

PSHA 模型的算法改进与中国地区未来地震概率评估

路志越¹ F. Benjamin Zhan² 鄂栋臣¹

(1 武汉大学中国南极测绘研究中心,武汉市珞喻路 129 号,430079)

(2 德克萨斯州立大学圣马克斯分校地理系, TX 78666)

摘要:考虑到 PSHA 模型对未来地震预测特别是大震预测能起到很好的参考分析作用,尝试结合能量分布模型对其进行了改进,并以中国及周边地区 200 a 内的地震历史记录为基础,利用原始 PSHA 模型以及改进后的 PSHA 模型分别计算了中国及邻近地区 2000~2010 年间 5 级以上地震发生的概率,并对两种结果进行了比较分析,最终证明改进后的 PSHA 模型具有更好的可靠性。

关键词:地震概率评估;大震预测;改进的 PSHA 模型

中图分类号:P315.7

在众多地震预测研究中,大部分学者都侧重于地震预报算法的研究,而往往忽视了根据地震历史记录对特定区域进行未来地震风险评估的工作^[1]。未来地震风险评估的研究最早由 Cornell (1968)开始利用历史数据记录的方法来推测震源区域的可能分布。后来诸多利用历史数据集进行地震灾害研究的工作大都基于 Cornell 的研究基础之上^[2]。美国学者 Frankel (1995)利用 PSHA (probabilistic seismic hazard assessment) 模型并结合高斯平滑算子,对美国中、东部地区进行了危害估算。此后,PSHA 模型成为震率评估的主要方法之一。自 1996 年起,美国 USGS 的国家地震灾害地图应用开始通过量化可视化的方式将美国全境包括阿拉斯加地区未来地震的可能性利用概率分布地图完全呈现出来,其核心概率预测算法依然是 PSHA 模型。

我国地震研究大多集中在地震预报算法以及地震动相关参数影响的研究等方面,而没有尝试对整个模型进行优化。如沈建文等人对 Gutenberg-Richter 震级频率分布关系中反映地震活动特性的 b 值参数进行了详细探讨和研究;秦长源等则侧重研究了震级误差在 Gutenberg-Richter 关系中的影响^[3]。为此,本文利用 1827~2000 年的地震历史记录,分别采用原始 PSHA 模型和改

进后的 PSHA 模型对中国地区 2000~2010 年间的地震概率进行了评估和比较分析。

1 地震数据预处理

利用 PSHA 模型进行未来地震概率的估算是以地震事件服从泊松分布为假设前提的,即

$$Pr = 1 - e^{-N(m)T} \quad (1)$$

其中, Pr 为未来 T 年内震级大于 m 的概率值; T 值可根据需求自行定义; $N(m)$ 为地震震级大于 m 的年震率。对于未来地震概率 Pr 的估算,其核心在于地震震率 $N(m)$ 的计算,而地震震率计算的可靠性又依赖于长时间可靠的地震历史记录。本文数据来自于 ANSS 和中国国家地震台网,数据范围为纬度 $20^\circ \sim 50^\circ$, 经度为 $72^\circ \sim 130^\circ$, 数据集为 1827~2010 年间的地震记录。该区域地震记录主要包括中国大部分地区以及中国边界外相邻地区的地震记录数据,共有 14 798 条。

在使用前,需要对地震目录进行相关处理。考虑到震率计算的可靠性,首先必须进行地震数据的完整性阈值确定。由于震率的计算以服从泊松分布的假设为前提,因此还必须进行地震数据的过滤(提取主震记录),以确保地震事件的独立性。然后进行震级误差的修正,消除震级误差在

计算中产生的影响。

1.1 地震记录的完整性阈值计算

中国早期由于地震测站数目较少,较小震级的地震震波能量较小,无法远距离传递,因此小地震的发生常常被忽略,造成地震记录不完整。近年来,地震观测网密集地增设测站,使小地震的侦测能力得到显著提高^[3],因此地震记录资料与早期相比,记录更为完整、可靠。在连同历史资料进行分析时,必须考虑到早期资料的不完整性,进行完整性阈值计算,再进行地震分析。

根据 Gutenberg-Richter 震级频率分布关系(下文统称 G-R 关系),地震记录频率的对数值应与震级呈线性关系,随着震级的升高而线性下降。据此,本文根据 1970~2000 年间的的海震数据记录绘出震级频率分布直方图,从插页 II 彩图 1 可以看出,震级大于 4.2 级地震数据的完整性较好,因而本文选取完整性阈值为 4.2。由于早期监测手段的局限,1827~1969 年间的的海震数据记录存在严重缺失的情况,这部分记录将在实际计算中利用统一归算的方法参与计算。在不同的研究区域中,由于仪器以及各种人为因素的影响,可能造成完整性程度的差异^[2]。因而在特殊区域中,根据需要可以单独计算完整性阈值。

1.2 主震记录提取

在地震记录中,通常会包含前震、主震以及余震记录。由于地震概率评估要求地震时间具有独立性,因而前震和余震记录必须被有效剔除。本文采用由 Davis 和 Frolich 建议的单连接聚类分析方法^[4],他们采用如方程(2)所示的原则定义时空关系窗口,从而进行主震提取:

$$d_{ST} = (d^2 + C^2 T^2)^{1/2} \quad (2)$$

其中, d_{ST} 为时空窗口阈值; d 为两个地震事件的空间距离; T 为两次地震时间的时间间隔(单位为d); C 为时间与距离的关联常数。通过不同地区数据的实验分析,Davis 和 Frolich 建议 C 值为 1 km/d,亚洲地区 d 值取值为 87.5^[4]。作为理想阈值进行主震分离,最终将 1826~2010 年的地震记录分离得到 $M_s > 4.2$ 的独立主震记录为 7 116 条,其中震级大于 4.2 小于 6.0 级的地震记录为 6 431 条,大于 6.0 级的地震记录为 685 条。利用 GMT 作出地震事件分布图如插页 II 彩图 2 所示,其中,小圆代表震级在 4.2~6.0 之间的地震;大圆代表震级高于 6.0 的地震。

1.3 震级误差修正

在利用 G-R 关系计算震率前,首先要考虑不同测站间精度不同所产生的震级记录误差,该误

差的大小会影响到 G-R 震级频率关系中参数 a 值的可靠性。根据 Felzer 的建议,震级误差可以通过将记录震级减去一个 ΔM 进行修正^[5]:

$$\Delta M = b^2 \sigma^2 / 2 \lg(e) \quad (3)$$

其中, σ 为震级误差标准差; b 为 G-R 关系中的 b 值。 ΔM 需在 b 值估算完成后进行计算。

2 改进后 PSHA 模型的震率计算

本文在原有的 PSHA 模型的基础上,结合地震能量释放模型对其进行了改进,最终得出更为可靠的地震率估算结果。

2.1 G-R 震级频率分布关系

G-R 震级频率分布关系是一种目前公认的能够较好描述地震活动区域中震级与频率分布情况的一种统计学关系^[5]。在实际应用中,G-R 关系通常采用如下公式表示:

$$v(m \geq M) = 10^{(a-bM)} \frac{1 - 10^{-b(M_{\max} - M)}}{1 - 10^{-b(M_{\max} - M_0)}} \quad (4)$$

其中, M_0 为地震震级阈值; M_{\max} 为地震震级上限; M_{\min} 为地震震级下限; M 为指定地震震级; v 表示地震震级大于 M 的震率; a 和 b 分别为 G-R 关系中的常数,其中 a 值显示指定区域地震活动的频繁程度,是一个时间空间相关变化量; b 值主要反映该地区的地质构造特性,其全球平均值为 1.0。

在处理地震数据记录方法中,对于地震次数的使用,分为间隔地震次数与累计地震次数两种方式来分析 b 值^[1]。一般采用后者进行分析,这种方式如同积分运算,可以降低局部震荡现象,将地震资料曲线平滑化。

2.2 b 值计算

b 值是 G-R 关系中的核心参数,其计算一般有最小平方和极大似然估计法,本文采用极大似然估计方法对 b 值进行估计。首先将 G-R 关系以自然对数表示为:

$$\ln N(M) = \alpha - \beta M \quad (5)$$

其中, $\alpha = a \ln(10)$; $\beta = b \ln(10)$ 。利用极大似然估计方法,最终通过迭代方法可以计算出 β 值。迭代极限为:

$$|(\beta_i - \beta_{i-1}) / \beta_{i-1}| > \epsilon \quad (6)$$

本文设定 $\epsilon = 10^{-4}$,最终根据 β 值计算出 b 值。

根据中国地区震源的分布特点,一般分为 4 个震区分别计算 b 值^[6],计算结果如表 1 所示。中国周边地区的 b 值统一采用全球平均值 1.0。

2.3 震率计算模型

原始 PSHA 模型只包含三个形式相似的子

表 1 不同地区 b 值及相应标准差Tab. 1 b -value and the Stand Deviation in Four Parts of China

区域	b 值	标准差	区域	b 值	标准差
华北	0.712	0.192	西北	0.707	0.189
华南	0.802	0.168	西南	0.864	0.191

模型,各个子模型都以假设地震发生服从泊松分布为前提,分别考虑到了不同的影响因素,基于不同的时空划分原则对 G-R 关系中的参数 a 进行计算,然后根据式(4)求出大于该震级的震率,最后将计算出的震率值分别采用不同尺度下的空间平滑方式进行平滑^[4]。本文在其基础上以地震能量分布模型作为附加模型来加权计算特定区域的震率值,并对平滑参数进行了重新设定。本文所附加的能量分布模型来自于 Lapajne 等在进行区域地震动峰值加速度评估的研究^[7],该能量模型充分考虑了大震效应,用以弥补 PSHA 模型本身震率估算的不足。

1) 子模型 1(标准模型)。这个模型包含了 1970~2010 年间震级范围在 4.2~7.0 之间的地震事件。计算这一期间的数据主要是假设未来的地震可能会集中发生在最近数十年内地震频发的区域。根据 G-R 关系,一个在过去一段时间内小规模地震集中发生的给定区域中,一定比例的较大规模地震发生次数的期望值会相对较大。在空间平滑时,模型 1 的相对距离选择为 17 km^[1]。

2) 子模型 2(中强震模型)。这个模型主要考虑计算 1862 年以后震级范围在 5~7 级之间的较大规模地震事件。模型主要以有较大规模地震发生过的区域地震复发的概率大于一般区域为前提假设,同时考虑到这段期间内数据的缺失,因而需要在统一计算过程中计算出一个统一因子。由于模型 1 的数据最为完整、准确,所以归一计算时,需以模型 1 为基准来计算因子,乘以该统一因子就可以确保总区域内大规模地震的年震率与模型 1 保持一致。模型 2 的平滑范围选择要大于模型 1,本文选为 25 km。

3) 子模型 3(背景模型)。模型 3 主要假设整个研究区域的地震服从统一分布,因而在平滑时是以整个研究区域为平滑范围的。由于计算性能所限,本文选择 500 km 作为平滑相对距离。其中数据选择为 1862 年以后 4.2 级以上的所有地震。其中 1862~1970 年间的地震记录同样需要进行统一归算。子模型 3 可以看作是震率计算的一致性背景震率区域,它考虑到了那些地震发生不频繁的区域,并对其未来发生地震的可能性作

出量化评价。

4) 模型 4(能量模型)。模型 4 是从地震总能量释放的角度来评估地震活动的。这个模型用来强化历史上大型地震的影响作用,其中能量释放与地震震级间的关系为 $\lg E_k = Am_k + B$,其中, A 和 B 为常数。则总释放能量为 $E = 10^B \sum_k 10^{Am_k}$ 。

将研究区域格网化,一个特定区域服从统一的 b 值,通过计算每个格网的地震释放能量,利用震率能量方程,可以得到研究区域的震率值。

能量模型主要体现大震事件对未来地震发生率的影响,因此平滑范围应当选择较大值,本文采用 100 km 进行平滑。

2.4 格网计算

本文对 4 种模型震率的计算结果分别利用高斯平滑算子进行平滑处理,然后将平滑结果进行格网化计算。将中国整个研究地区划分为 $0.1^\circ \times 0.1^\circ$ 的格网,对每个格网分别进行计算。PSHA 原始三种子模型的加权比值通常为 5 : 2.5 : 2.5 或者 6 : 2 : 2,本文在采用两种模型对比计算时,原始 PSHA 模型的权值比例选择为 5 : 2.5 : 2.5,改进后的 PSHA 模型采用 5 : 1.5 : 1.5 : 2 的权值比例将四种模型的结果进行加权求和。其中子模型 1 比例偏高,是由于该子模型所采用的近年地震记录最为准确、可靠。

3 计算结果

根据不同震级下所得的震率值 N 生成相应的格网文件,在 GMT 中重采样为 $0.05^\circ \times 0.05^\circ$,并根据泊松分布计算相应震级未来发生的概率值。通过生成格网文件,在服务器端利用 Linux 系统下的 GMT 自动重采样进行地图绘制。

为检验计算方法的可靠性,本文利用 2000 年 1 月 1 日以前的数据作为实验数据,用以估算中国范围内未来十年(2000~2010 年)5.0 级以上地震发生的概率,并以此结果与实际发生的地震情况作叠合比较,如插页 II 彩图 3(a)所示。可以看出,大部分地震事件都较好地落在概率超过 50% 的空间范围内。

为了将改进后算法的预测结果与原始算法进行比较评估,本文同样根据原始 PSHA 模型进行了同时空区域同震级的概率计算,计算结果如插页 II 彩图 3(b)所示。两图比较可以明显看出,在一些实际发生地震的区域,原始模型的计算结果明显小于改进后模型的计算结果。这是由于原始模型对大震效应的估计不足,从而低估了未来发

震的概率。通过插页 II 彩图 4 可以看出,根据原始模型的计算,只有不足 60% 实际发生地震区域的预测概率高于 50%,而改进后模型中,有超过 90% 的实际发生地震区域的预测概率高于 50%。可见,改进后模型的计算可靠性得到了显著提高。另外,由于改进模型考虑了大震效应的影响,所以在平滑过程中,与大震区域相连的区域地震概率也会相应提高。这与实际情况较为吻合,但也有可能出现对某一区域的概率评估较高,但实际并无地震发生的情况。由于被高估的地区往往都是与历史发生大震地区较为接近的区域,因而虽然在预测期间内并没有地震发生,其潜在风险仍然较大,所以这部分的概率高估问题应当在概率评估可以接受的范围之内。后续工作将会对概率高估区域的评估效用进行测试与评定。

4 结 语

中国地区由于自身地质结构的独特性,因而在评估地震潜在危害时所利用的地震模型参数等方面都需要通过实际历史记录严格测算,不能完全根据其他国家地区的经验值进行简单估算。特别是在计算中国区域 b 值时,若能够得到充分准确的各观测站的标准差,建议采用秦长源通过加权震级误差的修改方式修改震级,这种方法取得的最终效果较为明显。另外,由于青藏高原监测能力较弱,导致数据记录完整性较弱,因而可以考虑分区进行数据阈值估算。

参 考 文 献

- [1] Frankel A. Mapping Seismic Hazard in the Central and Eastern United States[J]. *Seismol Res Lett*, 1995,66(4): 8-21
- [2] Cornell C A. Engineering Seismic Risk Analysis [J]. *Bull Seismol Soc Am*, 1968,58:1 583-1 696
- [3] 秦长源. 地震震级误差对 b 值的影响[J]. *地震学报*, 2000, 22(4): 338-344
- [4] Davis S D, Frolich C. Single-link Cluster Analysis and Earthquake Aftershocks; Decay Laws and Regional Variations [J]. *J Geophys Res*, 1991, 96: 6 335-6 350
- [5] Felzer K R. Appendix I; Calculating California Seismicity Rates[R]. USGS Open File Report, Pasadena, California, 2007
- [6] 卢怡利. 1993 年至 2004 年嘉南地震之 b 值[D]. 台北:中国台湾国立中正大学,1994
- [7] Lapajne J K. Spatially Smoothed Seismicity Modeling of the Seismic Hazard in Slovenia[J]. *J Seismol*, 1997(1): 73-85
- [8] 汤皓,陈国兴. 基于 GIS 和神经网络模型的场地地震液化势风险评价[J]. *武汉大学学报·信息科学版*, 2007,32(8): 727-730
- [9] Fu Zhengxiang, Lu Xiaojian, Shao Huicheng, et al. Analysis on Statistical Characteristics of b -values of Aftershock Series in China Continent and Its Subregions[J]. *Earthquake*, 2008,28(3): 1-7

第一作者简介:路志越,博士生,主要从事 GIS 方面的研究。
E-mail: zhiyu283@hotmail.com

Probability Estimation of Future Earthquakes in China Based on Improved PSHA Model

LU Zhiyue¹ F. Benjamin ZHAN² E Dongchen¹

(1 Chinese Antarctic Center of Surveying and Mapping, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Department of Geography, Southwest Texas State University, San Marcos, TX 78666, USA)

Abstract: Considering the importance of PSHA model in earthquakes forecasting, by using the catalog in the 200 years in China, we tried to improve the original PSHA model with combining seismic energy distribution model, and computed the probability of earthquakes in a specific area in China with time-span T based on both the improved PSHA model and the original model. Finally the experimental results show that the improved model is more efficiency and reliable than the original one.

Key words: seismic rate estimation; earthquakes forecasting; improved PSHA model