

水库诱发地震危险性预测的模式识别方法初探

薛军蓉 李蓉川 韩晓光

(国家地震局地震研究所)

提 要 本文阐明了模式识别方法的原理及其在水库诱发地震危险性预测研究中应用的可行性和实用性。以国内22个水库为样本(其中11个是有震样本),提取水库诱发地震有关的地质、地震、水文等方面的因素及特征,对拟建长江三峡三斗坪和清江隔河岩高坝水库进行预测。初步判定这两个水库区存在着水库诱发地震的危险性背景,并且清江隔河岩高坝水库区都镇湾以西地段,不存在诱发地震的背景,而其以东库,存在着水库诱发地震的背景。

关键词 库水诱发地震 危险性预测 模式识别方法与标志 判定准则

水库诱发地震与水库工程的兴建直接有关。随着水库诱发地震特征、成因、环境条件和诱发机制及力学模型等方面研究的深入,水库诱发地震危险性预测已成为兴建高坝水库库坝安全必需虑及的问题。

水库诱发地震的预测和一般构造性地震的预测有其共性一面,但也有其差别。水库诱发地震的可能发震地点和发震时间受到水库库区和蓄水时间过程的制约,因此,在其危险性预测时的研究重点主要解决的问题是有无发生水库诱发地震的可能性,如果有,那么其强度可能有多大。

近年来不少作者利用构造类比、两极模糊综合判别模式和逻辑信息法等方法开展过水库诱发地震危险性预测的研究,对一些拟建的高坝水库提出了有益的预测意见^{〔1-4〕}。本文将在这些工作的基础上,采用判别分类的费歇尔判别准则和模式识别中的COR—3算法,探讨在水库诱发地震危险性预测研究中应用的可行性和实用性,并初步对拟建或在建的长江三峡三斗坪和清江隔河岩高坝水库进行预测。

一、水库诱发地震识别标志及特征选取

水库诱发地震的孕育及发生是一个十分复杂的过程,外部环境的影响和内在因素的作用,至今尚未全部认识,本文以水库诱发地震为研究对象(样本)的特征识别主要是众多研究中较为一致的影响因素,即与水库诱发地震有关的地质、地震、水文等方面的因素,并考虑到这些因素获取的可行性。

文献〔4〕给出了我国现已认为属于水库诱发地震的有震震例和尚未发生水库诱发地震的无震震例(表1)。并给出了水库诱发地震的特征识别标志,各特征的定义见表2。本文直接利用这些特征标志,对各水库给出的原始特征参数进行了预处理和判别,取“0”表示特征不存在或在该参数平均值以下,将特征标志存在并且在均值以上的,取值为“1”。

表 1 有震震例与无震震例样本

Tab. 1 Examples of earthquakes areas and non—earthquakes areas

(I) 类, 有震样本		(II) 类, 无震样本	
$0 < M < 4.7$	$M \geq 4.7$	$M = 0$	
1. 南 冲	9. 丹 江	1. 毛 家 村	7. 大 沙 河
2. 邓 家 桥	10. 新 丰 江	2. 汤 河	8. 高 州
3. 乌 江 渡	11. 参 窝	3. 刘 家 峡	9. 鹤 地
4. 乌 溪 江		4. 陆 水	10. 松 涛
5. 黄 石		5. 狮 子 山	11. 兴 安 江
6. 南 水		6. 拓 溪	
7. 前 进			
8. 柘 林			

二、水库诱发地震综合模式原理

水库诱发地震与多种因素有关, 现有研究表明, 各种因素与每次地震不是一一对应的。

表 2 特 征 标 志

Tab. 2 Characteristic signs

标志序号	特 征 定 义
1	地震活动背景值 (以坝址为圆心, 100Km 为半径范围内地震能量累积值)
2	水库区主要断裂方位角
3	水库及水域外 25km 范围内 10km 以上断裂总长度
4	水库及水域外 25km 范围内 10km 以上断裂交点数目
5	白垩—第三系盆地边缘的影响系数 (坝址到盆地边缘的最小距离与 25km 之比值)
6	新生代盆地边缘的影响系数 (坝址到盆地边缘的最小距离与 25km 之比值)
7	三级以上断块的影响系数 (坝址到断块边界的最小距离与 25km 之比值)
8	三级以上断块角顶的影响系数 (坝址到角顶的距离与 25km 之比值)
9	三级以上隆起与凹陷过渡带的影响系数 (坝址到过渡带的最小距离与 25km 之比值)
10	水库水域范围内碳酸盐类及花岗岩类岩石出露面积与该范围总面积之比值
11	水库及水域外 25km 范围内碳酸盐类及花岗岩类岩石出露面积与该范围总面积之比值
12	水库 100km 范围内新生代以来活动大断裂总长度
13	水库 100km 范围内第三纪以来活动大断裂总长度
14	水库 100km 范围内第四纪以来活动大断裂总长度
15	水库及水域外 25km 范围内 10km 以上正断层总长度
16	水库及水域外 25km 范围内 10km 以上逆断层总长度
17	水库及水域外 25km 范围内 10km 以上平推断层总长度
18	水库坝高
19	水库库容

需寻求合理的方法,进行特征检验,综合多方面因素对水库未来的发震背景作出最佳判定。

1. 各特征作用检验与初始分类

在各特征检验过程中,对表2中选取的19个特征标志进行分类判别,用“贡献系数”进行检验,考察各特征所起作用的大小。对“贡献系数”较小的特征,说明其作用不大,可以删去。根据计算结果,删去了表2中的X₁、X₅、X₁₃、X₁₄、X₁₆、X₁₇等六项。

对于两类样品各特征的初始分类,即训练情况,利用统计量F进行检验,如通过F检验,说明初始分类合理。

2. COR A—3 方法

本文水库诱发地震进行特征提取和分类采用COR A—3方法和最大隶属原则。将已发生地震的水库作为I类样品(D类或危险类),将未发生地震的水库作为II类(N类或安全类)。

从数学上看,决策判别可采用模糊数学或模式识别方法。这两类问题的线性判别函数的一般形式及决策规划为:

$$g(X) = W \cdot X \begin{cases} > RO \\ < RO \end{cases} \quad X \in \begin{cases} W_1 \\ W_2 \end{cases} \quad (1)$$

其中W称为权向量,RO为阈值。令X和W为d维向量,即

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} \quad W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_d \end{bmatrix} \quad (2)$$

方程 $g(X) = RO$ 定义了一个决策面,它是一个超平面,将特征空间分成两个半空间,使得W₁类的点和W₂类的点分别隶属于不同空间,即W₁的决策域为R₁,W₂的决策域为R₂。当X在R₁一侧时属于W₁类,当X在R₂一侧时属W₂类。这样,识别分类问题取放于所确定的 \bar{W} 和RO。

公式(2)中d个特征向量的回答采用二进制描述。如果有P个样本(水库区),则对(1)式的回答构成了P行d列,其元素为0或1的系数矩阵。然后按全组合原则将原向量集构成一维、二维、三维新向量集,其系数由原系数矩阵逻辑构成。随后计算新向量在I、II类区中的频度 N_j^I , N_j^{II} (j为所有新向量之总个数)。根据预先选定的阈值K₁、K₂、K₃、K₄来确定某个新向量能否作为I类或II类的特征。如果新向量X之 $N_j^I \geq K_1$ 、且 $N_j^{II} \leq K_3$,则X被确定为I类的一个特征。若新向量X_i之 $N_1^I \leq K_2$ 、且 $N_1^{II} \geq K_4$ 则X_i被确定为II类的特征之一。

选定了不同的K(i=1,2,3,4)后就可得到能描述I类、II类地区的可能性程度。

3. Fisher判别函数的分类与预测

依上法选定两类地区的特征向量后,就构成了如(1)式的线性判别函数,再按Fisher准则求出 \bar{W} 及RO

假设有一集合包括x₁、x₂、……、x_p共p个d维样本,其中p₁个属于w₁类的样本记为

子集 x_1 , p_2 个属于 w_2 类的样本记为子集 x_2 。如对 X 进行线性变换, 得到相应于 P 个样本的集合, 即:

$$Y = \bar{W}^T X = \frac{1}{P_i} \sum_{X \in X_i} \bar{W}^T X \quad (i = 1, 2) \quad (3)$$

Fisher准则函数定义为:

$$J(\bar{W}) = \frac{(\bar{m}_1 - \bar{m}_2)^2}{S_1^2 + S_2^2} \quad (4)$$

式中 m_i 是线性变换后各类样本之均值, \bar{S}_i^2 是线性变换后各类样本类内离散度。线性变换的目的是使同类的点尽可能集中, 而使不同的点相对离开, 即使 $J(\bar{W})$ 的分子尽可能地大, 分母尽可能地小。通过求极值可得到使 $J(\bar{W})$ 为极值的 \bar{W} 之估值 $\hat{\bar{W}}$:

$$\hat{\bar{W}} = S^{-1} \bar{W} (m_1 - m_2) \quad (5)$$

$$\text{其中 } S_{\bar{W}} = S_1 + S_2 = \sum_{X \in x_i} (X - m_i)(X - m_i)^T \quad (i = 1, 2) \quad (6)$$

$$m_i = \frac{1}{P_i} \sum_{X \in x_i} X \quad (i = 1, 2) \quad (7)$$

求得 $\hat{\bar{W}}$ 后, 可计算两类的临界判别值 R_0 及各类判别值 R_1 、 R_2 , 即:

$$\begin{cases} R_0 = (\sum_{X \in x_1} X + \sum_{X \in x_2} X) \hat{\bar{W}} / (P_1 + P_2) \\ R_1 = \sum_{X \in x_1} X \hat{\bar{W}} / P_1 \\ R_2 = \sum_{X \in x_2} X \hat{\bar{W}} / P_2 \end{cases}$$

如果预测判别值 $R_0 \leq R \leq R_1$ 则属 I 类区。如果 $R_0 \leq R \leq R_2$ 即属于 II 类区

三、拟建长江三峡三斗坪和清江隔河岩高坝水库诱发地震危险性预测

在水库诱发地震危险性预测中, 我们根据水库诱发地震震例资料的具体情况, 选取表 1 中给出的样本, 进行模式识别。

在拟建长江三峡三斗坪和在建清江隔河岩高坝水库的库区及水域外 25 公里范围内, 根据上述选取的 13 个特征标志值, 进行模式识别。取不同值的 K_i , 以不同的特征标志或特征标志组合进行分类, 作为区分 I、II 类地区的特征。在采用不同的 K_i 值进行识别后, 取一组稳定的, 按 Fisher 准则判别分类结果, 得到显著的 I 类区特征列于表 3。并得到 I、II 类地区的判别值 R_1 、 R_2 及临界判别值 R_0 如下:

$$R_1 = 17.0 \quad R_2 = 5.4 \quad R_0 = 10.9$$

同时得到 I 类地区的内检误识率为 0.02; II 类区的内检误识率为 0.05。

表3 $K_i=0.85, 0.50, 0.50, 0.85$ 时选定I类区的特征Tab. 3 Characters of No. 1 area while $K_i=0.85, 0.50, 0.50$ and 0.85

标志号	一维	二 维 组 合							二维组合			
2	1											
3		1			1				1	1	1	
4			1									
6				1								
7												1
8						1						
9	1	1					1		1	1		
10	1		1					1			1	1
11						1						
12					1							
15	1						1	1		1	1	1
18				1								
19									1			

对拟建的三峡水库与清江隔河岩高坝水库预测判别值 R 分别为11.5和17.0, 界于I、II类地区的判别值 R_1 、 R_2 之间。可以认为这两个预测水库均属于I类区, 即存在着水库诱发地震的危险性背景。为进一步判定清江隔河岩高坝水库存在水库诱发地震的危险地段, 以都镇湾为界将库区分为东段与西段分别进行预测。得到都镇湾以西地段 $R=0$, 属II类区, 不存在诱发地震的背景, 而其以东库段 $R=17.0$, 属于I类区, 存在着水库诱发地震的背景。

四、结 论

表3可以看出, 特征标志 x_2 、 x_7 、 x_8 、 x_{11} 项为I类地区分类的主要特征, 即库区及水域外25公里范围内10公里以上断层的交点数以及正断层总长度, 库区是否处在地壳差异运动表现最明显的隆起与凹陷过渡地带, 以及水库水域范围内碳酸盐类岩石及花岗岩类岩石出露面积, 是水库诱发地震危险性的四种主要特征标志, 它们分别对水库地震的诱发有着大小不同的贡献。其它各特征标志及其组合的作用不明显。

利用模式识别法对拟建的三峡水库与清江隔河岩高坝水库预测结果, 与文献〔4〕¹⁾ 逻辑信息法等对这两个大型水库的预测结果基本相同, 进一步为这两个水库诱发地震危险性判别提供了依据。

1) 湖北省地震局, 11988, 清江隔河岩水利枢纽工程地震综合研究报告。

模式识别方法能根据问题的物理本质进行综合, 这是其优特的一面, 但是同样需要较丰富的样本资料。由于以水库样本资料来源有限, 尤其是已诱发了地震的水库数量更少, 使得取样受到了一定的限制。有关特征标志的提取, 由于水库诱发地震的孕育和发生是个十分复杂的过程, 受到许多因素的影响和作用, 而且不少因素至今尚未被认识。本文研究仅为初步探索, 随着研究的深入, 积累更多的水库地震活动资料, 模式识别方法预测效果将会更好。

本文工作得到严尊国、于品清同志的帮助, 在此谨向他们表示感谢。

参 考 文 献

- 〔1〕 胡毓良, 1979, 我国水库地震及有关成因问题的讨论, 地震地质, (4)。
- 〔2〕 常宝琦, 水库诱发地震的预测, 华南地震, 4, (4)。
- 〔3〕 国家地震局地震研究所, 1984, 中国诱发地震, 地震出版社。
- 〔4〕 余永毓, 1987, 逻辑信息法在水库诱发地震研究中的应用, 地震学报, 9, (增刊)。
- 〔5〕 张郅珍等, 1985, 华北近期地震危险区的确定要素和综合图象识别, 中国地震, (4)。
- 〔6〕 王碧泉等, 1989, 模式识别理论、方法和应用, 中国地震出版社。

RESEARCH ON A METHOD OF HOW TO DISTINGUISH THE MODEL OF FORECASTING THE RISK OF RESERVOIR INDUCED EARTHQUAKE

Xue Junrong, Li Rongchuang and Han Xiaoguang

(Seismological Institute of State Seismological Bureau)

[Abstract] This paper describes the principle of how to distinguish the model and its feasibility and practicability in forecasting the risk of reservoir induced earthquake. Taking 22 reservoirs (among them, earthquakes occur in 11 reservoirs) as an example and picking up the factors and characters of geology, earthquake and hydrology which relate to the reservoir induced earthquake, we have forecasted the reservoirs which will be built in Sandouping and Qinggeheyan area of Three Gorge in the Yangtze River. It judges preliminarily that the risk of reservoir induced earthquakes are existed in these two reservoir areas, but induced earthquakes do not occur at the west of Dozhenwan region in Qinggeheyan reservoir area, they may occur in the east of this area.

[Key words] Reservoir induced earthquake; Anticipation of earthquake risk; Methods of how to distinguish the model; Distinguishing criterion