

模式识别方法在地震综合预测研究中的应用

张郢珍 邱竞男 粟生平 关华平

(国家地震局分析预报中心)

摘 要

本文阐述了模式识别方法的原理及其在综合地震预测中的实用性与可行性。文中以华北和西北地区为例,提取了中强地震前的前兆与地震活动异常特征,根据计算结果估计,未来两年内在西北甘青交界武威—西宁以西(37° — 38° N, 101° — 102° E)范围内可能发生6级地震的预测意见。

本文将判别分类的费歇尔判别准则与模式识别中的CORA—3算法应用于地震综合预测、预报中。根据地震孕育过程的物理模式、地震活动及某些前兆观测资料,清理和分析了华北1966—1984年,西北1970—1984年的 $M \geq 5.5$ 级地震前的异常特征,同时对有异常无震的情况也进行了统计。针对华北地区提出了地震活动性及某些前兆观测项目的定量异常标准,由此进一步提出了人机联合判断震情的工作流程,并将此工作方法及我们所编制的计算程序用于华北震情的综合分析及西北地区的地震活动的综合分析中。

一、模式识别方法简介及人机联合判断震情的方案及步骤

从数学上看,决策判别可采用模糊数学或模式识别(或称图象识别)方法。在模式识别中对于要预报或不预报某一等级的地震问题,可归结为两类模式识别分类问题。

两类问题的线性判别函数的一般形式及决策规划为:

$$g(x) = W^T x \begin{cases} \geq R_0 \rightarrow x \in \omega_1 \\ < R_0 \rightarrow x \in \omega_2 \end{cases}$$

其中 W^T 称为权向量, R_0 称为阈值, x 为多维空间的点(样本)。方程 $g(x) = R_0$ 定义了一个决策面,它是一个超平面,将特征空间分成为两个半空间,把属于 ω_1 类的点同属于 ω_2 类的点划分开来。即对 ω_1 的决策域为 R_1 ,而对 ω_2 的决策域为 R_2 ,当 x 在 R_1 一侧时属 ω_1 类,当 x 在 R_2 一侧时属 ω_2 类,因此模式识别的分类问题就取决于如何确定 W 和 R_0 。

本文采用模式识别中CORA—3算法来确定判别函数中的特征向量,对所识别的问题

震形势中起主要作用。表1给出了上述12项识别标志在有震区与有异常而无震地区出现次数的比例。

各项异常在有震样本与无震样本中出现的比例 表1

识别标志序号	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
有震样本区	0.80	0.80	0.80	0.50	0.50	0.80	0.50	0.70	0.40	0.80	0.40	0.70
无震样本区	0.25	0.37	0.25	0.50	0	0.50	0.37	0.50	0.25	0.50	0	0.50

由费歇尔准则下的判别函数计算得到

$$R_0 = 6.313, \quad R_1 = 9.120, \quad R_2 = 2.803$$

按 R_0 判断已知样本得到：有震区的误识率为0%，无震区的误识率为25%。在9个待预测区的R值中，8个区之值落在 R_0 与 R_2 一侧，由此可确定这些地区发生大震的可能性不大。仅第II区（即丰镇所在区）的R值为6.812，稍大于 R_0 ，则该区应判为有可能发生5.8级以上地震的地区。由于计算中所选取的识别标志有限，加之已知样本数较少，因此计算结果可能与实际情况有差别。但上述模式识别与判别分类的方法可以实现在某种物理模式的前提下用人机联合的办法来判断震情，从而避免了人们由于某些心理作用造成的虚报。

三、西北地区地震活动性的模式识别分析

1. 地震活动性指数的扫描计算结果

利用国家地震局分析预报中心数据库的西北地震目录（1970年1月—1984年6月），取 $33^\circ-40^\circ N, 95^\circ-106^\circ E$ 范围内 $3^\circ \times 3^\circ$ 的区域进行地震活动性指数（b值、频度、 \sqrt{E} ）的扫描计算。扫描步长为 2° （即各区在经、纬度重叠 1° ），起算震级为2.0级。由于该地区地震台网分布的不均匀，对地震的控制能力相差较大，我们在求各区地震活动性指数多年平均值时适当进行了一些舍取。图2给出了西北各区地震活动性指数的多年平均值，以此作为该区地震活动性背景。

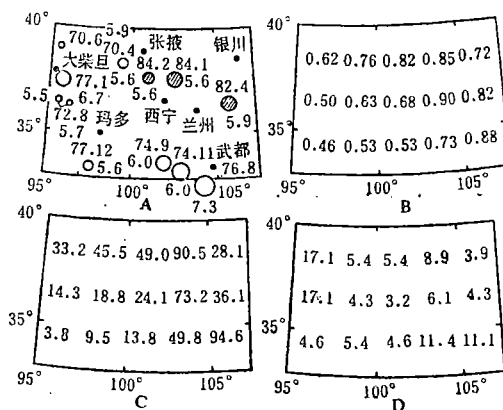


图2 西北地区地震活动背景参数

A. $M_L \geq 5.5$ 级地震震中分布图 B. 滑动b值多年平均值 C. 滑动频度多年平均值

D. 应变释放多年平均速率 (10^9 尔格 $^{1/2}$ /年)

Fig. 2 The background parameters of earthquake activity in the North-west area

图2 A绘出了1970—1984年 $M_L \geq 5.5$ 级地震分布, 1977年以前地震活动主要在青海西部与甘青川交界地区, 80年代开始地震主要集中在本区中部甘宁青交界地区。平均应变释放速率(图2 D)较高的是本区西北角与东南角, 应变释放速率分别为 17.1×10^9 尔格 $^{1/2}$ /年与 11.4×10^9 尔格 $^{1/2}$ /年, 而其它各区约为 5×10^9 尔格 $^{1/2}$ /年。从b值与小震频率来看, 本区东北部地震频度较西南高, 前者每年平均频度最低、最高值分别为33—94, 而西南部分别为4—24。高b值区也在本区东北部, 其数值为0.68—0.90, 而西南部b值为0.46—0.63, 这可能与地震台网稀疏小地震缺记有关。

2. 识别标志的选取

根据上述西北地区地震活动背景性参数, 我们选取如下11项相对变化量作为识别标志, 由此挑选和确定与地震发生有关的地震活动性指数的异常特征。对西北地区15个 $3^\circ \times 3^\circ$ 的滑动窗, 以每两年的平均值作为统计量, 若满足下列11项标识异常量, 则以“1”表示, 否则以“0”表示。并规定在滑动窗中发生了 $M_L \geq 5.5$ 级地震的地区为有震区, 其它为无震区。

十一项识别标志为:

(1) b_i 值是否 $\leq \bar{b} - 10\% \bar{b}$ (\bar{b} 为该统计窗的b值背景基值, 如图2所示, b_i 为该地区在统计年份的b值); (2) b_i 值是否 $\leq \bar{b} - 15\% \bar{b}$; (3) b_i 值是否 $\geq \bar{b} + 10\% \bar{b}$; (4) b_i 值是否 $\geq \bar{b} + 15\% \bar{b}$; (5) N_i 是否 $\geq 1.5 \bar{N}$ (\bar{N} 为该统计窗的多年平均年频度, N_i 为该窗在统计年份中的年频度); (6) N_i 是否 $\geq 2.0 \bar{N}$; (7) N_i 是否 $> \bar{N}$; (8) $\sqrt{E_i}$ 是否 $< 1.0 \sqrt{\bar{E}}$ ($\sqrt{\bar{E}}$ 为该统计窗口的多年平均年应变释放速率, $\sqrt{E_i}$ 为统计年份中的平均年应变释放速率); (9) $\sqrt{E_i}$ 是否 $\geq 2 \sqrt{\bar{E}}$; (10) $\sqrt{E_i}$ 是否 $\geq 3 \sqrt{\bar{E}}$; (11) $\sqrt{E_i}$ 是否 $> \sqrt{\bar{E}}$ 。

考虑到1970—1971年西北地区地震记录可能不全, 统计从1972年开始到1984年止, 同时顾及到中强震发生后对该窗口的影响, 震后两年的时段未进行统计, 并以1983—1984年之值作为各窗口的预测值, 由此得到10个有震样本窗、46个无震样本窗以及11个预测样本窗。

3. 西北地区中强地震的活动性异常特征及未来地震活动趋势估计

按前述模式识别及判别准则的方法, 对西北地区十二年来的地震活动性异常特征的模式识别得到, 当 K_1 、 K_2 、 K_3 、 K_4 分别取为0.60、0.20、0.60、0.20时, 我们可选取17个危险性特征与49个不危险特征。由表2看出识别标志中1、2、5、7、9、11等项可作为中强震的异常特征, 即中强震前该区b值下降到小于原来基值的10%—15%, 地震活动频度高于多年平均值水平的1~1.5倍, 且年应变释放速率大于多年平均应变释放速率的1—2倍。

由费歇尔准则下的判别函数计算得到:

$$R_0 = 2.224, \quad R_1 = 7.052, \quad R_2 = 1.175$$

内检得到有震区的误识率为10%, 无震区的误识率为18%。在待预测的11个窗中, 有二个窗的判别值接近 R_1 (图3)其判别值均为6.798, 落在 $R_0 \sim R_1$ 之间。由于2个区互相重叠, 我们取其重叠部分作为未来危险地区, 即甘青交界的武威—西宁一带($37^\circ - 38^\circ N$, $101^\circ - 102^\circ E$)为未来2年内可能发生6级地震的地区。

选取作为中、强地震的危险性特征表

表 2

识别标志编号	一维变量			二维变量组合						三维变量组合					
	1	2	3	1 1	1 2	1 3	2 1	2 2	2 3	1 1 1	1 1 2	1 1 3	1 2 1	1 2 2	1 2 3
1				1						1	1				
2	1			1	1	1				1	1	1			
3					0					0		0			
4						0					0	0			
5		1					1	1	1				1	1	1
6															
7							1						1	1	
8								0					0		0
9									1	1				1	1
10															
11			1												

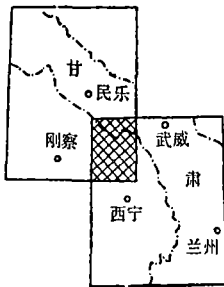


图 3 西北地区地震活动性的模式识别预测结果

Fig. 3 The earthquake activity in the North-west area predicted by using pattern Recognition

四、讨 论

由以上分析与实例看出,模式识别方法在探索地震预报的综合研究中是一个重要的、有希望的途径。该方法能根据问题的物理本质来综合与地震发生有关的诸因素,随着地震研究工作的深入,提供的与地震活动有关的有效因素的资料愈多,模式识别就愈准确。地震资料处理及震情判断工作系统化、计算机化无疑是我们地震预报工作者努力的方向之一、在这方面模式识别方法的应用和发展将有广阔的前景。

(本文1985年10月8日收到)

参 考 文 献

- [1] 张郢珍、粟生平、李格平,华北地区近期地震危险区的确定要素与危险区的综合图象识别,中国地震, №1, 1986.
- [2] 高旭等,我国地震前兆特征的统计分析与研究,地震, №5, 1984.
- [3] 梅世蓉等,一九七六年唐山地震,地震出版社,1982.
- [4] 马宗晋等,1966—1976年中国九大地震,地震出版社,1982.

- [5] I.M.Gelfand, et al, Criteria of high seismicity, determined by pattern recognition, Tectonophysics, Vol.13, №41, 1972.
- [6] I.M.Gelfand, et al, Recognition of places where strong earthquakes may occur, Computational Seismology, Acad Sei USSR (Moscow), №6, 1973.
- [7] 王碧泉, 大震前地震活动的图象识别, 地震学报, Vol. 4, №2, 1982.
- [8] 刘承祚、孙惠文, 数学地质基本方法及应用, 地质出版社, 1981.

APPLYING THE PATTERN RECOGNITION FOR THE RESEARCH OF COMPREHENSIVE EARTHQUAKE PREDICTION

Zhang Yingzhen Qu Jingnan Su Shengping Cuan Huaping
(*Centre for Analysis and Prediction State Seismological Bureau*)

Abstract

In this paper, the principle of the pattern recognition and its practicability in comprehensive earthquake forecasting research has been described. We take the North China area and North west area as examples to extract the precursory features of the moderate-strog earthquakes occurred in these two areas. According to the calculated results, we forecast that an earthquake with magnitude 6 will strike the boundary ares (37° — 38° N 101° — 102° E) of Gansu and Qinghai provinces in two years.