

刘春国,余丹,陶志刚,等.基于数据挖掘的甘肃武都殿沟泉氡异常与地震活动统计关系研究[J].地震工程学报,2021,43(6):1263-1268.DOI:10.3969/j.issn.1000-0844.2021.06.1263

LIU Chunguo, YU Dan, TAO Zhigang, et al. Statistical relationship between the radon anomaly of Diangou spring and seismic activity in Wudu, Gansu based on data mining technology[J]. China Earthquake Engineering Journal, 2021, 43(6): 1263-1268. DOI:10.3969/j.issn.1000-0844.2021.06.1263

基于数据挖掘的甘肃武都殿沟泉氡异常 与地震活动统计关系研究

刘春国,余丹,陶志刚,樊俊屹

(中国地震台网中心,北京 100045)

摘要: 殿沟泉气氡浓度在汶川地震后经常出现成簇高频震荡异常。应用数据挖掘方法,采用正规化、聚类、特征值计算等方法对氡观测数据和地震目录进行预处理,提取气氡异常、震级、震源深度和地震分区等参数,在此基础上采用贝叶斯网络模型开展殿沟泉氡异常与周边区域地震活动统计关系研究。利用聚类开展的地震分区结果显示与构造分区基本吻合,反映出不同构造单元地震活动的丛集性特征;贝叶斯网络模型结果表明泉点气氡异常与周边地震活动有关,特别是与汶川地震主震区龙门山断裂中段和临近区青川断裂区域的地震活动有关。研究结果有助于了解泉点气氡对周边地震响应的差异性特征,对未来该泉点气氡异常的判断与地震预测有参考意义。为地震前兆异常分析和地震活动分区提供了一种新的技术思路和方法。

关键词: 氡浓度异常;地震活动;数据挖掘;聚类分区

中图分类号: P315.723

文献标志码: A

文章编号: 1000-0844(2021)06-1263-06

DOI:10.3969/j.issn.1000-0844.2021.06.1263

Statistical relationship between the radon anomaly of Diangou spring and seismic activity in Wudu, Gansu based on data mining technology

LIU Chunguo, YU Dan, TAO Zhigang, FAN Junyi

(China Earthquake Networks Center, Beijing 100045, China)

Abstract: High-frequency oscillatory anomalies of radon concentration in Diangou spring in Wudu City, Gansu Province often clustered after the Wenchuan earthquake. The radon observation data and seismic catalogue were preprocessed by using the data mining method, then the parameters such as radon anomaly, magnitude, focal depth, and seismic zonation were extracted. On this basis, Bayesian network model was adopted to study the statistical relationship between radon anomaly in Diangou spring and seismic activity in the surrounding area. The results of seismic

收稿日期: 2020-09-20

基金项目: 中国地震局地震科技星火计划攻关项目(XH20002);中国地震局专项“全国地球物理台网数据跟踪分析与产出”

第一作者简介: 刘春国(1968—),女,湖南涟源人,硕士,正高级工程师,主要研究方向为地下流体数据处理与观测方法。

E-mail: liuchguo@126.com.

zonation obtained from the clustering method were basically consistent with those of tectonic division, reflecting the clustering characteristics of seismic activities in different structural units. The results of Bayesian network model showed that the gas radon anomalies were related to the surrounding seismicity, especially the seismic activities in the middle segment of Longmenshan fault and the Qingchuan fault. The research results are helpful to understand the differential characteristics of gas radon response in Diangou spring to surrounding earthquakes, thus having reference significance for the judgment of gas radon anomaly and earthquake prediction in the future. It also provides a new technical idea and method for earthquake precursor anomaly analysis and seismic activity zonation.

Keywords: radon concentration anomaly; seismic activity; data mining; cluster zoning

0 引言

氡是目前国际上普遍认可的七大地震前兆之一。长期的观测结果表明氡存在显著的震前异常变化。2008年5月12日发生的汶川M8.0地震引起了中国大陆多个井泉的氡浓度出现脉冲式、台阶式的异常变化,氡的震后效应才开始引起人们的关注和研究^[1-2]。

殿沟泉位于甘肃武都县城西约3 km处的殿沟村,距汶川地震8.0级震中302 km。泉水出露于NEE向的断层破碎带上志留系的灰岩中,属于裂隙岩溶上升泉。水温年变化范围为17~18℃,流量为0.7~1.4 L/s。自2001年3月27日开始该泉采用SD-3A测氡仪,开始了连续自动观测气氡浓度,其采样率为1次/h。氡浓度年动态为夏高冬低型,汶川地震气氡同震效应明显,同时流量变化很大^[3-4]。汶川地震后经常出现成簇高频震荡异常变化,经现场检查与台站调研,观测系统运行正常,周边观测环境无明显变化,仪器观测的气氡浓度异常变化真实可信。

汶川地震发震断层为NE向龙门山断裂带,与殿沟泉构造上同属于我国南北带地质构造中段。汶川地震引发的数以万计的余震空间展布尺度大,殿沟泉正好位于汶川地震余震分布的东北末端。殿沟泉出现的高频震荡异常出现的密集时段正好是汶川地震余震密集发生时段。

殿沟泉氡异常是否因汶川地震余震引起?泉点气氡对不同区域的地震活动响应是否存在差异性?为此,我们拟采用数据挖掘方法开展殿沟泉氡异常与地震活动统计关系研究。

数据挖掘技术已广泛应用在工业、农业、商业、医学、金融、服务及气象等多个领域,而在地震监测预报领域尚处于起步阶段,主要研究集中在地震目录及地震活动性等方面,而用于前兆异常研究尚未

见报道,难点在于地震监测数据为高维时间序列,其预处理方法是数据挖掘的关键技术。

研究思路是:首先对氡观测数据和地震目录数据进行预处理。预处理内容包括数据的清洗、数据转换与分箱处理、异常特征值的提取、地震目录聚类与地震空间参数的获取等。然后对异常特征值、地震震级、地震空间参数等集成、统计分析研究异常与地震活动分区的关系。

1 预处理与集成

1.1 数据来源

殿沟泉气氡数据和地震目录数据均来自于地震科学数据共享网站(<https://data.earthquake.cn>)。时间从2008年1月1日至2010年12月30日。殿沟泉气氡为小时值时间序列数据,共计26 304个数据。本文采用中国大陆地震速报地震目录,共计116 263条记录(含汶川地震余震)。

1.2 氡浓度观测数据预处理

气氡观测数据的预处理包括观测数据清洗和异常特征值的提取。气氡观测数据的清洗主要包括错误数据剔除和缺数标志的统一。气氡异常特征值用变量 σ_3 表示,表征气氡观测数据是否异常, σ_3 为标志变量,取值1(异常)和0(无异常)。气氡异常参数 σ_3 提取方法如下:

采用移动平均法消除气氡时间序列 $\{X_i\}$ 的趋势变化,得到消除趋势的气氡残差序列 $\{\Delta X_i\}$ ($i=0,1,2,\dots,n$),选取相对较为平静时段:2009年1月1日—12月31日,计算该段残差序列的均值 $\overline{\Delta x}$ 和标准差 δ 。当 $\{\Delta X_i\}$ 中的 ΔX_i 满足条件: $\Delta X_i > \overline{\Delta x} + 3 \times \delta$ 或 $\Delta X_i < \overline{\Delta x} - 3 \times \delta$ 时,认为 ΔX_i 对应的观测值 X_i 出现显著异常, σ_3 为1,否则为0。

气氡观测时间序列与残差序列曲线如图1所示。

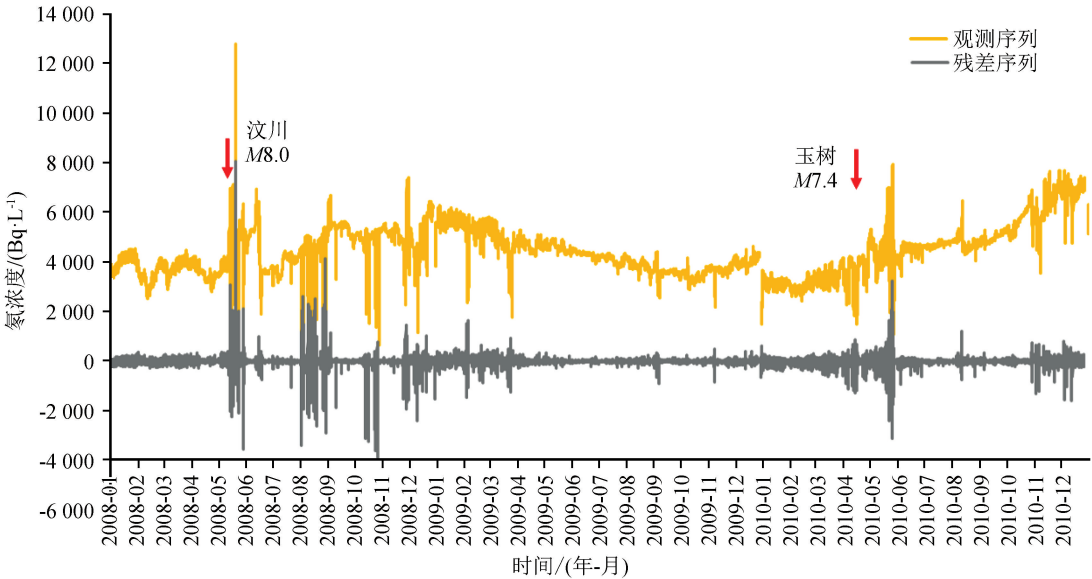


图 1 气氡小时值观测序列与残差序列对比曲线

Fig.1 Comparison between curves of radon observation series and residual series

1.3 地震目录预处理

地震目录预处理包括地震目录的筛选、震级的正规化和地震空间参数提取。

(1) 地震目录筛选

根据地震能量的衰减以及气氡浓度对地震响应的经验范围,按 3 级以下地震影响范围 50 km 以内,3.0~3.9 级影响范围 100 km 内,4.0~4.9 级影响范围 300 km 内,5.0~5.9 级影响范围 600 km 内,6.0~6.9 级影响范围 1 000 km 内,大于 7 级为全国范围内,选取 2008 年 1 月 1 日至 2010 年 12 月 30 日期间可能与殿沟泉气氡变化相关的地震目录 1 124 条。

(2) 震级与震源深度处理

按照表 1,对选取的地震目录按震级划分为 9 个区段,分别投射到整数 0、1、2、3、4、5、6、7、8,并把这个整数称为类震级。类似地,对地震深度也进行了处理,划分 4 个区段分别投射到整数 1、2、3、4,把这个整数称为类深度。

表 1 类震级与地震震级、类深度与地震深度

Table 1 Regular magnitude as well as earthquake magnitude and regular depth as well as earthquake depth			
类震级	地震震级范围	类深度	地震震源深度/km
0	<3.0	1	[1,10)
1	[3.0,4.0)	2	[10,20)
2	[4.0,5.0)	3	[20,30]
3	[5.0,5.5)	4	>30
4	[5.5,6.0)		
5	[6.0,6.5)		
6	[6.5,7.0)		
7	[7.0,7.5)		
8	[7.5,8.0]		

(3) 地震空间参数提取

考虑到地震呈现条带状分布和沿地震断裂带分布的特点,地震空间分布参数没有采取将整个地震分布区域划分为若干区域并编号的方式来提取^[5],而是采用 K-平均聚类方法(K-Means),对 1 224 个地震目录样本数据的地震经度、纬度进行聚类。地震空间参数即为地震所在区域的聚类编号。

K-平均聚类算法:首先随机地选择 k 个对象,每个对象初始代表了一个簇的平均值或中心。对剩余的每个对象根据其与各个簇中心的距离,将被赋给最近的簇,然后重新计算每个簇的平均值。这个过程不断重复,直至准则函数收敛^[6]。K-平均聚类算法的准则函数定义为:

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - \overline{x_i}\|$$

(1)

式中: x 是空间中的点,表示给定的数据对象; x_i 是簇 C_i 的平均值。这个准则可以保证生成的簇尽可能的紧凑和独立。

第一次聚类结果分出 10 个聚类区域。除聚类 J_1 区和 J_7 区之外,其他聚类区域地震个数较少,均小于 5 个,远离研究泉点区域。 J_1 区紧邻研究泉点,包含的地震最多,又是本次研究的重点区域,对该区域进行第二次聚类,又分出 J_{1-1} 、 J_{1-2} 、 J_{1-3} 三个聚类区域。两次聚类分区结果如表 2 所列,可见聚类分区基本与构造分区吻合,反映了不同构造单元地震活动的丛集性特征。

特别是地震数目较多的 J_7 、 J_{1-1} 、 J_{1-2} 和 J_{1-3} 区,这也是汶川地震及其余震分布密集区域,聚类分区结

果与依据长度约 60 km 的余震稀疏段为标志划分的基本段落一致^[7]。J₇ 与 J₁ (J₁₋₂) 的分区界线在北川—江油一线,而这正好是龙门山断裂中段与北段

的分界线,也是第四纪活动强弱的分界线^[8-9],J₁₋₃ 与 J₁₋₁ 分界正好是龙门山与西秦岭构造单元的分界。泉点紧邻区域聚类分区地震分布如图 2 所示。

表 2 地震聚类分区结果一览表

Table 2 The clustering results of earthquakes in the study area			构造位置
聚类分区编号	地震数目		
J ₁	J ₁₋₁	118	文县断裂的北侧、西秦岭北缘断裂南侧
	J ₁₋₂	133	映秀—北川断裂中央北西侧
	J ₁₋₃	455	平武—青川断裂的北段青川断裂
J ₇		501	映秀—北川断裂中央
J ₂		1	南天山与昆仑山交汇的地质构造带上,卡兹克阿尔特断裂
J ₃		4	近西南-东北排列:大盈江断裂、磨盘山—绿汁江断裂、南华—楚雄—建水断裂、华蓥山深断裂西南端延伸
J ₅		3	祁连地震带的南边界断裂:大柴旦—宗务隆山断裂带
J ₆		1	于田、策勒与西藏交界的昆仑山分水岭、郭扎错断裂
J ₈		3	甘孜—玉树断裂
J ₉		3	西秦岭北缘断裂及其延伸以北地区
J ₁₀		1	祁连山地震带中东段地区的大黄山隆起东缘

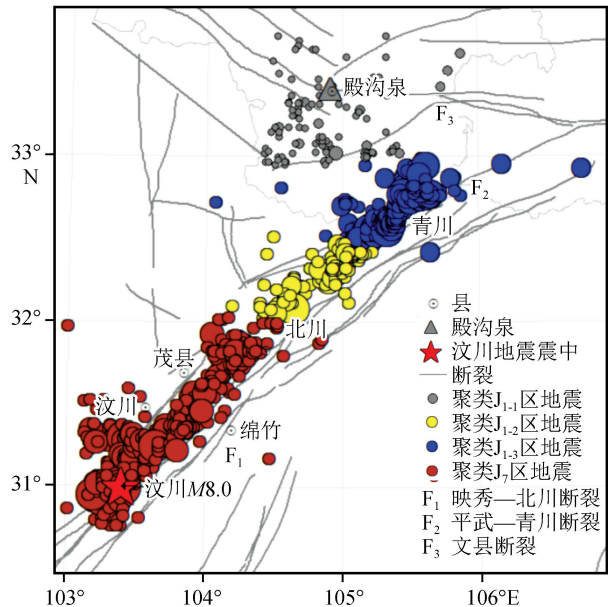


图 2 泉点紧邻区域分区地震分布图

Fig.2 Distribution of earthquakes in the surrounding areas of Diangou spring

1.4 挖掘数据集的形成

经过预处理的气氡观测数据和地震目录对比分析发现,σ₃=1 的时段均出现有地震,而出现地震,气氡 σ₃=1 并不一定等于 1。以地震目录为基准,将地震目录与气氡异常信息集成作为数据挖掘集。数据集包括类震级、类深度、地震空间参数(聚类区域编号)、气氡异常指标 σ₃。

2 氡异常与地震活动关系分析

采用回归、主成份分析法作初步分析发现,气氡异常与周边地震三要素的关系不是一种简单的线性

关系,也不能用一种显式的关系式来表达。本文选用贝叶斯网络模型 (Bayesian Belief Networks, BBN) 从统计概率的角度进行分析。

采用 Clementine12.0 软件进行统计分析。贝叶斯网络模型共设置 3 个属性变量(类震级、类深度、聚类区类别)作为输入节点,1 个标志变量(气氡异常指标 σ₃)作为输出节点。采用 TAN 贝叶斯网络模型,参数学习方法采用最大似然法。

统计分析结果显示,气氡显著异常与震级、地震发生的区域和震源深度均有关关系,引起殿沟泉气氡数据出现显著性异常的重要因素是地震所在的特定构造区域,其次是地震震级和地震深度。

不同聚类区域的概率分布如表 3 所列。聚类区域 J₂、J₃、J₅、J₆、J₉ 及 J₁₀ 区发生的地震,殿沟泉气氡均未出现显著异常;聚类区域 J₄ 和 J₈ 区地震少,均有 1 次 7 级以上地震,对应这些大震各出现 1 次显著异常。J₁ 和 J₇ 区区域发生的地震,殿沟泉气氡出现的异常次数最多(表 3),其中以 J₇ 区最多,其次为 J₁₋₃、J₁₋₂ 区。当地震震级小于 5.5 级时,气氡异常与否的概率基本上相当的,而当震级大于 5.5 级以后,气氡异常的概率明显比无异常概率要大。

不同的聚类区域引起氡浓度异常的地震的能量是否存在明显差异? 我们又计算了考虑地震震级 (M) 和泉与震中距离(r) 的地震能量密度(e)。其计算公式^[10] 为:

lg(e)=(0.48 * M - lg(r) - 1.4)/0.33 (2)

以发震时间为横轴,分别绘制 J₇、J₁₋₁、J₁₋₂、J₁₋₃ 地震能量密度散点图(图 3),用颜色标识氡浓度异常

与否。从图 3 可以看出汶川地震余震一直延续到 2010 年 12 月,其中在大震后一年内尤为密集。引起氡异常的地震集中分布在大震及震后 3 个月内。为了能够看到密集时段引起氡异常的地震情况,又

以地震序号为横轴,分别绘制了 J_7 、 J_{1-1} 、 J_{1-2} 、 J_{1-3} 区地震能量密度散点图(图 4)。从图 4 可以清晰地看到 J_7 和 J_{1-3} 区对应氡异常的地震能量密度较小,约 10^{-6} J/m³,而 J_2 区明显偏高,约为 10^{-5} J/m³。

表 3 不同聚类区域的概率分布表

Table 3 Probability distribution table of different clustering zones								
震级	地震个数	氡异常个数	不同区域氡异常出现概率					
			J_{1-1}	J_{1-2}	J_{1-3}	J_7	J_4	J_8
[3.0,4.0)	353	41	0.05	0	0.95	0	0	0
[4.0,5.0)	658	206	0	0.13	0.14	0.74	0	0
[5.0,5.5)	66	27	0	0.15	0.07	0.78	0	0
[5.5,6.0)	6	2	0	0	0	1	0	0
[6.0,6.5)	12	3	0	0.33	0	0.67	0	0
[6.5,7.0)	1	1	0	0	0	0	0	1
[7.0,7.5)	2	2	0	0	0	0	0.5	0.5
[7.5,8.0]	1	1	0	0	0	1	0	0
地震总个数			20	132	448	495	1	3
氡异常个数合计			2	31	69	178	1	2

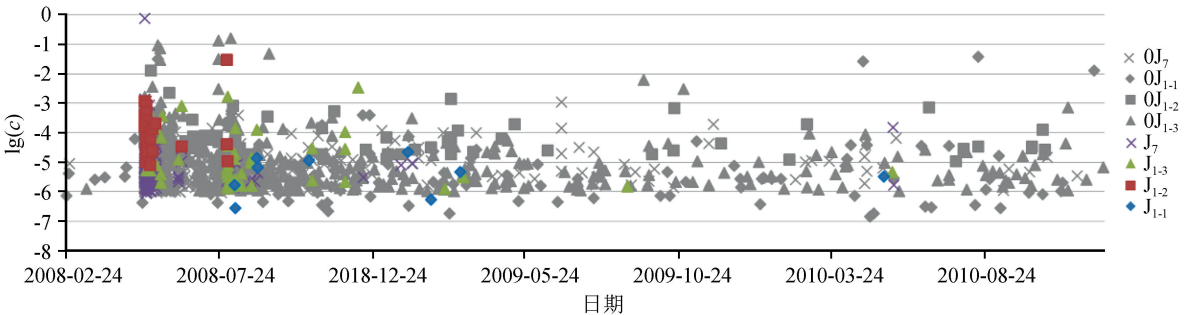


图 3 J_7 、 J_{1-1} 、 J_{1-2} 、 J_{1-3} 聚类区域发震时间与地震能量密度的散点图

Fig.3 Scatter plot of occurrence time and seismic energy density in clustering areas J_7 , J_{1-1} , J_{1-2} and J_{1-3}

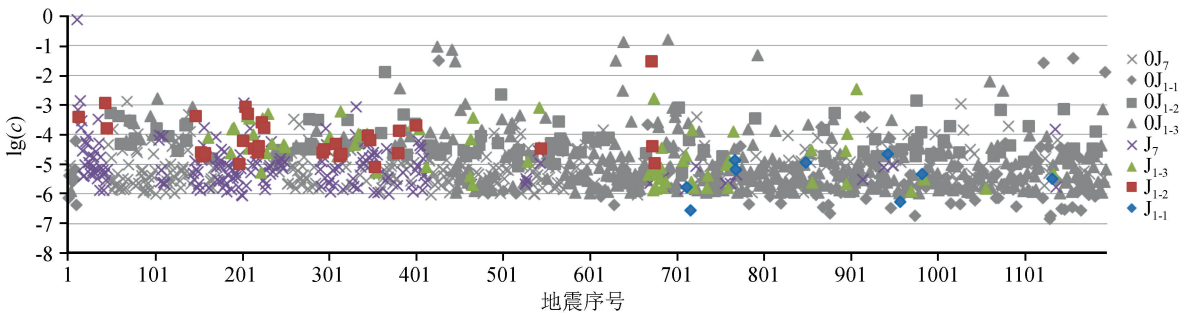


图 4 J_7 、 J_{1-1} 、 J_{1-2} 、 J_{1-3} 聚类区域发震序号与地震能量密度的散点图

Fig.4 Scatter plot of seismic occurrence number and seismic energy density in clustering areas J_7 , J_{1-1} , J_{1-2} and J_{1-3}

3 结论

本文引入聚类、贝叶斯网络模型等数据挖掘方法,对汶川地震后殿沟泉氡出现异常与周边地震的发震时间、震级与构造位置的统计关系进行了研究。获得了如下认识:

(1) 2008—2010 年殿沟泉气氡的多次异常显

著变化是由泉点区周边地区地震活动引起,主要是龙门山断裂中段(汶川地震主震区)和北段(青川断裂)汶川地震及其余震活动引起。当地震震级大于 5.5 级,气氡出现异常的概率明显增大。

(2) 氡值异常受地震所在构造单元的影响较为显著,其次是地震震级、地震深度。龙门山断裂中段和北段地震活动更易引起殿沟泉气氡的异常。

(3) 气氡的地震响应异常机理研究目前还处于探索阶段。考虑到氡异常源在地下水中的迁移遵循水动力弥散机制^[11-12], 气氡地震响应异常形成可能是一种水动力学机制^[13-15], 与地震瞬时流量变化有关。

本项研究结果有助于了解泉点气氡对周边地震响应的差异性特征, 对未来殿沟泉气氡异常的判断与分析预测有参考意义。本项研究是数据挖掘方法应用于前兆异常分析研究的一次尝试, 其数据处理方法及流程具有借鉴意义。此外, 利用聚类开展的地震分区结果令人满意, 为地震活动分区提供了一种新的思路。

参考文献(References)

- [1] 刘耀炜, 任宏微. 汶川 8.0 级地震氡观测值震后效应特征初步分析[J]. 地震, 2009, 29(1): 121-131.
LIU Yaowei, REN Hongwei. Preliminary analysis of the characteristics of post-seismic effect of radon after the Wenchuan 8.0 earthquake[J]. Earthquake, 2009, 29(1): 121-131.
- [2] 任宏微, 刘耀炜, 马玉川. 汶川 8.0 级地震氡震后效应机理讨论[J]. 中国地震, 2010, 26(1): 73-83.
REN Hongwei, LIU Yaowei, MA Yuchuan. Discussion on mechanism of the post-seismic effect of radon in the Wenchuan Ms8.0 earthquake[J]. Earthquake Research in China, 2010, 26(1): 73-83.
- [3] 张昱, 刘宝勤, 徐辉, 等. 武都殿沟泉流体观测资料的分析[J]. 高原地震, 2011, 23(2): 1-6.
ZHANG Yu, LIU Baoqin, XU Hui, et al. Analysis and application of fluid observational data in the Diangou spring in wudu[J]. Plateau Earthquake Research, 2011, 23(2): 1-6.
- [4] YE Q, SINGH R P, HE A H, et al. Characteristic behavior of water radon associated with Wenchuan and Lushan earthquakes along Longmenshan fault[J]. Radiation Measurements, 2015, 76: 44-53.
- [5] 吴绍春, 吴耿锋, 王伟, 等. 寻找地震相关地区的时间序列相似性匹配算法[J]. 软件学报, 2006, 17(2): 185-192.
WU Shaochun, WU Gengfeng, WANG Wei, et al. A time-sequence similarity matching algorithm for seismological relevant zones[J]. Journal of Software, 2006, 17(2): 185-192.
- [6] 李爱国, 匡向阳. 数据挖掘原理、算法及应用[M]. 西安: 西安电子科技大学出版社, 2012.
LI Aiguo, SHE Xiangyang. Principle, algorithm and application of data mining[M]. Xi'an: Xidian University Press, 2012.

- [7] 徐锡伟. 5·12 汶川 8.0 级地震地表破裂图集[M]. 北京: 地震出版社, 2009.
- [8] 邓起东, 陈社发, 赵小麟. 龙门山及其邻区的构造和地震活动及动力学[J]. 地震地质, 1994, 16(4): 389-403.
DENG Qidong, CHEN Shefa, ZHAO Xiaolin. Tectonics, seismismity and dynamics of Longmenshan Mountains and its adjacent regions[J]. Seismology and Geology, 1994, 16(4): 389-403.
- [9] 陈国光, 计凤桔, 周荣军, 等. 龙门山断裂带晚第四纪活动性分段的初步研究[J]. 地震地质, 2007, 29(3): 657-673.
CHEN Guoguang, JI Fengju, ZHOU Rongjun, et al. Primary research of activity segmentation of Longmenshan fault zone since Late-quaternary[J]. Seismology and Geology, 2007, 29(3): 657-673.
- [10] WANG C Y. Liquefaction beyond the near field[J]. Seismological Research Letters, 2007, 78(5): 512-517.
- [11] 刘春国. 水动力弥散与水化学前兆信息传递特征初步研究[J]. 地震地质, 1997, 19(4): 358-362.
LIU Chunguo. Hydrodynamic dispersion and transmission characteristic of hydrogeochemistry precursor information[J]. Seismology and Geology, 1997, 19(4): 358-362.
- [12] 王康, 姚玉霞, 李松林, 等. 气温气压与氡浓度短期变化的相关性分析[J]. 华南地震, 2018, 38(3): 91-98.
WANG Kang, YAO Yuxia, LI Songlin, et al. Correlativity between short-term change of radon concentration and air temperature & air pressure[J]. South China Journal of Seismology, 2018, 38(3): 91-98.
- [13] 车用太, 鱼金子, 刘五洲. 水氡异常的水动力学机制[J]. 地震地质, 1997, 19(4): 353-357.
CHE Yongtai, YU Jinzi, LIU Wuzhou. The hydrodynamic mechanism of water radon anomaly[J]. Seismology and Geology, 1997, 19(4): 353-357.
- [14] 王俊, 缪阿丽, 李军辉, 等. 2011 年安徽安庆 M4.8 地震前苏、皖典型前兆异常特征分析[J]. 地震研究, 2018, 41(4): 568-576, 657.
WANG Jun, MIAO Ali, LI Junhui, et al. Analysis on characteristics of precursory anomalies of the 2011 Anqing M4.8 earthquake in Anhui and Jiangsu[J]. Journal of Seismological Research, 2018, 41(4): 568-576, 657.
- [15] 田雷, 余怀忠, 周志华. 基于水化学分析的滇 17 井流体异常识别及应用[J]. 地震研究, 2018, 41(4): 551-558, 657.
TIAN Lei, YU Huaizhong, ZHOU Zhihua. Fluid anomalies detection of the Dian-17 well based on geochemical analysis and its application[J]. Journal of Seismological Research, 2018, 41(4): 551-558, 657.