

用模式识别方法判定青海地区潜在震源区

马文静 邹文卫

(青海省地震局)

摘 要

本文用模式识别中修改的CORA—3方法和加权的Hamming方法^[1],对青海地区大量的地质、地貌、地球物理及地震资料进行了分析和研究,分别判定出了该地区6.0、6.5和7.0级地震的潜在震源区,并定量地给出了整体识别的可信度和每个潜在震源区的危险概率。1988年11月5日唐古拉6.8级和1990年1月14日茫崖6.7级地震均发生在本文圈定的潜在震源区内。

一、前 言

模式识别方法目前已在各个领域内得到广泛应用。七十年代我国学者开始将该方法引入地震学领域内^[2-6]。在应用过程中对该方法不断地进行改进¹⁾。在本文中,我们利用修改的CORA—3方法和加权的Hamming方法对青海地区不同震级的潜在震源区进行了判定,同时对每个危险点的稳定性和可信度进行了分析,并给出了数学表达式。本文使用的“PR”程序包是由国家地震局地球物理研究所马秀芳、王碧泉等移植的。

需要说明的是,本文给出的潜在震源区仅是构造意义上的,无时间概念。它只说明某区存在发生某一级地震的条件,其危险概率仅表示它们属于危险域的相对置信度。

二、青海地区的构造特征

青海地处青藏高原中北部,由于受到印度板块的强烈挤压作用,该区活动构造发育并影响到深部,具有独特的地形地貌和地质构造特征。根据区域地貌和构造形态,可将该区分三部分:北部为祁连山褶皱隆起带;中部为柴达木盆地、共和盆地、化隆盆地沉降带;南部为昆仑山和唐古拉山隆起带^[7]。

该区内地壳厚度由北向南加厚。阿尔金山和祁连山一带,地壳厚度变化明显,形成一地壳厚度陡变带。往南在柴达木盆地和共和盆地广大区域,地壳厚度变化平缓,形成地壳变化的平台区。沿库玛断裂带,地壳厚度又开始陡变,由该带向南地壳厚度均匀增加。伴随着地壳厚度的变化,重力异常由北向南也依陡变—平台—陡变—斜坡的规律变化。深部构造的这种

1) 叶洪等,华北潜在震源区的统计图象识别及图象识别方法的改进。

变化规律,同该区的地形地貌形态是相一致的。

该区的现代构造运动,在北部以相对沉降为主,分布有一系列新生代盆地,而南部以相对抬升为主。比较两区的地震活动,北部除阿尔金断裂带史前有过大震外^[8],其余地区均无大震发生。青海地区有文字记载的4次7级以上地震均发生在南部,唐古拉地震带是近期6级地震的多发区。

三、特征和对象的选取

1. 对象的选取

考虑到青海地区的构造特征,我们将其分为两个大区分别进行识别(图1)。北部(I区)包括阿尔金断裂带、祁连山构造带和柴达木盆地;南部(Ⅱ区)包括托素湖地震带和可可西里—三江地震带。

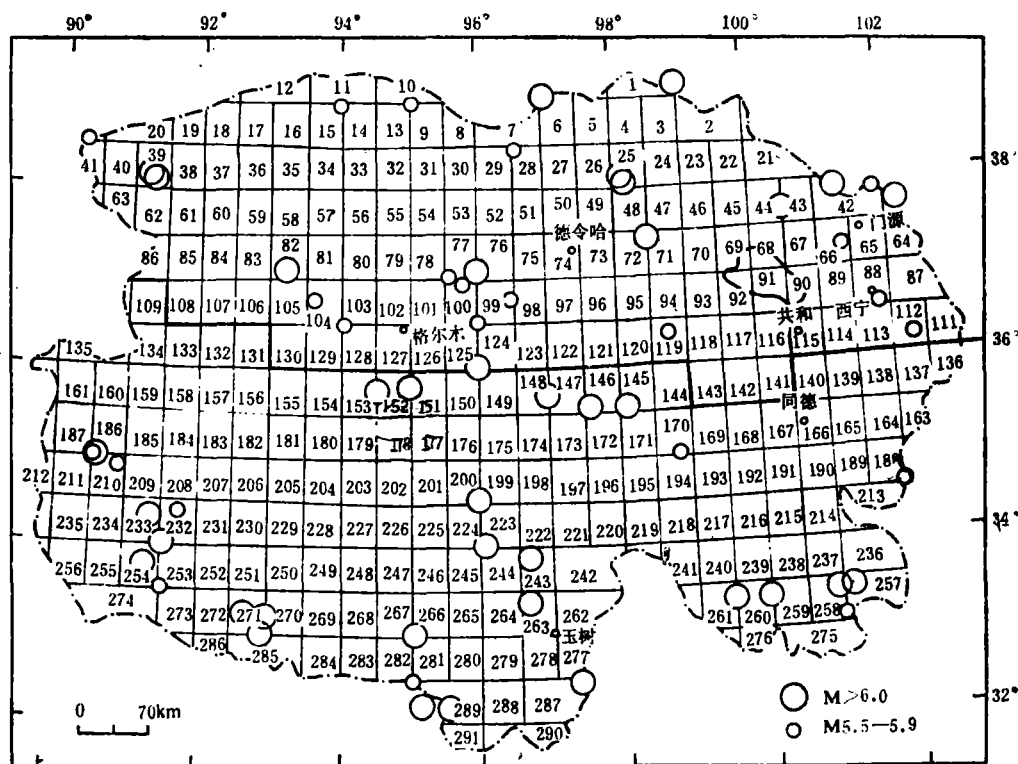


图1 青海地区分区及网格图

Fig. 1 Dividing area and checkerboard in Qinghai Province

由于该区活动构造发育,断裂带密集,很难用其它方法选取对象,因此本文中的对象是按经纬度进行划分的,每一个经纬度所包含面积的四分之一为一个对象。全区共划分出291个对象,其中I区有131个,Ⅱ区有160个(图1)。

为了划分不同震级的危险区,我们进行了两次识别。

第一次识别的震级下限 $M_1 = 6.0$ 级,其分类原则为:

(1) 以已发生过6.0级以上地震的地点为圆心,以35公里为半径划圆,在此圆内的对

象或部分区域在此圆内的对象为D类（危险类），紧邻此圆的对象为X类（未知类）。

（2）考虑到震级测定的误差，以 $5.5 \leq M_s < 6.0$ 地震的震中为圆心，35公里为半径划圆，在此圆内的对象为X类。除上述D类和X类外，其余的归为N类（安全类，即远离震中的对象）。

第二次识别的样本为第一次识别出的所有6.0级以上的地震的危险区。I区的识别震级 $M_1 = 6.5$ ，II区的识别震级 $M_1 = 7.0$ 。其分类原则为：

（1）将发生过 $M_s \geq M_1$ 地震的对象归为D类。

（2）将发生过 $6.0 \leq M_s < M_1$ 地震的对象归为N类。

其余的为X类。

两次识别时对象的分类及识别结果见表1。本文所用的地震资料截止到1987年底。根据最新考察结果我们增补了3个7级以上地震^[7,9]，这样43、44、151、152、153号对象则归为D类。

表1 识别前后各类对象的分布

	原始分类				识别结果							
	I区		II区		I区				II区			
	第1次	第2次	第1次	第2次	第1次	第2次	第1次	第2次	第1次	第2次	第1次	第2次
对象总数	131	56	160	87	总 131	新	总 56	新	总 160	新	总 87	新
D类对象数	31	15	57	15	42	21	21	9	71	24	31	13
N类对象数	70	16	83	40	89		35		89		56	
X类对象数	30	25	20	22								

2. 特征的选取

以每个对象的几何中心为圆心，分别以 $r_1 = 25\text{ km}$ 、 $r_2 = 35\text{ km}$ 为半径作圆，然后根据地质构造、地球物理特征、地形地貌、地形变和地震活动性等资料选取特征（见表2）。特征采用S编码^[1]。

四、潜在震源区的判定

1. K值的选取及示性特征的选定

本文采用正交试验法^[10]取16组K值对两区分别进行识别，根据不同K值下投票结果的误识率和示性特征的分布，分别选出各自的最优K值（表3）。

分别取表3中的K值对学习集进行识别。发现在对不同地区或不同震级的对象进行识别时，其示性特征不同：

（1）北部的中强地震越活跃，垂直形变速率和莫霍面起伏程度越大的地区，越易发生6.0级以上地震。该区内6.5级以上地震危险区除具有上述特点外，还具有在活动断裂交汇点和活动断裂拐点越多越易发震的特征。

（2）南部6.0级以上地震易发地点除具有与（1）相同的条件外，还具有活动断裂交

表 2 特 征 选 取 表

序 号	特 征 内 容	序 号	特 征 内 容
1	r ₁ 范围内的活动断裂交汇点数目	11	r ₂ 范围内莫霍面深度最大与最小值之差
2	r ₁ 范围内的活动断裂拐点数	12	r ₂ 范围内布格异常的最大值
3	r ₁ 范围内是否有新生代盆地边缘	13	r ₂ 范围内布格异常的最小值
4	r ₂ 范围内是否有新生代盆地边缘	14	r ₂ 范围内布格异常最大值与最小值之差
5	r ₁ 范围内是否有温泉分布	15	r ₂ 范围内最大垂直形变速率
6	r ₂ 范围内地形的最大高程	16	r ₂ 范围内最小垂直形变速率
7	r ₂ 范围内地形的最小高程	17	r ₂ 范围内垂直形变速率最大与最小值之差
8	r ₂ 范围内地形最大高程与最小高程之差	18	r ₂ 范围内6级以上地震的次数
9	r ₂ 范围内莫霍面深度最大值	19	r ₁ 范围内5级以上地震的次数
10	r ₂ 范围内莫霍面深度最小值	20	r ₂ 范围内是否有重力异常零值线

表 3 两次识别的最优K值及误识率

		阈 值				D类误识率	N类误识率	综合误识率
		K ₁₁	K ₁₂	K ₂₁	K ₂₂			
第一次识别	I 区	0.7	0.4	0.7	0.4	23%	16%	0.097
	II 区	0.6	0.3	0.6	0.2	14%	24%	0.107
第二次识别	I 区	0.6	0.3	0.6	0.3	20%	19%	0.097
	II 区	0.6	0.1	0.6	0.3	0	10%	0.053

汇点越多地形高程越高,有温泉分布的地区越易发展的特点。该区7.0级以上地震危险区的示性特征与上述诸特征有明显差别。如增加了重力布格异常、重力异常变异带、新生代盆地边缘等反应深部和边缘构造的特征。另外6级地震的多发区也是7级地震可能发生的地区。

2. 有关稳定性与危险概率的讨论

(1) 识别结果的稳定性

为了验证结果的稳定性,我们进行了 $N = \text{MAX}(N_D, N_N)$ 次删除对象试验。检验表明识别结果基本稳定。统计在N次实验中,某一对象被识别为D类的百分率Q,则在不同的Q值下各种误识率的分布见表4。图2是 ε -Q分布曲线。

图2显示出在第一次识别中两区均在 $Q = 100\%$ 时 ε 达到最低点,第二次识别的 ε_{min} 对应的Q值分别为75%(I区)和85%(II区)。应该说明的是第二次识别的样本是选取 $Q = 70\%$ 时第一次识别判定的所有危险对象,这是因为考虑到要尽可能降低D类误识率,最大限度地保证D类危险区不遗漏。

(2) 潜在震源区的危险概率

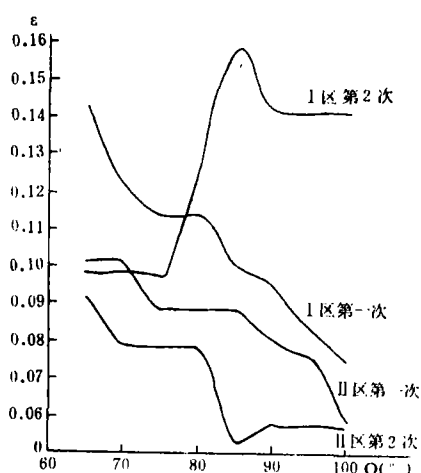
通过分析和研究,对于整体识别的可靠程度以及每一个危险对象的危险概率,我们给出以下定义。

1) 整体识别的可靠性 $P = (1 - \varepsilon/\varepsilon_{\text{max}}) \times 100\%$, 其中 ε_{max} 为该方法的 最大误识

表4

不同Q值下各种误识率的分布

Q (%)			100	95	90	85	80	75	70	65
区号										
误识率										
第一次识别	D类(%)	I	58	42	39	35	35	32	32	29
		II	17	17	17	17	17	17	17	17
	N类(%)	I	1	6	8	10	13	13	14	18
		II	8.4	13	14	17	17	17	20	20
	ϵ	I	0.0758	0.0824	0.0956	0.1003	0.1135	0.1135	0.1220	0.1435
		II	0.0588	0.0757	0.0799	0.0884	0.0884	0.0884	0.1011	0.1011
第二次识别	D类(%)	I	53	53	47	47	33	20	13	7
		II	7	7	7	0	0	0	0	0
	N类(%)	I	6	6	12	19	19	19	25	25
		II	10	10	10	10	15	15	15	18
	ϵ	I	0.141	0.141	0.142	0.159	0.128	0.097	0.098	0.098
		II	0.058	0.058	0.058	0.053	0.079	0.079	0.079	0.092

图2 ϵ -Q分布曲线Fig. 2 The ϵ -Q distribution curve

率, 为0.25^[11]。

2) 单个对象的危险概率 $P_i =$

$$\frac{(\Delta_i + \Delta_{max}) - 2\Delta_i}{2(\Delta_{max} - \Delta_i)}, \text{ 其中 } \Delta_i \text{ 为该对象}$$

在识别时得到的D类和N类示性特征的票数之差 ($\Delta_i = n_D - n_N$); Δ_{max} 为对象点获D类的全票与N类票0票时的票数之差; Δ_i 为投票的阈值。以上各参数均为N次试验的平均值。

当 $\Delta_i = \Delta_{max}$ (即 $n_D = n_{Dmax}$, $n_N = 0$)

时, $P_i = 1$;

当 $\Delta_i = \Delta_i$ 时, $P_i = 0.5$ 。

3. 潜在震源区的划分

根据以上两个定义, 我们计算出两次识别判定出的危险对象的危险概率和整体识别

的可靠性。表5给出的是新判定的危险区的分布和危险概率。6级以上地震潜在震源区的划分见图3 (取 $P_i \geq 0.4$ 的对象), 7级以上地震 (北部为6.5级以上) 的潜在震源区见图4 (取 $P_i \geq 0.50$ 的对象)。

我们用加权的Hamming方法对第一次识别结果做了验证, 选用 $Q = 100\%$ 时的结果比较, 两者稍有差异。

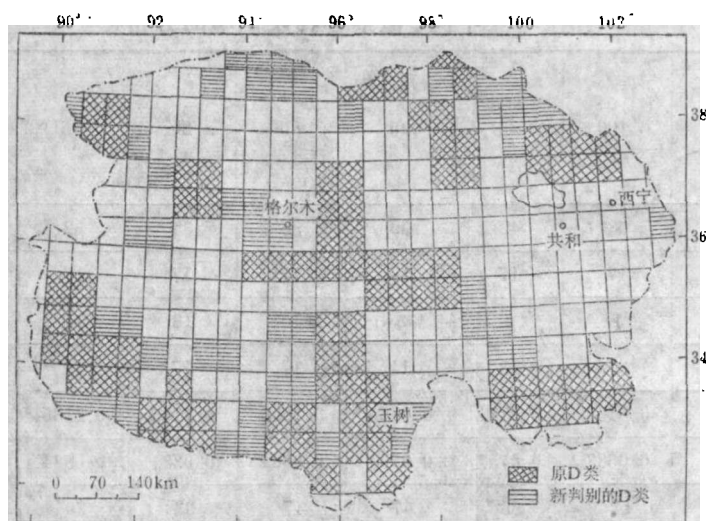


图3 用修改的CORA—3方法判定的6级以上地震潜在震源区

Fig. 3 $M \geq 6.0$ potential earthquake source zones divided by the revised CORA—3 method

五、问题与讨论

1. 本文选取资料的截止日期为1987年底。在此之后青海省内分别发生了1988年11月5日唐古拉6.8级和1990年1月14日茫崖6.7级两次强震。这两次地震均发生在本文所圈定的危险区内(233、61号对象), 前地震在原D类区, 后地震在新判别的危险区内。

2. I、II区分属不同的构造单元, 对于相同震级的地震, 其影响因素存在着明显的差别。在判定6.0级以上地震危险区时, I区显示的影响因素为地形高程、活动断裂交汇点、温泉分布等。而II区在识别震级提高到6.5级时才出现上述特点。另外, II区识别结果明显优于I区, 其影响发展的因素比

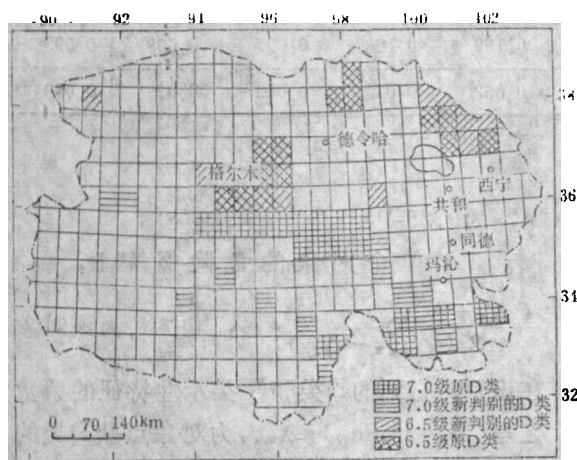


图4 用修改的CORA—3方法判定的6.5级(I区)、7.0级(II区)以上地震潜在震源区

Fig. 4 $M \geq 6.5$, $M \geq 7.0$ potential earthquake source zones divided by the revised CORA—3 method

I区多而突出。由此可以认为I区6级地震的发生随机性较大。

3. 同一地区、不同震级的地震危险区的条件也不同。震级越高, 一些反映深部构造和较大规模构造运动的特征越突出, 识别结果越好。

值得注意的是垂直形变速率在不同地区、不同震级的识别中均起着重要的作用。

4. 与其它地区的研究结果比较, 地形高差、莫霍面深度的变化和垂直形变速率是青海地区地震危险区的突出特征。这说明青藏高原北部地区, 因板块、块体之间相互碰撞、挤压造成的掀斜运动及垂直差异运动是强震发生的重要条件。

脾脏危险及分布的危险因素

I 对 象 区 号	6.0级以上 P=54.6%		6.5级以上 P=61.2%		I 对 象 区 号	6.0级以上 P=54.6%		6.5级以上 P=61.2%		I 对 象 区 号	6.0级以上 P=54.6%		6.5级以上 P=61.2%	
	Q(%)	P _i	Q(%)	P _i		Q(%)	P _i	Q(%)	P _i		Q(%)	P _i	Q(%)	P _i
2	95	0.56			22		0.64			84		100	0.66	0.68
4	90	0.77	90	0.60	23		0.66			103		90	0.47	100
9	100	0.97			29		0.74			104		70	0.42	
10	95	0.66			39				90	111		95	0.64	0.83
11	95	0.66			41		0.77			119				0.62
13	95	0.77			45		0.58			126				0.62
14	80	0.59			61		0.43			127		80	0.43	90
16	70	0.49			42				85	128		70	0.42	
21	95	0.46	100	0.74	66				100					

I 对 象 区 号	6.0级以上 P=69.7%		7.0级以上 P=78.8%		I 对 象 区 号	6.0级以上 P=69.7%		7.0级以上 P=78.8%						
	Q(%)	P _i	Q(%)	P _i		Q(%)	P _i	Q(%)	P _i					
132			90	0.52	217		0.49	100	0.94	259			100	0.89
133	100	0.57	90	0.52	224			100	0.72	262		95	0.53	
170	70	0.47	100	0.69	228		0.50	90	0.58	268		100	0.82	
193	80	0.50			229		0.50			273		100	0.90	
194	70	0.47	100	0.72	231		0.58			274		100	0.90	
201	95	0.55			239			100	0.97	277		95	0.57	
202	100	0.96	100	0.87	243			100	0.69	280		100	0.84	
208	100	1.00			246		0.61	100		283		100	0.77	
209	100	0.61			247		0.52	90		284		100	0.77	
216	70	0.47	100	0.86	256		0.73	100						

另外, 该区的地震危险区与活动断裂带的数目、端点无关。这可能与该区断裂密集、野外地质工作不详细有关。

5. 该方法判定的潜在震源区与青海地区主要活动构造比较吻合。

本项工作是青海省第三代烈度区划图编制的基础工作之一。在计算过程中国家地震局地球物理研究所的马秀芳同志给予了大量的指导和帮助, 青海省地震局的张晓东同志对本文也提出了一些有益的建议, 在此一并致谢。

(本文1990年6月20日收到)

参 考 文 献

- [1] 马秀芳、王玉秀、王碧泉, 几种有监督的模式识别方法及其程序包, 中国科学技术出版社, 1986.
- [2] 杨锦英, 中国西南地区地震活动的模式识别结果, 地震, No. 4, 1984.
- [3] 马文静, 祁连山地震带地震活动的图象识别, 内陆地震, Vol. 1, No. 4, 1987.
- [4] 马秀芳等, 用模式识别方法确定北京及邻近地区易发展的地点, 地震学报, Vol. 7, No. 4, 1984.
- [5] 马秀芳等, CORA—3 修改算法在北京及邻区易震地点识别中的应用, 地震研究, 1986.
- [6] 王碧泉等, 用模式识别方法综合分析多项震兆, 地震研究, Vol. 9, No. 6, 1986.
- [7] 曾秋生, 青藏高原的地震活动与地震构造, 高原地震, Vol. 1, No. 1, 1989.
- [8] 涂德龙, 阿尔金断裂中段构造新活动与地震相关性研究, 高原地震, Vol. 1, No. 1, 1989.
- [9] 曾秋生等, 东西大滩地震形变带的发现与研究, 高原地震, Vol. 1, No. 1, 1989.
- [10] 中国科学院数学研究所运筹室优选小组, 优选法, 科学出版社, 1975.
- [11] 王碧泉、陈祖荫, 模式识别理论方法和应用, 地震出版社, 1989.

DIVIDING THE POTENTIAL EARTHQUAKE SOURCE ZONES IN QINGHAI AREA BY THE METHODS OF PATTERN RECOGNITION

Ma Wenjing, Zou Wenwei

(Seismological Bureau of Qinghai Province, Xining, China)

Abstract

In this paper, by using the revised CORA—3 and weighted Hamming methods of pattern recognition, based on the analysis and study of geological, geomorphological, geophysical and seismological data in Qinghai area, the M6.0, M6.5 and M7.0 potential earthquake source zones are divided respectively in this area. The confidence level for whole recognition and danger probability for each potential earthquake source zone are given quantitatively. The M6.8 Tanggula earthquake on Nov. 5, 1988 and the M6.7 Mangya earthquake on Jan. 14, 1990 are all within the potential earthquake source zones divided by this paper.