

龙门山地区的地震活动性广义帕累托模型构建^{*}

任梦依

(中国地震局地球物理研究所, 北京 100081)

摘要: 利用 1930—2016 年龙门山地区 $M \geq 4.5$ 历史地震目录, 通过时空窗法减少余震对统计结果的影响, 构建了该地区广义帕累托分布的超出量分布模型, 估计了龙门山地区震级上限。结果表明: 广义帕累托分布较好地拟合了龙门山地区强震数据, 形状参数估计均为负值, 最大震级分布存在有限上界, 震级上限为 8.3。

关键词: 地震活动性; 广义帕累托分布; 震级上限; 龙门山地区

中图分类号: P315.5

文献标识码: A

文章编号: 1000-0666(2018)02-0226-07

0 引言

地震活动性模型的构建方法研究可为地震区划和地震危险性分析提供重要的技术支撑。对于地震活动性模型的改进方法研究, 一直是地震学领域的研究热点。概率方法是目前构建地震活动性模型中最常使用的分析方法(胥广银, 2003), 如基于截断 $G-R$ 关系模型的方法(Cornell, 1968)以及部分学者在此基础上提出的改进方法(徐伟进, 高孟潭, 2012), 美国现行的地震区划图以及第五代《中国地震动参数区划图》中均主要应用该方法构建地震活动性模型(潘华等, 2009; 潘华, 李金臣, 2016)。然而, 使用基于截断 $G-R$ 关系的方法, 需要先验地选取起算震级和设定震级上限。大量的实际地震资料分析也表明, 用 $G-R$ 关系模型检验地震震级分布情况时, 在许多情况下震级高端或低端会出现明显的“掉头”现象(胡聿贤, 1999)。由于实际资料对 $G-R$ 关系的偏离, 特别是强震级段的偏离, 一定程度上影响了利用该关系式推测未来强震发生危险性的可信度, 因此, 如何对上述关系进行改善, 特别是针对强震数据的拟合一直是地震学家研究关注的焦点。

Balkema 和 Haan (1974), Pickands (1975) 研究发现, 对于充分大的阈值, 随机变量超过阈

值超出量的极值分布为广义帕累托分布, 这为研究随机变量尾部分布特征提供了理论依据。近年来, 由极值模型发展而来的广义帕累托分布模型逐渐被应用到潜源区地震活动性模型的构建中, 如 Pisarenko 和 Sornette (2003) 利用广义帕累托分布分析了哈佛地震目录中 18 个地震区的浅层地震地震矩分布; Huyse 等 (2010) 利用太平洋地震工程研究中心的地面峰值加速度数据和残差数据比较了对数正态分布与广义帕累托分布的拟合优度; 钱小仕等 (2013) 利用广义帕累托分布分析了各活动地块边界带的强震震级分布特征; 田建伟等 (2017) 利用广义帕累托分布估计了马尼拉海沟俯冲带潜在地震海啸源区的地震危险性。该方法利用一段时限内潜源区的大震震级数据, 不仅无需先验地选定震级下限和震级上限, 而且可充分利用大震级区段的信息, 适合于历史地震记录时间长但低震级地震记录缺失的地区, 便于构建潜源区强震活动模型。

在构建广义帕累托模型前, 需充分了解、分析潜源区地震地质信息, 确定潜源区边界范围, 对于地震样本的选取也是十分重要的。另外, 在利用广义帕累托分布构建模型时, 需提前排除余震的影响, 尽量满足事件相互独立的条件。本文选取地震活动较为频繁的龙门山地区为研究区, 在分析该地区地震地质背景以及地震活动特征的基础上, 利用 1930—2016 年历史地震数据的大震

^{*} 收稿日期: 2018-01-15.

基金项目: 中国地震局地球物理研究所基本科研业务费专项《潜源区地震活动性广义极值模型构建方法研究——以龙门山地震带为例》(DQJB17B07) 资助.

级区段, 确定震级阈值, 构建了龙门山地区广义帕累托分布的超出量分布模型, 用于估计该区域的震级上限。

1 广义帕累托模型构建方法

1.1 广义帕累托分布

设 X_1, X_2, \dots, X_n 为地震震级随机变量列, 假设它们相互独立且服从同一分布, 记其最大震级为 $M_n = \max(X_1, X_2, \dots, X_n)$, 若存在 $\{a_n > 0, b_n \in R\}$ 和非退化分布函数 $G(x)$, 使得:

$$\Pr \left\{ \frac{M_n - b_n}{a_n} \leq x \right\} \rightarrow G(x), n \rightarrow \infty \quad (1)$$

则 G 称为极值分布。Fisher 和 Tippet (1928) 证明了极值分布的 3 大类型定理, 指出 G 必属于 Gumbel 分布、Frechet 分布、Weibull 分布之一。广义极值理论将这 3 种分布类型统一为 1 个分布类型, 由位置参数 μ 、尺度参数 σ 、形状参数 ξ 来表示 (Coles, 2001; 史道济, 2006), 如式 (2) 所示:

$$G(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (1 + \frac{\xi(x - \mu)}{\sigma} > 0) \quad (2)$$

选定震级阈值 u , 若 $X_i > u$, 称 $Y = X_i - u$ 为震级超出量。则超出量的分布函数可表示为:

$$F_u(y) = \Pr \{ X - u \leq Y | X > u \} = \frac{F(u+y) - F(u)}{1 - F(u)} (y > 0) \quad (3)$$

对于充分大的震级阈值 u , 震级超出量 Y 的分布函数近似服从广义帕累托分布:

$$H(y) = 1 - (1 + \frac{\xi y}{\tilde{\sigma}})^{-1/\xi} \quad (4)$$

式中: $y > 0$, $(1 + \frac{\xi y}{\tilde{\sigma}}) > 0$, 且 $\tilde{\sigma} = \sigma + \xi(u - \mu)$ 。

广义帕累托分布与广义极值分布具有相同的形状参数 ξ 。当 $\xi \geq 0$ 时, 广义帕累托分布均没有上边界; 只有当 $\xi < 0$ 时, 广义帕累托分布具有有限上界。

1.2 震级上限估计方法

震级上限是指潜在震源区内可能发生的最大地震的震级, 可以认为在潜在震源区内未来发生超过该震级地震的概率几乎为零 (胡聿贤, 1999)。当选定震级阈值 u , 且 $X > u$ 时, 则有:

$$\Pr \{ X > x | X > u \} = [1 + \xi \left(\frac{x - u}{\tilde{\sigma}} \right)]^{-1/\xi} \quad (5)$$

令 $\zeta_u = \Pr(X > u)$, 表示为超过阈值的样本个数与总样本的比值, 式 (5) 可写为:

$$\Pr \{ X > x \} = \zeta_u [1 + \xi \left(\frac{x - u}{\tilde{\sigma}} \right)]^{-1/\xi} \quad (6)$$

考虑到潜在震源区的地震事件的震级必为有限值, 因此只有当形状参数 $\xi < 0$ 时, 广义帕累托分布具有有限上界, 才符合震级存在上限的条件, 震级上限可表示为:

$$x = u - \tilde{\sigma} / \xi \quad (7)$$

利用最大似然估计方法可得到 3 个参数 ($\zeta_u, \tilde{\sigma}, \xi$) 的估计值。根据 Delta 定理可知, 3 个参数的最大似然估计值近似符合正态分布, 因此根据方差协方差矩阵可以获得置信水平为 $1 - \alpha$ 的震级上限的置信区间如式 (8) 所示:

$$\hat{x}_N \pm Z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(\hat{x}_N)} \quad (8)$$

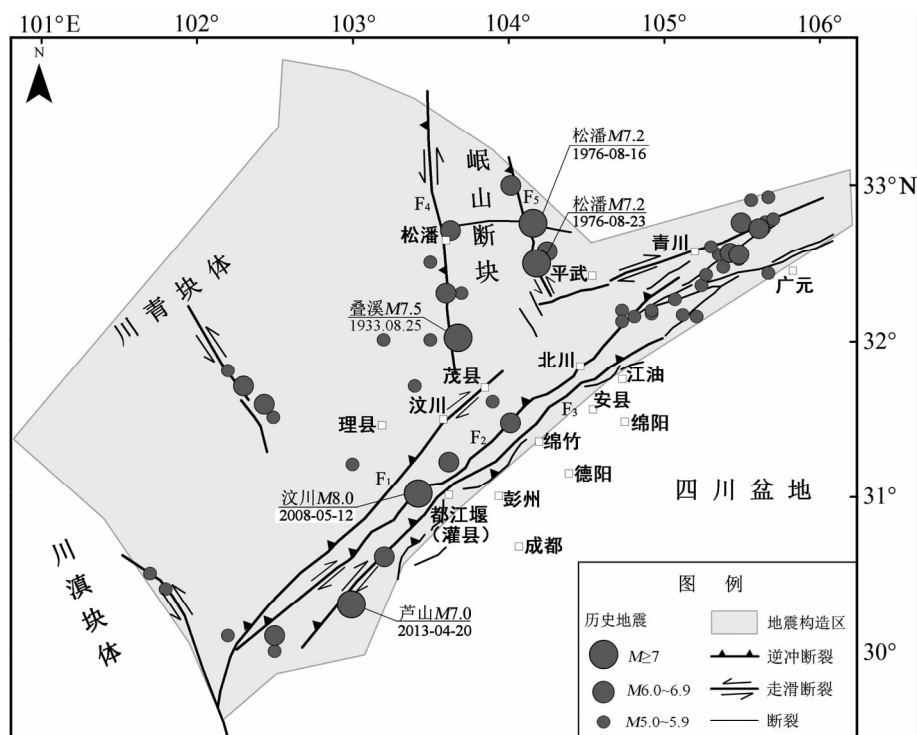
式中: $\text{Var}(\hat{x}_N) \approx \nabla x_N^T \mathbf{V} \nabla x_N$; \mathbf{V} 是 $(\hat{\zeta}_u, \hat{\tilde{\sigma}}, \hat{\xi})$ 参数估计的协方差矩阵; ∇x_N^T 为 $(\frac{\partial x_N}{\partial \zeta_u}, \frac{\partial x_N}{\partial \tilde{\sigma}}, \frac{\partial x_N}{\partial \xi})$ 在 $(\hat{\zeta}_u, \hat{\tilde{\sigma}}, \hat{\xi})$ 处的值; $Z_{1-\frac{\alpha}{2}}$ 为标准正态分布 $1 - \frac{\alpha}{2}$ 的分位数 (Coles, 2001)。

2 龙门山地区地质构造

第五代《中国地震动参数区划图》中将龙门山地区划分为地震构造区, 区内的地震活动具有相似的地震构造环境和发震构造模型, 因此, 本文选择龙门山地区为研究区。该区域块体滑移方向为南南东, 区域主压应力呈北西西—南西西向的水平挤压, 具有发生强震的构造与动力学环境。该区主要包括岷山断块和龙门山断裂带 (图 1), 为巴颜喀拉地块

的东边界,该边界不仅是强烈的活动构造带,也是我国南北地震带中段的一部分 (Densmore, 2007;

闻学泽等, 2011; 易桂喜等, 2011; 孔军, 周荣军, 2014; 高孟潭, 2015)。



F₁: 汶川—茂县断裂; F₂: 映秀—北川断裂; F₃: 安县—灌县断裂; F₄: 岷江断裂; F₅: 虎牙断裂

图1 龙门山地区地质构造及去余震后 $M \geq 5$ 地震震中分布图

(据李海兵等, 2009; 邓起东, 2007; 高孟潭, 2015 修改)

Fig. 1 Map of geological structures and distribution of $M \geq 5$ earthquakes after removing aftershocks in Longmenshan region (revised from Li *et al.*, 2009; Deng, 2007; Gao, 2015)

岷山断块东、西边界分别为虎牙断裂和岷江断裂。虎牙断裂主要呈北北西向延伸, 自晚第四纪以来表现为逆—走滑运动性质, 且以水平走滑运动为主。岷江断裂是划分西部川青块体和东部岷山断块的界限, 自晚更新世以来显示为逆—走滑运动性质, 且水平位错量与垂直位错量大致相当。川青块体向南东东方向的滑移过程现今仍在继续, 导致了岷山断块区强烈的近代地震活动 (陈智梁等, 1998; Chen *et al.*, 2000; 周荣军等, 2006)。

龙门山断裂带是青藏高原东缘边界山脉, 不仅是中国东部和西部地形地貌的界限, 还地处我国南北地震带的中段, 地震活动频度高, 强度大, 是青藏高原北部地震亚区主要强震活动带之一 (邓起东等, 1994; 李勇等, 2006)。该断裂带主要由一系列大致平行的北东—南西向的逆冲断裂组成, 由西向东分别为: 汶川—茂县断裂、映秀

—北川断裂和安县—灌县断裂。汶川—茂县和映秀—北川断裂是控制龙门山快速隆升的重要断裂。龙门山断裂带具有长期活动性, 自晚三叠世至古近纪时期, 以逆冲—左旋走滑作用为特征, 而中生代以来均显示由北西向南东的逆冲运动, 并具有显著的右旋走滑分量 (周荣军等, 2006; 李勇等, 2006)。

3 龙门山地区震级上限估计

3.1 地震活动

本文数据来源主要为国家地震科学数据共享中心 (<http://data.earthquake.cn>) 的中国历史地震目录数据库以及国家地震台网地震目录数据库, 黄玮琼等 (1994) 对中国大陆大部分地区的历史地震资料作了较详细的分析与研究, 显示龙门山区域 $M_s \geq 4 \frac{3}{4}$ 地震资料在 1931 年之后基本完整。

因此，笔者收集了1930—2016年龙门山地区($101^{\circ}\sim 106^{\circ}\text{E}$, $29.5^{\circ}\sim 33.5^{\circ}\text{N}$) $M\geq 4.5$ 的历史地震目录。经统计,1930年至今,龙门山地区发生过5次 $M\geq 7.0$ 强震,其中在龙门山断裂带上发生过2次,分别是2008年汶川8.0级地震和2013年芦山7.0级地震;北部岷山地区发生过3次,分别是1933年叠溪7.5级地震和1976年松潘2次7.2级地震。

在利用广义帕累托模型进行地震危险性估计之前需要尽量删除地震目录中的余震。目前国内外地震危险性分析研究中,考虑到计算的简便和工程应用上的可操作性,普遍采用Gardner和Knopoff(1974)提出的时空窗法来删除余震(徐伟进,2012)。在该方法中,空间窗口根据式(9)确定(Knopoff, Gardner, 1972):

$$\lg R = aM + b$$

(9)

式中: a , b 为固定参数值,在我国 a , b 的取值分别为0.5、-1.78(刘杰等,1996)。

时间窗口如表1所示。删除余震后,共收集到龙门山地震构造区内地震事件118个,对收集到的地震资料统计发现,龙门山地震构造区内以浅源地震为主,其中 $M\geq 7$ 地震5次, $6\leq M<7$ 地震14次, $5\leq M<6$ 地震34次。从图1可看出,地震震中分布位置与断裂带的展布形态密切相关。

表1 不同震级地震的余震时间窗(Gardner, Knopoff, 1974)
Tab. 1 Aftershock time windows of different magnitude earthquakes (Gardner, Knopoff, 1974)

主震震级 M_s	持续天数/d	主震震级 M_s	持续天数/d
4.5	83	7	915
5	155	7.5	960
5.5	290	8	985
6	510	8.5	985
6.5	790		

3.2 震级上限估计

利用广义帕累托分布估计龙门山地区震级上限,首先需确定震级阈值 u 。阈值选取通过2种方法:(1)依据震级平均超出量分布函数图,震级平均超出量与震级阈值应满足线性关系;(2)随着阈值选取不断增大,形状参数 ξ 和修饰的尺度参数 $\tilde{\sigma}$ 估计值应基本保持稳定不变(Coles, 2001)。

由图2可以看出,当震级阈值在(5.1, 5.5)、(5.7, 6.0)以及(6.3, 6.5)范围内时,平均超出量近似呈线性分布。为了对阈值进行最合理的选取,还需要综合考虑样本数据的实际情况,阈值选取过大,超出量样本数过少,误差增大。结合图3分析,阈值在(5.1, 5.5)范围内相应的形状参数和修饰的尺度参数的估计值基本保持不变,根据广义帕累托分布应用条件,选取尽可能大的阈值 $u=5.5$ 。

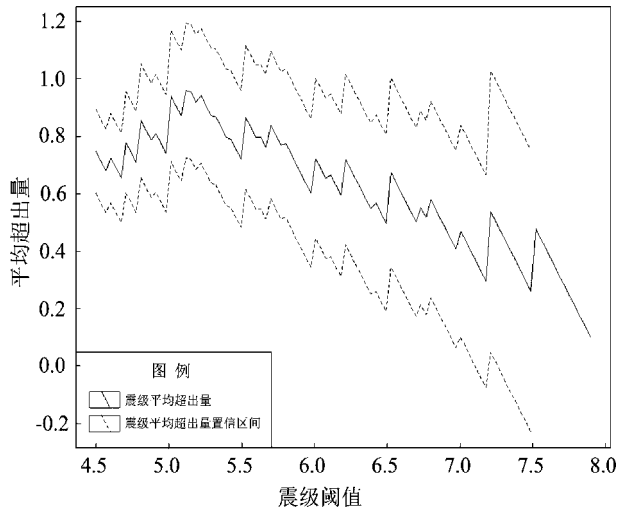


图2 震级平均超出量分布函数图
Fig. 2 Diagram of mean excess function of seismic magnitude

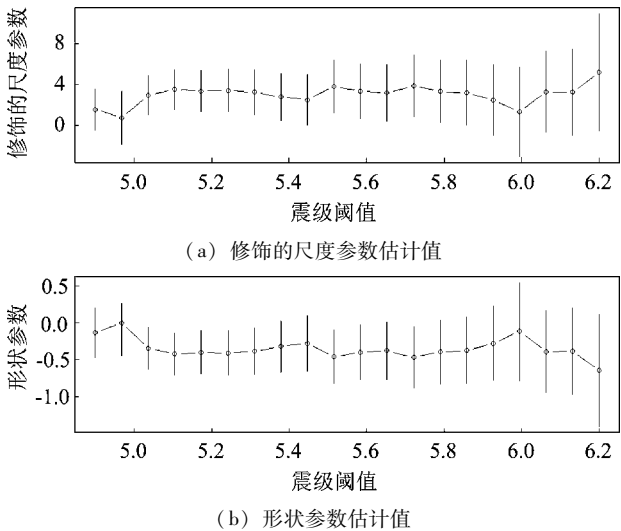


图3 参数估计值随震级阈值选取的变化图
Fig. 3 Variation diagram of parameter estimates according to the seismic magnitude threshold variation

在确定震级阈值 u 后,通过概率图、分位数图、重现水平图以及概率密度图可以对所拟合的广义帕累托分布模型进行检验。当数据符合理论假设分布时,概率图与分位数图应近似为直线。重现水平图可作为模型检验的诊断工具,当 $\xi = 0$ 时为直线,当 $\xi < 0$ 时为凸曲线,当 $\xi > 0$ 时为凹曲线。模型诊断如图 4 所示。

图 4a 是概率图,横坐标表示实际数据的累积概率,纵坐标表示选用分布模型的累积概率;图

4b 是分位数图,横坐标表示所选分布模型的分位数,纵坐标表示实际数据利用经验分布函数计算得到的分位数。从图 4a, b 可看出,大部分点均沿直线分布,因此不能否认所拟合的模型。由于 ξ 的估计值为负,因此最大震级分布存在有限上界,故重现水平图中的曲线为凸曲线,渐近地趋于有限值(图 4c)。概率密度图中的密度曲线估计与直方图也显示一定的拟合(图 4d)。因此 4 个诊断图均支持拟合的广义帕累托分布模型。

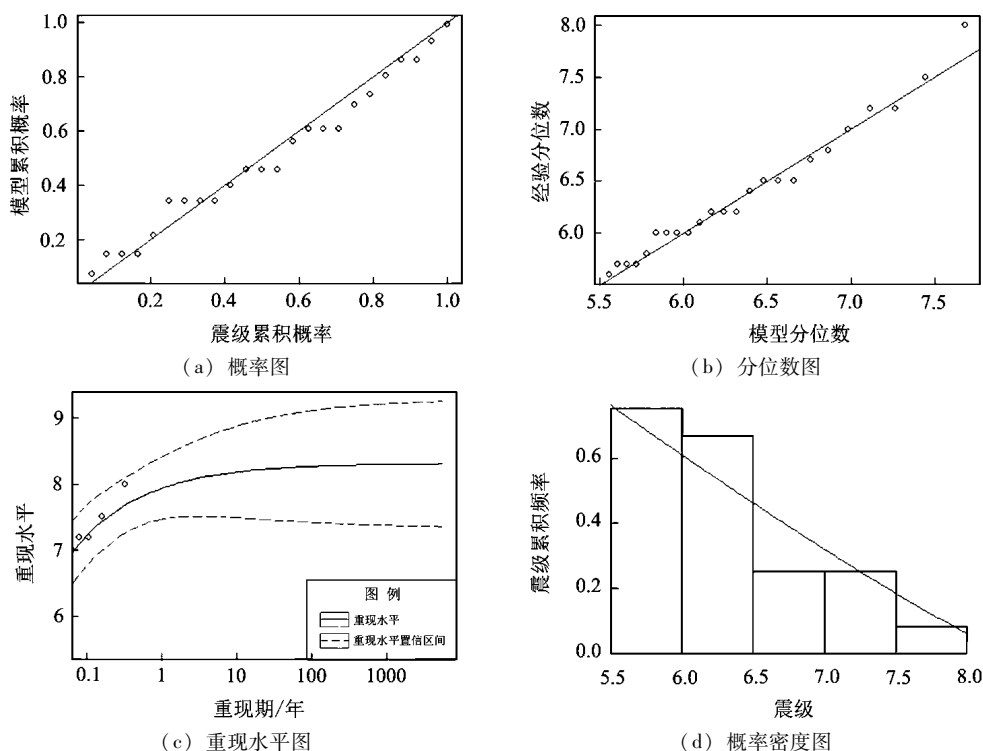


图 4 广义帕累托分布拟合诊断图

Fig. 4 Diagnostic plots of the Generalized Pareto Distribution fitted to seismic magnitude

利用最大似然法估计广义帕累托分布的参数,得到修饰的尺度参数和形状参数的估计值为 $(\hat{\sigma}, \hat{\xi}) = (1.31, -0.47)$, 其中形状参数估计的标准差为 0.18, 其 95% 的置信区间为 $[-0.82, -0.12]$, 形状参数的最大似然估计与置信区间都显示为负值, 这表明龙门山地区在研究时间尺度范围内震级超出量分布具有上限值, 这也与潜在震源区一定具有震级上限的认识相符。

将阈值、形状参数与修饰的尺度参数估计值代入式 (7) 得到震级上限为 8.3。对震级上限的误差进行分析, 首先需计算参数估计的协方差矩阵:

$$V = \begin{Bmatrix} 0.0014 & 0 & 0 \\ 0 & 0.1134 & -0.0573 \\ 0 & -0.0573 & 0.0333 \end{Bmatrix} \quad (10)$$

将式 (10) 协方差矩阵代入式 (8), 得到震级上限的标准差为 0.50, 由上限估计值和标准差可知震级上限置信度为 95% 的置信区间为 $[7.3, 9.3]$ 。

4 结论与讨论

通过对龙门山地区历史强震记录进行统计, 从拟合结果可以看出震级超出量分布服从广义帕累托分布。依据震级平均超出量分布函数图以及

参数估计值变化图, 综合分析选取 5.5 为震级阈值, 利用最大似然估计方法, 获得修饰的尺度参数以及形状参数的估计值分别为 1.31、-0.47。对龙门山地区强震危险性进行分析, 计算得到龙门山地区的震级上限为 8.3, 置信度为 95% 的置信区间为 [7.3, 9.3]。

在利用广义帕累托分布计算龙门山地区震级上限时, 用于统计的震级样本具有相似的地震构造环境和发震构造模型, 满足构建广义帕累托模型时样本属于同分布的条件。另外, 采用时空窗法相对地减少余震对统计结果的影响, 也更加符合事件相互独立的条件。虽然时空窗法的经验公式在删除余震时难免会存在少量误差, 但本文在计算龙门山地区震级上限时, 震级阈值设定为 5.5, 余震震级大于 5.5 的数量较少, 因此, 删除余震的误差对于计算结果的影响相对较小。关于龙门山地区地震目录余震删除方法在模型中的应用还需进一步的研究与探讨。

本文利用广义帕累托分布模型计算出的震级上限为 8.3, 高于第五代《中国地震动参数区划图》中龙门山构造区的震级上限 8.0。广义帕累托分布可以充分利用历史地震大震级区段的数据资料, 在小震资料不充足的情况下也可构建潜源区的地震活动模型, 其统计结果可作为地震区划以及地震危险性分析研究工作的参考。

本文在撰写过程中得到蒋长胜研究员、边银菊研究员以及孔韩东博士的帮助, 在此表示衷心感谢。

参考文献:

陈智梁, 刘宇平, 赵济湘, 等. 1998. 青藏高原东部地壳运动的 GPS 测量[J]. 中国地质, (5): 32-35.

邓起东, 陈社发, 赵小麟. 1994. 龙门山及其邻区的构造和地震活动及动力学[J]. 地震地质, 16(4): 389-403.

邓起东. 2007. 中国活动构造图(1:400 万)[M]. 北京: 地震出版社.

高孟潭. 2015. GB 18306—2015《中国地震动参数区划图》宣贯教材[M]. 北京: 中国质检出版社.

胡聿贤. 1999. 地震安全性评价技术教程[M]. 北京: 地震出版社, 183-193.

黄玮琼, 李文香, 曹学锋. 1994. 中国大陆地震资料完整性研究之二: 分区地震资料基本完整的起始年分布图像[J]. 地震学报, 16(4): 423-432.

孔军, 周荣军. 2014. 龙门山和成都地震构造区的划分[J]. 震灾防御技术, 9(1): 64-73.

李海兵, 司家亮, 付小方, 等. 2009. 2008 年汶川地震同震滑移特征、最大滑移量及构造意义[J]. 第四纪研究, 29(3): 387-402.

李勇, 周荣军, Densmore A L, 等. 2006. 青藏高原东缘龙门山晚新生代走滑—逆冲作用的地貌标志[J]. 第四纪研究, 26(1): 40-51.

刘杰, 陈棋福, 陈颢. 1996. 华北地区地震目录完全性分析[J]. 地震, (1): 59-67.

潘华, 高孟潭, 李金臣. 2009. 新版美国地震区划图源及其参数模型的分析与评述[J]. 震灾防御技术, 4(2): 131-140.

潘华, 李金臣. 2016. 新一代地震区划图的地震活动性模型[J]. 城市与减灾, 108(3): 28-33.

钱小仕, 蔡晓光, 任晴晴. 2013. 中国大陆活动地块边界带强震震级分布特征研究[J]. 地震工程与工程振动, 33(1): 212-220.

史道济. 2006. 实用极值统计方法[M]. 天津: 天津科学技术出版社.

田建伟, 刘哲, 任鲁川. 2017. 基于广义帕累托分布的马尼拉海沟俯冲带地震危险性估计[J]. 地震, 37(1): 158-165.

闻学泽, 杜方, 张培震, 等. 2011. 巴颜喀拉块体北和东边界大地震序列的关联性 with 2008 年汶川地震[J]. 地球物理学报, 54(3): 706-716.

胥广银. 2003. 潜在震源三维空间模型及其在地震危险性概率分析中的应用研究[D]. 北京: 中国地震局地球物理研究所.

徐伟进, 高孟潭. 2012. 根据截断的 G-R 模型计算东北地震区震级上限[J]. 地球物理学报, 55(5): 1710-1717.

徐伟进. 2012. 地震危险性分析中地震时空统计分布模型研究[D]. 北京: 中国地震局地球物理研究所.

易桂喜, 闻学泽, 辛华, 等. 2011. 2008 年汶川 $M_s 8.0$ 地震前龙门山—岷山构造带的地震活动性参数与地震视应力分布[J]. 地球物理学报, 54(6): 1490-1500.

周荣军, 李勇, Alexander, 等. 2006. 青藏高原东缘活动构造[J]. 矿物岩石, 26(2): 40-51.

Balkema A A, Haan L D. 1974. Residual Life Time at Great Age[J]. Annals of Probability, 2(5): 792-804.

Chen Z, Burchfiel B C, Liu Y, et al. 2000. Global Positioning System measurements from eastern Tibet and their implications for India/Eurasia intercontinental deformation[J]. Journal of Geophysical Research Solid Earth, 105(B7): 16215-16227.

Coles S. 2001. An Introduction to Statistical Modeling of Extreme Values [M]. Berlin: Springer-Verlag.

Cornell C A. 1968. Engineering seismic risk analysis[J]. Bulletin of the Seismological Society of America, 58(5): 1583-1606.

Densmore A L, Ellis M A, Li Y, et al. 2007. Active tectonics of the Beichuan and Pengguan faults at the eastern margin of the Tibetan Plateau[J]. Tectonics, 26(4): TC4005.

Fisher R A, Tippett L H C. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample[J]. Mathematical Proceedings of the Cambridge Philosophical Society, 24(2): 180-190.

Gardner J K, Knopoff L. 1974. Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? [J]. Bulletin of

- the Seismological Society of America, 64(5):1363–1367.
- Huyse L, Chen R, Stamatakis J A. 2010. Application of Generalized Pareto Distribution to Constrain Uncertainty in Peak Ground Accelerations[J]. Bulletin of the Seismological Society of America, 100(1): 87–101.
- Knopoff L, Gardner J K. 1972. Higher seismic activity during local night on the raw worldwide earthquake catalogue[J]. Geophysical Journal of the Royal Astronomical Society, 28(3):311–313.
- Pickands J. 1975. Statistical Inference Using Extreme Order Statistics[J]. Annals of Statistics, 3(1):119–131.
- Pisarenko V F, Sornette D. 2003. Characterization of the Frequency of Extreme Earthquake Events by the Generalized Pareto Distribution[J]. Pure & Applied Geophysics, 160(12):2343–2364.

The Establishment of Generalized Pareto Distribution Model of Seismicity in Longmenshan Region

REN Mengyi

(Institute of Geophysics, China Earthquake Administration, Beijing 100081, China)

Abstract

Choosing Longmenshan region as the case study area, using the historical earthquake catalogue with magnitude 4.5 or above, from the year 1930 to 2016, reducing the impact of aftershocks by the method of space and time window, this paper establishes an excess distribution model of Generalized Pareto Distribution, and it estimates the upper magnitude limit of Longmenshan region. The result indicates that the model is well fitting for the earthquake data and the estimate value of shape parameter is less than zero, which means that the maximum magnitude distribution has a limited upper bound. The upper magnitude limit is 8.3. The result can be a reference for earthquake zoning and seismic hazard analysis researches.

Keywords: seismic activity; generalized pareto distribution; upper magnitude limit; Longmenshan Region