性的聚类过程

Fig.5 Clustering process of stability in goaf-collapse site

4 稳定性评价结果与区划分析

经主成分层次聚类后,研究区域4个主要采空塌陷区域稳定性等级分属3类,将拟规划建设采空塌陷场地的稳定性分为A区(不稳定)、B区(基本稳定)、C区(稳定)(图6)。



图6 采空塌陷场地稳定性区划示意图

Fig.6 Diagram of site stability zoning of goaf-collapse site

根据评估范围内的采空塌陷特征,采用开采条件判别法对采空塌陷场地稳定性进行评价,以终采时间为主控因素,综合考虑地表变形特征、采深、顶板岩性

和松散层厚度等因素,该场地稳定性综合评价结果(表4)。

表4 按开采条件判别法评价采空塌陷场地稳定性

Table 4 Stability evaluation based on discrimination method of mining conditions in goaf-collapse site

采空区编号	采厚/m	采深/m	停采时间/a	场地稳定性评价
7p	7.28	440	4.5	不稳定
8p	3.10	120	3.8	基本稳定
11p	2.42	440	5.2	基本稳定
13p	3.49	440	7.1	稳定

对比分析主成分层次聚类模型与开采条件判别法的评价结果,聚类性能度的外部指标 Rand 指数为 0.75。研究区域场地稳定性总体呈现中部稳定性差,周边区域稳定性好的情况。其中,采空塌陷场地的中部 7P 和 8P 采空塌陷场地为不稳定区域,其场地仅能满足一般城市绿地的用地需求;11P 和 13P 采空区为稳定与基本稳定区,其特点是地表沉陷变形小,具有建设使用价值和经济价值。总体上,采空塌陷场地能够与研究区域城市规划的不同用地性质的土地承载力相适应,主成分层次聚类模型判别场地稳定性分区符合实际情况。

5 结论

(1)主成分层次聚类模型对研究区域采空塌陷场地稳定性的判别结果与采用开采条件综合判别方法的场地稳定性结果基本一致,模型针对急倾斜矿层消除了各地质采矿影响因素间的相关性,有较强的判别能力。

(2)主成分分析对高维特征数据集进行降维,前4项主成分包含了原始数据序列的主要特征,为后续层次聚类算法提供了数据基础,提高了计算效率。层次聚类算法以最短欧式距离作为度量簇间的链接准则,模型检验的 Rand 指数为 0.75,聚类谱系图包含了整个算法过程的信息,结果直观。

(3)通过对比采空塌陷场地稳定性评价结果验证了主成分层次聚类判别方法能够处理非线性高维度数据,为地下开采覆岩破坏与地表塌陷场地时空演化机理研究提供一种新途径,在采空塌陷场地稳定性评价缺少判别准则时提供了理论支撑。

参考文献:

- [1] 李怀展,查剑锋,元亚菲. 关闭煤矿诱发灾害的研究现状及展望[J]. 煤矿安全, 2015, 46(5): 201-204. [LI H Z, ZHA J F, YUAN Y F. Research status and prospect of hazards caused by close coal

- mine [J]. Safety in Coal Mines, 2015, 46 (5): 201 - 204. (in Chinese)]
- [2] 国家安全监管总局. 建筑物、水体、铁路及主要井巷煤柱留设与压煤开采规范[S]. 北京:煤炭工业出版社, 2017. [State Administration of Work Safety. Code for coal pillar leaving and coal mining under building, water body, railway and main shaft [S]. Beijing: China Coal Industry Press, 2017. (in Chinese)]
- [3] GUO G L, ZHU X J, ZHA J F, et al. Subsidence prediction method based on equivalent mining height theory for solid backfilling mining[J]. Transactions of Nonferrous Metals Society of China, 2014, 24(10): 3302 - 3308.
- [4] 谭志祥, 邓喀中. 采动区建筑物地基、基础和结构协同作用模型[J]. 中国矿业大学学报, 2004, 33(3): 264 - 267. [TAN Z X, DENG K Z. Coordinating work model of ground, foundation and structure of building in mining area[J]. Journal of China University of Mining & Technology, 2004, 33(3): 264 - 267. (in Chinese)]
- [5] 张永波. 老采空区建筑地基稳定性及其变形破坏规律的研究[D]. 太原: 太原理工大学, 2005. [ZHANG Y B. Research on the stability of building foundation above old mine goafs and its damage regular [D]. Taiyuan: Taiyuan University of Technology, 2005. (in Chinese)]
- [6] 沈瑾, 赵铁政. 棕地与绿色空间网络——唐山南湖采煤沉陷区空间再利用[J]. 建筑学报, 2006(8): 28 - 30. [SHEN J, ZHAO T Z. Brown land and green space network: spacial reutilization of the Nanhu coal mining settlement area in Tangshan[J]. Architectural Journal, 2006(8): 28 - 30. (in Chinese)]
- [7] 廖湛娅, 吴克宁. 矿区农田景观生态适宜性评价——以徐州市潘安采煤塌陷区为例[J]. 资源与产业, 2012, 14(3): 36 - 42. [LIAO C H, WU K N. A case study on xuzhou's pan'an coal mine subsidence: assessment of farmland landscape ecological suitability in mining area[J]. Resources & Industries, 2012, 14(3): 36 - 42. (in Chinese)]
- [8] 熊彩霞, 梁恒昌, 马金荣, 等. 煤矿采空区建筑场地地基适宜性分析[J]. 采矿与安全工程学报, 2010, 27(1): 100 - 105. [XIONG C X, LIANG H C, MA J R, et al. Stability analysis of the building foundation over goaf area[J]. Journal of Mining and Safety Engineering, 2010, 27(1): 100 - 105. (in Chinese)]
- [9] 张曦沐, 张国锋, 马靖华. 关于采煤沉陷区人居环境建设的思考[J]. 建筑科学, 2010, 26(11): 103 - 105. [ZHANG X M, ZHANG G F, MA J H. Thinking of settlement environment construction in mining subsidence area[J]. Building Science, 2010, 26(11): 103 - 105. (in Chinese)]
- [10] 周翠竹, 朱建军, 石岩. 一种基于双重距离约束的多层次空间聚类方法[J]. 测绘科学, 2014, 39(10): 98 - 101. [ZHOU C Z, ZHU J J, SHI Y. A multi-level spatial clustering method based on dual distance constraints [J]. Science of Surveying and Mapping, 2014, 39(10): 98 - 101. (in Chinese)]
- [11] 李光强, 邓敏, 程涛, 等. 一种基于双重距离的空间聚类方法[J]. 测绘学报, 2008, 37(4): 482 - 488. [LI G Q, DENG M, CHENG T, et al. A dual distance based spatial clustering method [J]. Acta Geodaetica et Cartographica Sinica, 2008, 37(4): 482 - 488. (in Chinese)]
- [12] 宫凤强, 刘科伟, 李志国. 矿区采空塌陷危险性预测的 Bayes 判别分析法[J]. 采矿与安全工程学报, 2010, 27(1): 30 - 34. [GONG F Q, LIU K W, LI Z G. The Bayes discriminant method for forecasting the stability of underground goaf[J]. Journal of Mining and Safety Engineering, 2010, 27(1): 30 - 34. (in Chinese)]
- [13] 孙云华, 郭涛, 崔希民, 等. 基于行为聚类算法的土地利用聚类适宜性分析研究[J]. 地球信息科学学报, 2016, 18(3): 396 - 405. [SUN Y H, GUO T, CUI X M, et al. Suitability analysis on behavior-based aggregation of land use classification in Yunnan Province [J]. Journal of Geo-information Science, 2016, 18(3): 396 - 405. (in Chinese)]
- [14] CUI X M, ZHAO Y L, WANG G R, et al. Calculation of residual surface subsidence above abandoned longwall coal mining [J]. Sustainability, 2020, 12(4): 1528.