

王岩, 曹凤娟, 王亮, 张博, 翟丽娜. 2016. 基于数据挖掘的东北华北 $M_s \geq 5.0$ 地震活动参数综合指标分析. 地震学报, 38(6): 906-913. doi:10.11939/jass.2016.06.010.

Wang Y, Cao F J, Wang L, Zhang B, Zhai L N. 2016. Comprehensive index analyses on seismic activity parameters based on data mining for $M_s \geq 5.0$ earthquakes in northeast China and North China. *Acta Seismologica Sinica*, 38(6): 906-913. doi:10.11939/jass.2016.06.010.

基于数据挖掘的东北华北 $M_s \geq 5.0$ 地震活动参数综合指标分析^{*}

王 岩[†] 曹凤娟 王 亮 张 博 翟丽娜

(中国沈阳 110034 辽宁省地震局)

摘要 根据地震活动参数的数据特征, 以 2013 年内蒙古科尔沁 $M_s 5.3$ 地震为例, 引入主成分分析和因子分析两种数据挖掘方法, 在信息损失尽可能少的条件下, 实现对参数变量的约简降维, 并提取综合指标 W 。研究显示, 2013 年科尔沁 $M_s 5.3$ 地震前两年, W 值变化显著。在进一步的单因子分析中, 从综合后的参数变量信息中重新提取了具有物理意义的多个单因子综合指标, 消除了大量信息重叠所造成的不一致, 实现了对综合指标的细化分析。在此基础上, 对东北华北 29 次 $M_s \geq 5.0$ 地震的 11 个地震活动参数(频次 N , b 值, η 值, $A(b)$ 值, M_f 值, AC 值, C 值, D 值, E 值, R_m 值, 响应比 Y)进行主成分分析, 其结果显示, 主成分综合指标 W 在震前 1—2 年均出现了明显的异常变化, 这充分说明综合指标 W 可以用作地震预报研究的综合异常参考判据。

关键词 地震活动参数 主成分分析 因子分析 综合指标

doi:10.11939/jass.2016.06.010 中图分类号: P315.5 文献标志码: A

Comprehensive index analyses on seismic activity parameters based on data mining for $M_s \geq 5.0$ earthquakes in northeast China and North China

Wang Yan[†] Cao Fengjuan Wang Liang Zhang Bo Zhai Lina

(Earthquake Administration of Liaoning Province, Shenyang 110034, China)

Abstract: According to the data characteristics of seismic activity parameters, taking the $M_s 5.3$ Horqin, Inner Mongolia earthquake in 2013 for a typical case, the present paper introduces principal component analysis method and factor analysis method to reduce the dimension of parameter variables on the condition of information loss as little as possible, and extracts the comprehensive index. The result shows that the comprehensive index W had been changed significantly two years before the earthquake. Furthermore, by conducting factor analysis on

* 基金项目 辽宁省地震局专项项目(LZ-201606)和震情跟踪定向工作任务(2015010112)共同资助。

收稿日期 2016-01-15 收到初稿, 2016-07-01 决定采用修改稿。

† 通讯作者 e-mail: wangyancau@126.com



Horqin earthquake, several single factor comprehensive indices with physical meanings are extracted, not only eliminating the inconsistency caused by the overlapping of information, but also achieving the refinement of the comprehensive index. In further research, the principal component analysis result about the eleven seismic activity parameters ($N, b, \eta, A(b), M_f, AC, C, D, E, R_m, Y$) of 29 earthquakes with $M_s \geq 5.0$ in northeast China and North China shows that comprehensive index W of principal component analysis had obvious precursor changes in 1—2 years before the earthquakes. This suggests that the index W can be taken as a precursory in earthquake prediction research.

Key words: seismic activity parameters; principal component analysis; factor analysis; comprehensive index

引言

地震活动参数是指地震活动性研究中经多年实践积累得到的与地震时、空、强以及地下介质变化等信息紧密相关的变化指标(国家地震局预测预防司, 1997; 中国地震局监测预报司, 2002). 常用的地震活动参数中: b 值, η 值和 M_f 值反映 G-R 关系拟合中的偏离信息, R_m 值反映固体潮的调制作用, C 值反映地震空间集中度, 响应比 Y 反映介质的稳定程度, 这些参数均与地震孕育地点的地下介质变化相关; D 值和 AC 值分别反映地震在时间分布上的丛集程度和演化特征; 能量 E , 频度 N 和 $A(b)$ 值则反映地震强度信息(陆远忠等, 1999; 华爱军等, 2001; 韩渭宾, 2003; 中国地震局监测预报司, 2007). 运用不同的地震活动参数能够从不同侧面挖掘震前震后的异常变化信息, 但由于参数自身固有的特性, 在实际应用过程中也常常出现异常特征不同步、不一致的现象(王炜等, 2006a, b; 李永振, 2011), 难以判定震情趋势. 对众多地震活动参数的分析结果(韩渭宾, 2003)显示, 虽然不同参数所反映的地震活动性特征的角度不同, 但大都存在一定的相关性. 综合众多参数信息, 统一不同参数的判定结果, 是目前亟需解决的问题. 鉴于此, 本文拟引入主成分分析和因子分析两种数据挖掘方法(Hotelling, 1933; Gunopoulos, Das, 2001; Moghadam, 2002), 将众多参数变量降维归纳为少数几个弱相关的综合变量, 以达到重复信息约简、综合指标提取分析过程简化的目的(王岩等, 2016).

1 主成分分析与因子分析

为避免遗漏重要信息, 研究中通常会分析尽可能多的指标, 这在保证信息完整的同时却增加了分析问题的复杂性, 而且也会造成大量信息的重叠. 通过主成分分析法和因子分析法可将具有错综复杂关系的变量归结为少数几个综合因子.

1.1 主成分分析法

主成分分析法是通过正交变换将一组可能存在相关性的变量(指标)转换为一组线性不相关的新变量(综合指标)的数学降维方法.

设 x_i ($i = 1, 2, \dots, p$) 为地震活动参数变量, 进行主成分分析后得到的主成分 Z_i ($i = 1, 2, \dots, m$ 且 $m < p$). Z_i 为 x_i 的线性组合, 包含绝大多数的参数变量信息, 相互之间基本不相关. 第一主成分 Z_1 , 对应于数据变异(贡献率 e_1)最大的方向, 所包含的信息最为丰富. 对于 Z_2, Z_3, \dots, Z_m , 依次有贡献率 $e_2 \geq e_3 \geq \dots \geq e_m$. 基于各主成分的贡献率, 按照主

成分分析方法的步骤(王炜, 2006a; 李永振, 2011), 计算得到主成分综合指标

$$W = \sum_{i=1}^m e_i Z_i, \quad (1)$$

该指标是能够反映地震时、空、强和地下介质异常的综合判定指标.

1.2 因子分析法

因子分析法是主成分分析法的发展, 其关键在于正交因子模型的建立和应用.

地震活动参数正交因子模型的建立如下: 存在 $m (m \leq p)$ 个公共因子 $f_i (i = 1, 2, \dots, m)$, 使参数变量 x_i 可用其线性组合表示为

$$\mathbf{X} = \mathbf{AF} + \boldsymbol{\varepsilon}, \quad (2)$$

式中: $\mathbf{X} = (x_1, x_2, \dots, x_p)^T$; \mathbf{F} 为公共因子, $\mathbf{F} = (f_1, f_2, \dots, f_m)^T$; $\boldsymbol{\varepsilon}$ 为误差或特殊因子, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)^T$; \mathbf{A} 为因子载荷矩阵 $\mathbf{A} = (a_{ij})_{p \times m}$, 其中 a_{ij} 为第 i 个指标在第 j 个公共因子上的载荷, 是第 i 个变量与第 j 个公共因子的相关系数; 特殊因子 ε_i 相互独立, 且服从正态分布 $N(0, \sigma^2)$.

设原始数据矩阵为

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix},$$

这里 n 为样本数, p 为变量数.

由式(2)可以得到初始因子载荷矩阵 \mathbf{A} . \mathbf{A} 中通常有一些变量在几个公共因子上均有较大载荷, 使得初始公共因子的物理意义难以得到合理解释, 这时可旋转载荷矩阵 \mathbf{A} , 用一个正交阵右乘, 使旋转后的因子载荷阵结构简化, 便于对公共因子进行解释. 结构简化就是使每个变量仅在一个公共因子上有较大的载荷, 这种变换因子载荷的方法称为因子旋转.

基于主成分分析的过程, 根据各公共因子得分 f_i 和相应的权值(贡献率) e_i , 因子分析的综合指标为

$$W_{fa} = \sum_{i=1}^m e_i f_i, \quad (3)$$

将多参数中的信息合并后按照类别再分类, 可用单因子综合指标表示为

$$W_{fai} = e_i f_i. \quad (4)$$

与主成分综合指标相比, 单因子综合指标可以根据需求进行因子筛选, 且筛选后所提取的单因子综合指标能够更细化地表现多类别的特征.

2 2013 年科尔沁 $M_{\text{S}}5.3$ 地震综合指标分析

2.1 地震活动参数分析

以 2013 年 4 月 22 日内蒙古科尔沁 $M_{\text{S}}5.3$ 地震为例, 计算得到地震发生前后震中周围 200 km 范围内的 11 个地震活动性参数(频次 N , b 值, η 值, $A(b)$ 值, M_{f} 值, AC 值, C 值, D 值, E 值, R_{m} 值, 响应比 Y)随时间的变化曲线, 如图 1 所示. 参数计算震级为

$M_L 2.5—6.0$, 累计时间为 18 个月, 滑动步长为 1 个月。从该图中可见, 不同参数在地震发生前变化规律明显不一致。

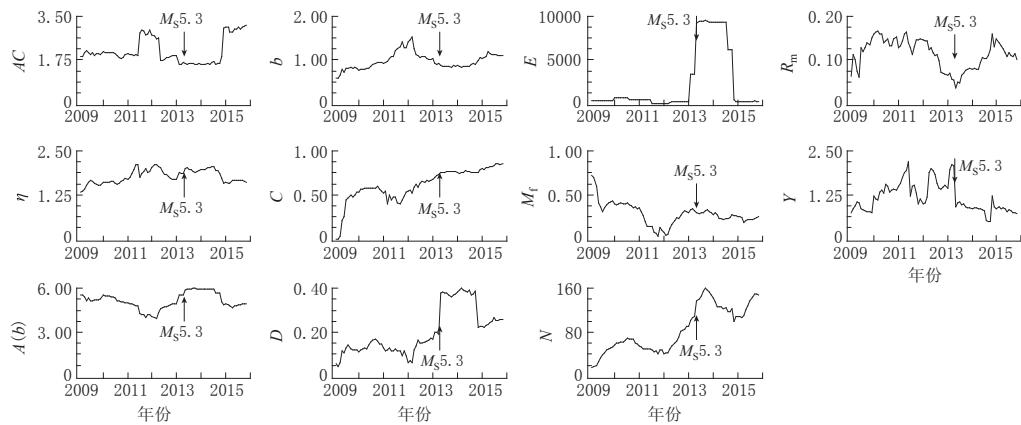


图 1 2013 年科尔沁 $M_s 5.3$ 地震 11 个地震活动参数的变化曲线

Fig. 1 Variations of 11 seismic activity parameters before and after 2013 Horqin $M_s 5.3$ earthquake

对选取的 11 个参数进行线性相关性计算, 结果列于表 1, 表中相关系数的数值大小代表相关性强弱, 正负代表相关方向。可以看出, 除响应比 Y 与各参数的相关性较弱外, 其余各参数之间均存在较强的相关性。其中: 频次 N , C 值, D 值, 能量 E 等 4 个与地震强度相关性较大的参数之间的相关性相对更强; η 值, $A(b)$ 值, AC 值, R_m 值等 4 个与地震发生地地下介质变化相关较大的参数之间的相关性相对更高。

表 1 2013 年科尔沁 $M_s 5.3$ 地震 11 个活动性参数的相关性系数

Table 1 Correlativity between the 11 seismic activity parameters of 2013 Horqin $M_s 5.3$ earthquake

	N	b	η	$A(b)$	M_f	AC	C	D	E	R_m	Y
N	1.000	-0.201	0.205	0.497	0.395	0.184	0.882	0.850	0.635	0.169	-0.047
b	-0.201	1.000	-0.010	-0.299	-0.389	-0.146	-0.167	-0.244	-0.245	-0.230	0.043
η	0.205	-0.010	1.000	0.805	0.067	0.574	0.177	0.259	0.268	0.294	-0.097
$A(b)$	0.497	-0.299	0.805	1.000	0.533	0.713	0.438	0.540	0.386	0.388	-0.078
M_f	0.395	-0.389	0.067	0.533	1.000	0.465	0.391	0.403	0.114	0.187	0.018
AC	0.184	-0.146	0.574	0.713	0.465	1.000	0.176	0.200	-0.117	0.419	-0.005
C	0.882	-0.167	0.177	0.438	0.391	0.176	1.000	0.670	0.473	0.141	0.064
D	0.850	-0.244	0.259	0.540	0.403	0.200	0.670	1.000	0.606	0.247	-0.139
E	0.635	-0.245	0.268	0.386	0.114	-0.117	0.473	0.606	1.000	0.021	-0.141
R_m	0.169	-0.230	0.294	0.388	0.187	0.419	0.141	0.247	0.021	1.000	0.035
Y	-0.047	0.043	-0.097	-0.078	0.018	-0.005	0.064	-0.139	-0.141	0.035	1.000

进一步检验可知: 参数变量数据 KMO(Kaiser-Meyer-Olkin) 检验(Kaiser, 1960; Hill, 2011)的统计量结果为 0.663, 说明各参数之间的相关性较强; Bartlett's 球型检验(Bartlett, 1937)的相伴概率 P 值小于设定显著性水平 0.05, 独立性假设不成立, 也说明各参数之间存在相关性, 即研究震例的地震活动参数作为变量符合主成分分析法和因子分析法对数据的要求。

2.2 主成分分析

科尔沁地震活动参数主成分分析中的成分特征值、贡献率和主成分累计贡献率列于表

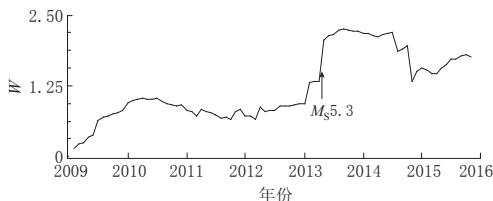
2, 可见少数几个主成分中已经涵盖了大部分的参数信息. 参数变量主成分载荷矩阵列于表 3, 可以看出表中与其它参数相关性较弱的响应比 Y 在对不同主成分的影响程度上是反

表 2 2013 年科尔沁 $M_s 5.3$ 地震活动

参数主成分特征值和贡献率

Table 2 Eigenvalues and contribution rate of Horqin $M_s 5.3$ earthquake in 2013

主成分	特征值	贡献率	累计贡献率
1	4.331	39.371%	39.371%
2	1.920	17.451%	56.822%
3	1.232	11.201%	68.023%
4	1.049	9.532%	77.555%
5	0.820	7.457%	85.012%
6	0.726	6.600%	91.613%
7	0.370	3.366%	94.979%
8	0.257	2.335%	97.314%
9	0.199	1.809%	99.123%
10	0.054	0.488%	99.611%
11	0.043	0.389%	100.000%

图 2 2013 年科尔沁 $M_s 5.3$ 地震前后综合指标 W 的变化Fig. 2 Variation of the comprehensive index W before and after 2013 Horqin $M_s 5.3$ earthquake表 4 2013 年科尔沁 $M_s 5.3$ 地震因子载荷系数Table 4 Load coefficient of factors of Horqin $M_s 5.3$ earthquake in 2013

参数	因子载荷系数			
	1	2	3	4
N	0.940	0.136	0.153	0.059
b	-0.122	-0.050	-0.822	0.190
η	0.195	0.853	-0.272	-0.210
$A(b)$	0.418	0.837	0.187	-0.101
M_f	0.271	0.312	0.705	0.179
AC	-0.025	0.871	0.227	0.113
C	0.854	0.125	0.138	0.229
D	0.844	0.205	0.205	-0.093
E	0.784	-0.008	-0.016	-0.326
R_m	0.022	0.546	0.287	0.089
Y	-0.026	-0.020	-0.055	0.898

表 3 2013 年科尔沁 $M_s 5.3$ 地震活动

参数主成分载荷系数

Table 3 Load coefficient of principal component of Horqin $M_s 5.3$ earthquake in 2013

参数	主成分载荷系数				
	1	2	3	4	5
N	0.832	-0.461	-0.003	0.157	-0.059
b	-0.395	-0.009	0.499	0.570	-0.326
η	0.558	0.512	0.544	0.122	0.095
$A(b)$	0.859	0.400	0.151	0.002	-0.049
M_f	0.600	0.137	-0.546	-0.146	-0.422
AC	0.547	0.718	-0.053	0.075	-0.228
C	0.748	-0.407	-0.103	0.285	-0.112
D	0.823	-0.354	0.048	-0.004	0.006
E	0.596	-0.514	0.300	-0.111	0.270
R_m	0.417	0.438	-0.151	-0.009	0.555
Y	-0.092	0.055	-0.486	0.751	0.275

向增加的, 但其在选取 5 个主成分的情况下, 仍能发挥较大作用, 说明主成分综合分析中的信息涵盖是全面的.

选用 5 个主成分, 在信息涵盖率高于 85% 的情况下, 计算科尔沁地震发生前后的综合指标 W , 得到 W 随时间变化的曲线, 如图 2 所示. 可以看出, 在地震发生前两年, 综合指标有明显的异常变化.

2.3 因子分析

为了进一步解析综合指数的异常变化, 在主成分分析的基础上, 对参数变量矩阵进行旋转, 重新整合信息, 得到因子载荷矩阵, 如表 4 所示. 与表 3 中主成分载荷对比可知, 经过旋转的因子对不同变量的敏感度发生了变化, 在承载多参数统计意义信息的同时, 较好地表达出了综合信息单因子的不同物理倾向. 因子分析将单一参数的载荷尽可能集中于单一因子上, 在表达统计意义的基础上, 更便于不同因子物理意义的解释. 其中: 因子 1 对频次 N , C 值, D 值和能量 E 这 4 个与地震强度相关较大的参数更为敏感; 因子 2 对 η 值, $A(b)$ 值, AC 值和 R_m 值这 4 个与地

震发生地地下介质变化相关较大的参数更为敏感; 因子 3 受到 b 值和 M_f 值不同方向的较大幅度影响, 也与介质变化特征相关; 因子 4 则受到响应比 Y 的影响最大, 变化趋势与响应比单参数变化相似。

选择 4 个因子计算单因子综合指标 W_{fai} 的结果如图 3 所示, 可以看出不同单因子综合指标的变化规律均有其自身的特点。与地震强度相关的 W_{fa1} 在地震发生前两年开始有显著变化; 与地下介质变化相关的 W_{fa2} 和 W_{fa3} 则从地震发生前 4 年开始有变化; 由于因子 4 受到响应比 Y 的影响很大, W_{fa4} 的变化规律与图 1 中响应比 Y 的变化类似。单因子综合指标排除了信息重叠对分析结果可靠性的影响, 同时实现了信息最小完整集内特征的合理分类。不同单因子综合指数的时间变化规律, 代表了不同物理特征的先后变化规律, 符合地震孕育发生过程中由介质变化、能量聚集及地震强度所反映的普遍特征。

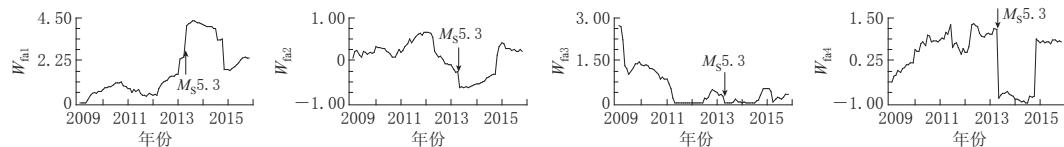


图 3 2013 年科尔沁 $M_s 5.3$ 地震单因子综合指标 W_{fai} 的变化曲线

Fig. 3 Variations of the single factor comprehensive index W_{fai} before and after 2013 Horqin $M_s 5.3$ earthquake

3 东北华北中强震综合指标分析

为了进一步对综合指标方法的区域可用性展开分析, 在典型震例的研究基础上, 本文对 1970 年以来东北华北地区($34^{\circ}\text{--}55^{\circ}\text{N}$, $110^{\circ}\text{--}135^{\circ}\text{E}$)地震目录中较为完整的 28 次中强震进行了 11 个地震活动参数(地震频次 N , b 值, η 值, $A(b)$ 值, M_f 值, AC 值, C 值, D 值, R_m 值, 能量 E , 响应比 Y)的主成分分析。参数计算中, 按照地震孕育过程的可能影响范围, 取以震中为圆心的圆域($M_s 5.0\text{--}5.4$, 半径为 200 km; $M_s 5.5\text{--}5.9$, 半径为 250 km; $M_s 6.0\text{--}7.0$, 半径为 300 km; $M_s \geq 7.0$, 半径为 350 km), 选取 $M_t \geq 2.5$ 地震序列, 以窗长 1.5 年、步长 1 个月进行计算。图 4 给出了 28 次中强震前后主成分综合指标 W 随时间的变化曲线, 可以看出, 在地震发生前 1—2 年, 综合指标 W 均有显著的变化。发生在郯庐断裂带的 1975 年海城 $M_s 7.3$, 1978 年营口 $M_s 5.9$, 1995 年苍山 $M_s 5.2$, 1999 年岫岩 $M_s 5.4$, 2013 年灯塔 $M_s 5.1$ 这 5 次地震前 1—2 年综合指标 W 均呈明显下降的特征; 发生在张渤断裂带的 1978 年唐山 $M_s 7.8$, 1995 年滦县 $M_s 5.0$ 和 1998 年张北 $M_s 6.2$ 地震前 1—2 年, 综合指标 W 均呈升高与降低交替出现的特征, 且震后 W 呈快速上升; 发生在山西断裂带的 1981 年丰镇 $M_s 5.5$, 1989 年大同 $M_s 6.1$, 1998 年张北 $M_s 6.2$ 和 1999 年广灵 $M_s 5.6$ 这 4 次地震中, 丰镇地震与广灵地震的综合指标特征相似, 即震前 1—2 年综合指标 W 呈显著下降, 而大同地震和发生在张渤断裂带、山西断裂带交界处的张北地震则与张渤断裂带地震的特征相似。由此说明, 不同断裂带地区地震综合指标的特征具有一定的固有特色, 主成分载荷变化特征也较为相似, 这与所选地震活动参数中包含反映地下介质信息的参数有关。

4 讨论与结论

传统的地震活动参数大都具有明确的物理意义, 包含着丰富多样的信息, 但由于数据

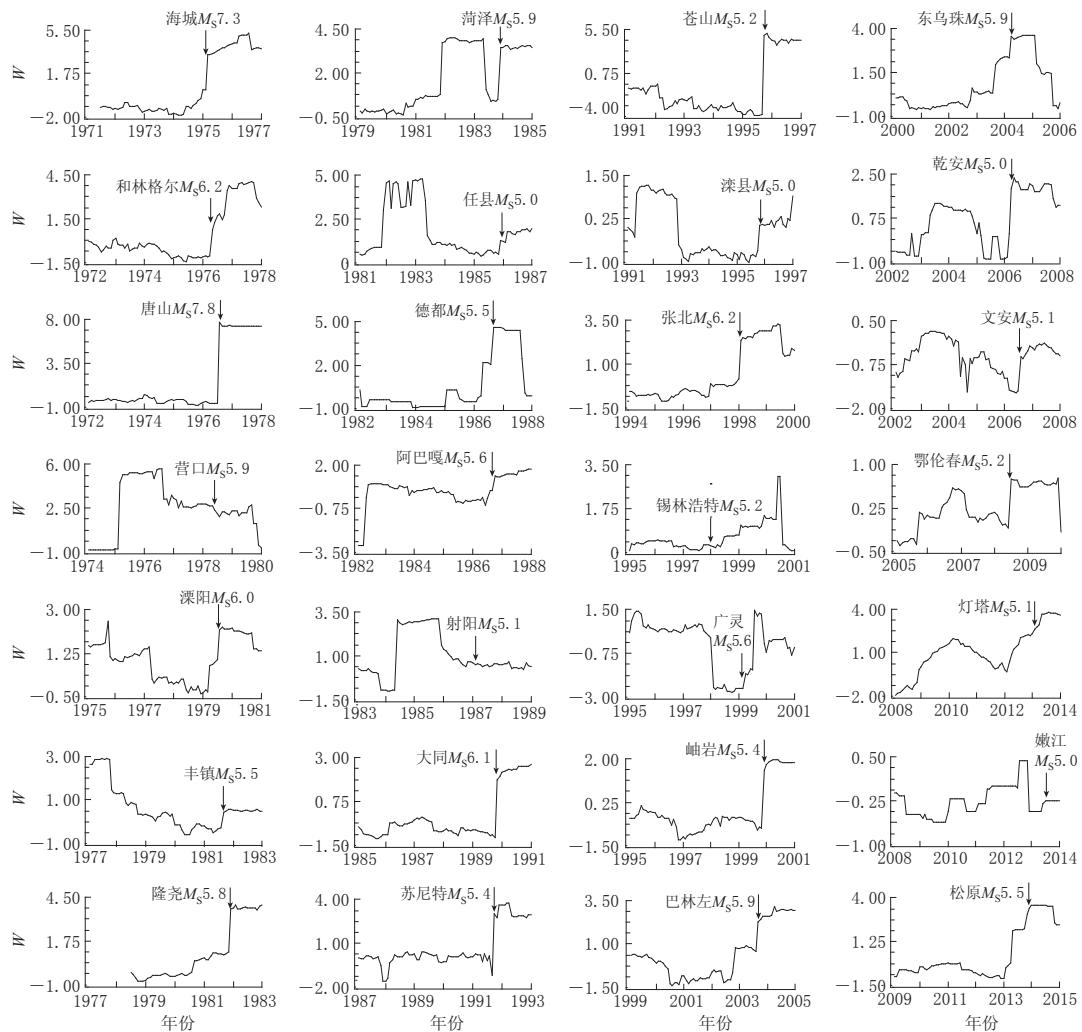


图 4 东北华北 28 次 $M_s \geq 5.0$ 地震综合指标 W 的变化曲线

Fig. 4 Variations of the comprehensive index W before and after 28 earthquakes with $M_s \geq 5.0$ occurred in northeast China and North China

信息大量重叠、分析手段单一、交互综合复杂等问题，其在应用上颇为受限。主成分分析和因子分析两种数据挖掘方法，依据不同参数本身的特点及其相互关系，对信息进行重新整合，完成了参数的降维，简化了多参数变量的分析过程，合理地避免了信息矛盾问题，实现了融合信息特征的再提取。1970 年以来，东北华北的 29 次 $M_s \geq 5.0$ 地震的综合指数 W 分析结果表明，综合指标在地震发生前均有明显的前兆变化。通常在地震发生前的 1—2 年内，综合指数均有下降再上升的趋势变化，且发震震级越大，影响时段越长，变化程度越显著，这一变化特征可以作为地震预报研究中的判据。进一步对单因子综合指标 W_{fai} 分析，从参数变量统计学变换中挖掘合理的物理解释，在综合指标约简降维的同时，细化了地震异常分析，使综合指数的应用更为广泛。

参 考 文 献

- 国家地震局预测预防司. 1997. 测震学分析预报方法[M]. 北京: 地震出版社: 49-116.
- Department of Prediction and Prevention, State Seismological Bureau. 1997. *Analysis and Prediction Methods of Seismometry*[M]. Beijing: Seismological Press: 49-116 (in Chinese).
- 韩渭宾. 2003. 地震活动性参数分类及其相关性初步研究[J]. 四川地震, (3): 1-5.
- Han W B. 2003. Primary study on the classifying among the seismicity parameters and their correlativity[J]. *Earthquake Research in Sichuan*, (3): 1-5 (in Chinese).
- 华爱军, 刘西林, 刁守中, 李红. 2001. 华北地区中等、大地震前算法复杂性 AC 值异常变化特征[J]. 内陆地震, 15(4): 326-330.
- Hua A J, Liu X L, Diao S Z, Li H. 2001. Abnormal characteristics of arithmetic complexity AC value before moderate, strong earthquakes in North China[J]. *Inland Earthquake*, 15(4): 326-330 (in Chinese).
- 李永振. 2011. 主成分分析法在辽宁地区地震预测中的应用[J]. 西北地震学报, 33(1): 76-79.
- Li Y Z. 2011. Application of primary components analysis method to earthquake prediction in Liaoning Province[J]. *Northwestern Seismological Journal*, 33(1): 76-79 (in Chinese).
- 陆远忠, 阎利军, 郭若眉. 1999. 用于中短期地震预报的一些地震活动性参数相关性讨论[J]. 地震, 19(1): 11-18.
- Lu Y Z, Yan L J, Guo R M. 1999. Discussion about correlation of some seismological parameters in medium- and short-term earthquake prediction[J]. *Earthquake*, 19(1): 11-18 (in Chinese).
- 王炜, 林命週, 赵利飞, 马钦忠. 2006a. 地震活动参数约简的因子分析方法[J]. 西北地震学报, 28(4): 303-308.
- Wang W, Lin M Z, Zhao L F, Ma Q Z. 2006a. Factor analysis method for reducing seismicity parameters[J]. *Northwestern Seismological Journal*, 28(4): 303-308 (in Chinese).
- 王炜, 刘锐, 李国正, 吴耿峰, 马钦忠, 赵利飞, 林命週. 2006b. 我国大陆强震预测的支持向量机方法[J]. 地震学报, 28(1): 29-36.
- Wang W, Liu Y, Li G Z, Wu G F, Ma Q Z, Zhao L F, Lin M Z. 2006b. Support Vector Machine method for forecasting future strong earthquakes in Chinese mainland[J]. *Acta Seismologica Sinica*, 28(1): 29-36 (in Chinese).
- 王岩, 邵媛媛, 曹凤娟, 郭晓燕. 2016. 基于因子分析的地震综合指数在辽宁地区地震预测中的应用[J]. 防灾减灾学报, 32(1): 1-7.
- Wang Y, Shao Y Y, Cao F J, Guo X Y. 2016. Application of comprehensive index to earthquake prediction of Liaoning using factor analysis method[J]. *Journal of Disaster Prevention and Reduction*, 32(1): 1-7 (in Chinese).
- 中国地震局监测预报司. 2002. 强地震中短期预报方法及其效能研究[M]. 北京: 地震出版社: 45-87.
- Department of Monitoring and Prediction, China Earthquake Administration. 2002. *Study on Medium- and Short-Term Forecast Methods and Their Efficiency*[M]. Beijing: Seismological Press: 45-87 (in Chinese).
- 中国地震局监测预报司. 2007. 中国大陆地震序列研究[M]. 北京: 地震出版社: 20-24.
- Department of Monitoring and Prediction, China Earthquake Administration. 2007. *Researches on Earthquake Sequences of Chinese Mainland*[M]. Beijing: Seismological Press: 20-24 (in Chinese).
- Bartlett M S. 1937. Properties of sufficiency and statistical tests[J]. *Proc R Soc A*, 160(901): 268-282.
- Gunopulos D, Das G. 2001. Time series similarity measures and time series indexing[J]. *AcM Sigmod Record*, 31(2): 624.
- Hill B D. 2011. *The Sequential Kaiser-Meyer-Olkin Procedure as an Alternative for Determining the Number of Factors in Common-Factor Analysis: A Monte Carlo Simulation*[D]. Stillwater, Oklahoma State: Oklahoma State University: 40-43.
- Hotelling H. 1933. Analysis of a complex of statistical variables into principal components[J]. *J Educ Psychol*, 24(6): 417-441.
- Kaiser H F. 1960. The application of electronic computers to factor analysis[J]. *Educ Psychol Measur*, 20(1): 141-151.
- Moghaddam B. 2002. Principal manifolds and probabilistic subspaces for visual recognition[J]. *IEEE Trans Pattern Anal Mach Intell*, 24(6): 780-788.