

徐伟进,高孟潭. 2014. 中国大陆及周缘地震目录完整性统计分析. 地球物理学报, 57(9): 2802-2812, doi: 10. 6038/cjg20140907.

Xu W J, Gao M T. 2014. Statistical analysis of the completeness of earthquake catalogs in China mainland. *Chinese J. Geophys.* (in Chinese), 57(9): 2802-2812, doi: 10. 6038/cjg20140907.

中国大陆及周缘地震目录完整性统计分析

徐伟进, 高孟潭

中国地震局地球物理研究所, 北京 100081

摘要 地震目录资料是进行地震预测、地震危险性分析、地震活动性研究等地震学研究的基础资料. 对地震目录资料的完整性可靠性分析是地震学研究的基础工作之一. 近几十年来我国积累的大量的仪器地震记录, 历史地震也在不断更新, 为地震活动性研究提供了更为丰富的样本, 因此有必要对新的地震目录进行完整性分析. 本文采用了新的统计方法对我国仪器记录地震目录和历史破坏性地震目录进行了完整性分析. 结果表明, 对于仪器记录地震目录($M \geq 3.0$)我国东部地区 1975 年后基本完整, 西部地区 1980 年后基本完整. 对于历史地震目录($M \geq 5.0$), 东部地区 1500 年后基本完整, 西部地区 1950 年后基本完整. 考虑到华北地震区历史破坏性地震($M \geq 4^{3/4}$)活动的时间非平稳性, 我们采用突变点分析法研究了华北地震区历史破坏性地震目录的时间特征, 得到了华北地震区地震活动周期性变化的突变点位置, 并揭示了华北地震区历史破坏性地震目录完整起始时间约为公元 1500 年前后.

关键词 地震目录; 完整性分析; 中国大陆

doi: 10. 6038/cjg20140907

中图分类号 P315

收稿日期 2013-10-15, 2014-08-27 收修定稿

Statistical analysis of the completeness of earthquake catalogs in China mainland

XU Wei-Jin, GAO Meng-Tan

Institute of Geophysics, China Earthquake Administrator, Beijing 100081, China

Abstract Earthquake catalogs are the basic data for earthquake prediction, seismic hazard analysis and other seismological studies. The analysis of the completeness and reliability of seismic catalogs is one basic work of seismology. In recent decades, our country has accumulated a large number of catalogs of instrumental and historical of destructive earthquakes. These catalogs provide rich samples for seismicity study. So it is necessary to analyze the completeness of these catalogs. In this paper, we use several statistical approaches to conduct this work. The results show that, the instrumental earthquake catalogs for magnitude equal to 3.0 or greater can be considered complete after 1975 for eastern Chinese mainland and after 1980 for western China. The historical catalog for magnitude equal to 5.0 or greater can be considered complete after 1500 for eastern Chinese mainland and after 1950 for western China. Considering the temporal non-stationary of the historical destructive earthquake catalogs in North China, we analyze the completeness of earthquake catalogs for this region by using Change-point analysis method. The result shows that the historical destructive earthquake catalogs in North China can be considered complete after 1500.

Keywords Earthquake catalog; Analysis of the completeness; China mainland

1 引言

地震目录是指按照时间顺序,对地震的时、空、强等参数进行收录,编辑成册的目录资料. 在对一个地区的地震活动性进行研究时,对该地区地震目录的分析是基本前提,我国地震目录主要包括现代仪器记录目录和从古代文献记载中收集总结的历史破坏性地震目录. 在地震目录的收集过程中由于搜集程序、数据分析方法及可用信息不同,这些目录在不同的时空范围内往往具有不同的完整性水平. 对于仪器记录地震来说,由于不同时期台网密度的变化以及求取地震震级参数和方法的变化也导致汇编的地震目录在时空上具有不同的完整性水平. 对于历史地震目录,由于缺少完整的记录文献,以及遇战争、自然灾害等导致记录文献的丢失损毁,因此人们收集整理的目录往往是缺失的,年代越久远地震目录丢失的概率就越大.

在地震活动性研究中,不完整的地震目录会给地震活动性参数的估计带来偏差,因此在研究一个地区的地震活动性时分析这一地区的地震目录可靠性是非常重要的工作. 国内外学者提出了多种方法来分析地震目录的完整性:在地震危险性分析中最常用的是 Gutenberg-Richter 关系曲线(Gutenberg and Richter, 1944) (简称 $G-R$ 关系)和时间累计频度曲线的线性是否良好来确定某一起始震级的地震目录是否完整(Franekl et al., 1996; Kafka and Levin, 2000; 徐伟进和高孟潭, 2012b, 2012c). Wiemer 和 Wyss. (2000)及 Woessner 和 Wiemer(2005)等采用 $G-R$ 关系的拟合优度来分析仪器记录的地震目录完整性. 由于 $G-R$ 关系是一个对数线性公式,对于小震的缺失并不敏感,即使在小震有大量缺失的情况下也可能得出良好的 $G-R$ 关系曲线(任雪梅, 2011),因此采用 $G-R$ 关系来判别地震目录的完整性可能并不可靠,特别是对历史地震目录的完整性分析,这一方法可能并不适用.

此外还有其他学者提出了多种统计方法(Stepp, 1971; Caputo and Postpischl, 1974; Bath, 1983; Tinti and Mulargia, 1985a, b; Mulargia and Tinti, 1985; Mulargia et al., 1987a, b; Rydelek and Sacks, 1989; Rotondi et al., 1994)用于地震目录的完整性分析. 然而 Albarello 等(2001)认为上述统计方法中学者们均假设地震时间过程是一个平稳过程,并且忽略了以下两点:(1)在分析地震目录时未考虑所采用的地震目录是否具有代表性,所谓代

表性即地震目录的记录周期能够跨越地震的复发周期且能够反应地震发生过程的统计特征;(2)所分析的地震目录并未包含至少一个完整的子目录,且这一子目录的地震活动率不能代表实际的地震活动率. 因此 Albarello 等(2001)考虑了上述两点要素,改进了 Mulargia 等(1987a, b)的半定性方法,提出了一种新的统计方法.

我国具有丰富的仪器和历史地震记录,对于地震目录的分析也从来没有停止过. Lee 和 Brillinger (1979)采用点过程分析的手段企图建立一个不完整的地震目录数据库模型;刘杰等(1996)采用 Tinti 和 Mulargia(1985a)的方法分析了华北地区的地震目录完整性;汪雪泉和潘坚云(2002)采用了 Fisher 的时序分割法研究了华东地区的地震目录完整性;黄玮琼等(1994a, b)采用多种统计手段研究了华北地震区及全国地震带不同起始震级地震目录的完整起始时间. 此外,苏有锦(2003)、王海涛等(2006)采用 $G-R$ 关系拟合法分别研究了川滇地区和新疆地区地震目录的完整起始时间. Mignan 等(2013)采用 Bayesian magnitude of completeness (BMC)方法分析了中国大陆仪器记录地震目录的完整性. 显然上述分析方法大都是基于 $G-R$ 关系进行的,有些方法还需要用到地震台站的经纬度数据. 这些方法仅适用于仪器记录地震目录较为丰富的地区,对于历史地震目录,上述方法大都并不适用.

本文将采用 Albarello 等(2001)提出的新的统计方法分析了我国大陆地区历史强震记录和仪器记录地震目录的完整起始时间. 并且特别地采用变点分析法分析了华北地震区历史破坏性地震($M \geq 4.5$)的地震活动率的突变点及地震目录完整性.

2 地震目录及余震删除

文中主要分析了我国大陆仪器记录的地震目录和历史破坏性地震记录. 研究所用的历史强震资料来源于《中国历史强震目录(公元前 23 世纪—公元 1911 年)》(国家地震局震害防御司, 1995), 1912—1990 年的资料来自《中国近代地震目录(公元 1912—公元 1990 年)》(中国地震局震害防御司, 1999), 1990 年以后的资料来自中国地震局的地震目录, 台网数据来源于各省的地震监测报告和中国地震局的地震目录.

在进行地震目录完整性分析之前往往需要删除目录中的余震和前震. 目前世界各国的地震危险性

分析中,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)提出的余震删除方法.文中也采用这一方法进行余震删除.该方法中,空间窗口由下式确定(Knopoff and Gardner, 1972):

$$\lg R = aM + b$$

其中, R 为地震破裂长度, M 为地震震级, a 、 b 为固定参数值,根据搜集到的不同震级地震的破裂长度回归得到.在我国 a 、 b 的取值分别为 0.5、-1.78(刘杰等,1996).

时间窗口根据表 1 给出.震级不在表中的根据线性插值得出.

图 1 为我国大陆及周缘地区 1970 年以来 $M \geq 3.0$ 级以上主震空间分布图.图 2 为我国大陆及周缘地区

表 1 不同震级地震的余震时间窗
(Gardner and Knopoff, 1974)

Table 1 Time window algorithm for aftershocks
(Gardner and Knopoff, 1974)

主震震级(M_s)	持续时间(天)	主震震级(M_s)	持续时间(天)
2.5	6	6.0	510
3.0	11.5	6.5	790
3.5	22	7.0	915
4.0	42	7.5	960
4.5	83	8.0	985
5.0	155	8.5	985
5.5	290		

公元 1000 年以来 $M \geq 5.0$ 级历史强震空间分布图.

3 方法

3.1 方法简介

文中首先采用 Albarello 等(2001)提出的新的统计方法分析了我国大陆地区历史地震目录和仪器记录地震目录的完整起始时间.该方法简介如下:

假设一套地震目录是完整的(假设 C)、具有代表性的(假设 R).在分析的地震目录中存在时间跨度为 ΔT 的地震子目录,在 ΔT 内,(1)地震目录具有代表性,所谓代表性即地震目录的记录周期能够跨越地震的复发周期且能够反应地震发生过程的统计特征(假设 R).其数学表达将在下文中描述;(2)所分析的地震目录包含至少一个完整的子目录,且这一子目录的地震活动率能够代表实际的地震活动率(假设 C).下文中详细推导了关于上述两个假设的条件和非条件概率表达式.满足上述两个假设的概率最大的子目录则认为是完整的子目录.

把计算得到的每年的地震发生频次作为地震活动率,将其称之为样本估计.这里计算的地震活动率是符合泊松分布的,但是根据中心极限定理,可以近似将其看作是符合高斯分布的.由于高斯分布的对称性,在两个不同时间段 Δt_i 和 Δt_j 内观测到的地震活动率 n_i 和 n_j 应满足下式:

$$p(n_i > n_j) = p(n_i < n_j) = 0.5, \quad (1)$$

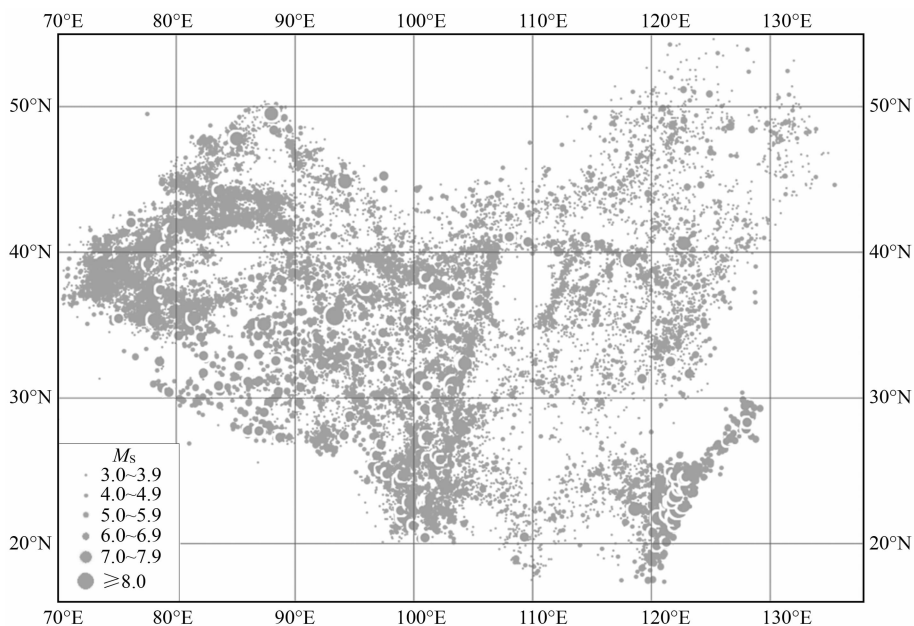


图 1 中国大陆及周缘公元 1970 年以来 $M \geq 3.0$ 级地震空间分布

Fig. 1 Epicentral distribution of earthquakes ($M \geq 3.0$) reported in the China's instrumental seismic catalog

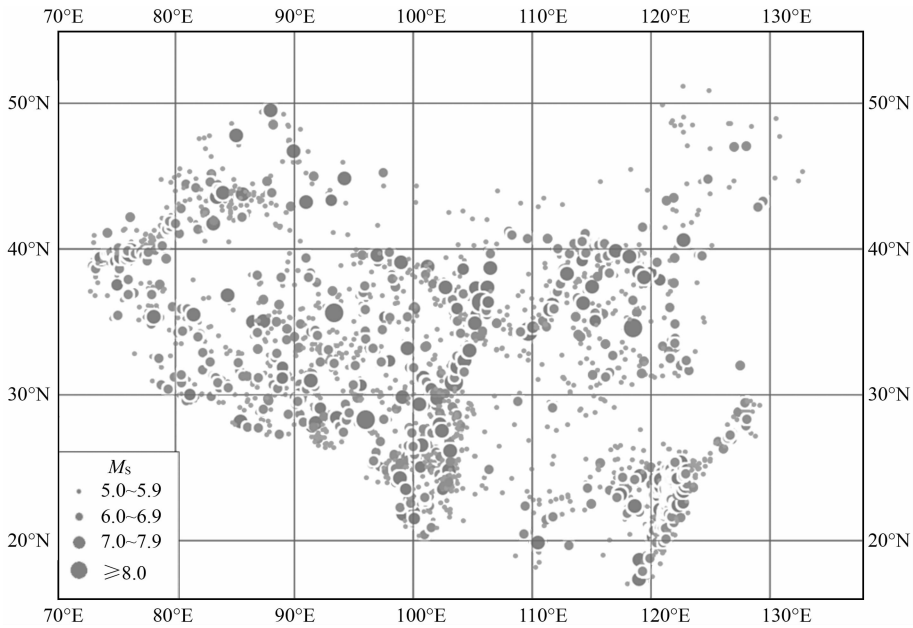


图 2 中国大陆及周缘公元 1000 年以来 $M \geq 5.0$ 级地震空间分布

Fig. 2 Epicentral distribution of earthquakes ($M \geq 5.0$) reported in the China's historical seismic catalog

其中 p 是相关概率. 如果接受假设 R , 即地震的发生是一个独立的过程, 那么就认为在 ΔT 时间段内地震目录是完整的.

为了实现上述假设, 首先把地震目录的时间段 ΔT 以 δt 为间隔分成 $2N$ 个不重合的子段, $2N = \Delta T / \delta t$. 对于每一子段 δt_i 计算该段内的地震活动率 n_i , 基于这些地震活动率定义一个二进制随机变量 D :

$$\begin{cases} D_i = 1, & n_i > n_{i+N} \\ D_i = 0, & n_i < n_{i+N} \end{cases} \quad (2)$$

实际上变量 D 是地震目录前一部分地震活动率和后一部分地震活动率的比较. 若 $D=0$ 的次数较多就意味着时间较早的地震有缺失, 地震目录完整程度下降.

对于 N' 个可用的地震目录子段, m 为 $D=0$ 的次数. 如果目录是完整的, 那么实际观测中计算得到的次数大于等于 m 的概率可用二项式计算:

$$P(C | R) = \sum_{j=m}^{N'} \frac{N'!}{j!(N'-j)!} p^j (1-p)^{N'-j}. \quad (3)$$

据(1)式 p 值取 0.5. 这一概率主要依赖假设 R 的置信程度. 概率值越小意味目录完整程度越低.

上述过程是地震子目录在整个目录中具有代表性的条件下其完整性的概率. 显然希望得到的是某一具体实际段 ΔT_i 内地震完整的非条件概率:

$$P_i(C) = P_i(R)P_i(C | R), \quad (4)$$

其中 $P_i(R)$ 表示地震目录在时间段 ΔT_i 内具有统计代表性, 即假设 R 的置信程度.

为了估计假设 R 的概率, 应当考虑地震孕育过

程的长期统计特征, 当然这一概率并非是一个先验概率. 假设 R 的概率简单写为

$$P_i(R) = \frac{\Delta T_i}{k \Delta T_{\max}}, \quad (5)$$

其中 ΔT_{\max} 是采用整个地震目录的记录周期. k 是一个大于 1 的定值. 因此随着 ΔT_i 的增大 $P_i(R)$ 是线性递增的.

联合(3)、(4)、(5)式, 则:

$$P_i(C) = \frac{\Delta T}{k \Delta T_{\max}} \sum_{j=m}^{N'} \frac{N'!}{j!(N'-j)!} p^j (1-p)^{N'-j}. \quad (6)$$

上述中 $\Delta T_i = T_p - T_i$, $\Delta T_{\max} = T_p - T_0$, 其中 T_p 为地震记录的截止时间, T_i 为每次循环计算时所采用的地震目录的起始时间, T_0 为所采用的整个地震目录的起始记录时间.

上述计算了 M 个 T_i 值时的 $P_i(C)$ 值, 本文目的是找出最优的 T_i 作为完整目录的起始时间 T_c . 这主要依赖于本文所采用的地震目录中需要有至少一个子目录是完整的. 即 $\sum_{i=1}^M P_i = 1$. 然而这是很难得到的, 因此需要将 P_i 值重新标度来得到新的完整性函数, 即:

$$P_i^* = \frac{P_i}{\sum_{i=1}^M P_i}, \quad (7)$$

P_i^* 的使用让我们在即使完整性函数不是很规则或二项分布的情况下也能很好的计算完整时间 T_c . 因

此基于上述情况, T_c 值可以用完整性函数 P_i^* 的中值来估计, 即:

$$\sum_{i=1}^c P_i^* = 0.5. \quad (8)$$

这一估计的不确定性可以用四分位来计算:

$$\delta T = T_u - T_l, \quad (9)$$

$$\text{其中} \begin{cases} \sum_{i=1}^u P_i^* = 0.75 \\ \sum_{i=1}^l P_i^* = 0.25 \end{cases}.$$

3.2 数值模拟

在将上述方法应用到实际地震目录之前, 先将这种方法应用到模拟地震目录中, 以验证这一方法在统计学上的可靠性和无偏性. 对于历史地震参照华北地震区破坏性地震活动率, 取 0.78 为模拟地震目录的活动率, 以公元 1000—2000 年为模拟时间段, 模拟了 1000 套泊松时间流, 将其看成模拟地震目录. 然后将 1800 年以前的地震目录进行人为删减, 具体做法是将 1800 年以前的每个模拟地震事件对应生成一个 $[0-1]$ 之间的随机数, 对随机数小于 0.7 的进行删除, 只考虑随机数大于 0.7 的地震事件, 这意味着 1800 年以前的地震目录中地震丢失率超过 70%. 对于仪器记录地震, 同样参照华北地震区地震活动率, 以 100 为模拟人造地震目录的活动率, 以 1970—2009 年为模拟时间段, 模拟了 1000 套人造地震目录, 将模拟记录中 1980 以前的地震事件按 30% 丢失率进行删减. 对模拟地震目录进行人为删减, 是为了确定一个已知的完整起始时间.

对进行删除后的每套模拟地震目录按上文中描述的方法进行计算, 寻找完整目录起始时间. 图 3 为在对 1000 套模拟目录进行完整性分析后得到的完

整起始时间的概率密度分布图. 从图 3a 中可以看出历史强震的完整时间主要集中在 1800 年左右, 从图 3b 中可以看出仪器记录地震的完整起始时间主要集中在 1980 年左右, 与已知的完整起始时间保持一致. 这一结果说明了采用上述方法进行地震目录完整性分析是有效的.

此外, 还计算了上述方法中样本尺度与样本丢失率对完整起始时间的影响. 图 4a 为目录完整起始时间及其不确定性与样本量的关系曲线, 可以看出在样本长度少于 40 时计算得到的完整起始年偏离理论值很远, 但理论值仍在估计值不确定性区间内. 当样本长度大于等于 40 时计算得到的估计值在理论值附近, 绝对差值不超过 20 年. 因此在下文的计算中以 40 为计算所需要的最低样本长度. 图 4b 为采用模拟记录计算得到的完整起始年估计值随样本丢失率的变化曲线, 可以看出在丢失率较小时, 估计值与理论值相差较大, 随着丢失率的增大, 计算得到的完整起始年估计值有变大的趋势, 特别是在丢失率达到 80% 时, 计算得到的估计值不确定性的下限已与理论值相当大小. 因此对于丢失率较大的历史地震事件, 计算得到的完整起始时间可能偏于保守; 而对于丢失率较小的仪器记录地震, 计算得到的完整起始年可能比实际值偏小. 然而考虑到计算结果的不确定性较大, 因此把计算得到的完整起始年估计值作为地震目录的完整起始时间是较为可行的, 可以用来分析研究.

4 中国大陆地震目录完整性分析

4.1 历史地震目录完整性分析

采用上一部分描述的统计方法, 分析了我国大

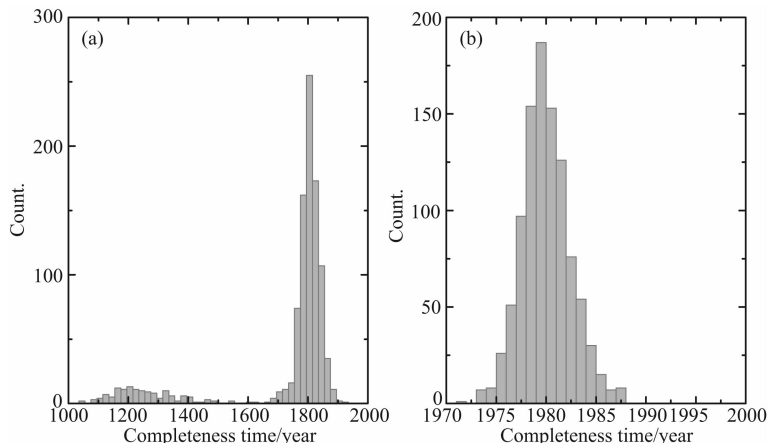


图 3 采用 1000 次模拟目录计算得到的完整起始年分布直方图

Fig. 3 Frequency of completeness time values T_c obtained after the analysis of 1000 synthetic catalogs

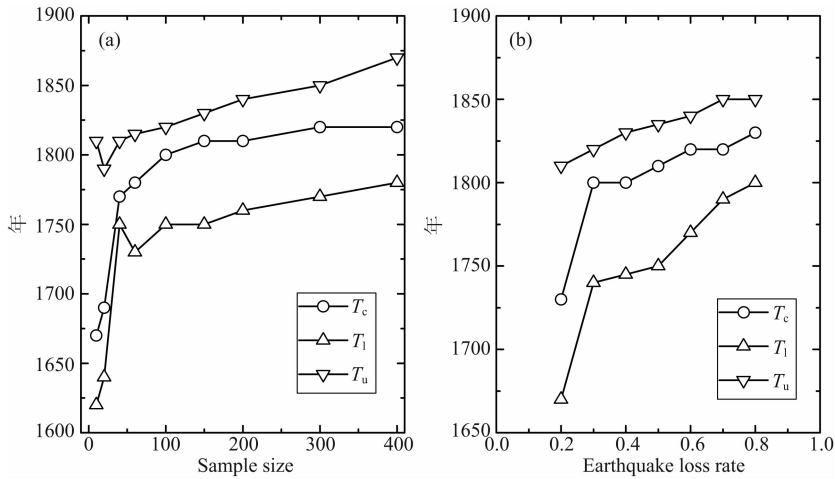


图4 (a)样本量及(b)地震丢失率对分析方法的影响

Fig. 4 Effects of (a) sample size and (b) loss rate on statistical method

陆地区历史地震的完整起始时间的空间分布. 将研究区域分成 $400 \text{ km} \times 400 \text{ km}$ 的网格, 以 200 km 为滑动窗口, 对于地震频次大于 40 的网格进行了分析. 图 5 可以看出华北地震区历史地震的完整性最好, 该地区的 $M \geq 5.0$ 历史地震在 1500 年左右时基本完整的; 华南地震带南部历史地震也在 1500 年以来基本完整; 我国西部绝大部分地区历史地震在 1900 年后才是完整的, 其中大部分地区是在 1920 年甚至 1950 年后 $M \geq 5.0$ 级地震才基本不漏. 东北地震区由于历史地震较少, 没有统计意义, 因此无法统计网格内的历史地震完整性, 对于整个东北地区的历史地震完整性在下文中有具体分析. 因此总体上来说我国东部地区 $M \geq 5.0$ 级地震在 1500 年以后是基本完整的, 可以用来分析研究. 西部地区 1950 年以后历史地震目录是完整的.

4.2 仪器记录地震目录完整性分析

从图 6 可以看出我国东部大部分地区 $M \geq 3.0$ 级地震在 1970—1975 年以后是完整的; 西部大部分地区特别是西藏和新疆地区 $M \geq 3.0$ 级地震在 1980 甚至 1985 年后是完整的. 川滇地区历来是地震研究的热点地区, 其观测能力较强, 可以认为这一地区 $M \geq 3.0$ 级地震 1975 年以后是完整的.

此外, 分析了东北、华北、华南、新疆、青藏高原及川滇地震区不同震级地震的完整起始年及其不确定性. 对于 $M \geq 3.0$ 级的仪器记录地震: 在华北、东北、华南及川滇地区可以认为 1975 年以来是完整的; 在新疆和青藏高原地震区可以认为 1980 年来是完整的. 对 $M > 5.0$ 级的破坏性地震: 华北和华南地区可以认为 1500 年以来是完整的; 东北地震区可以认为 1920 年以来是完整的; 新疆和青藏高原地震区

1950 年左右以来是完整的; 川滇地区 1900 年左右可以认为是完整的. 还计算了部分地区 $M \geq 6.0$ 、 $M \geq 7.0$ 级地震的完整起始年, 具体数值见表 2, 其中 T_c 为完整起始时间, T_u 为不确定性的上限值, T_l 为不确定性的下限值.

从表 2 中还可以看出震级越高其完整起始年的不确定性越大, 完整起始年的不确定性能达到几十甚至上百年. 这是由于随着震级的增大, 地震事件样本长度急剧减小, 从而导致了不确定性的增加, 其次对于大地震来说其复发周期很长, 目前的地震记录周期无法覆盖其复发周期, 从而也导致了其完整起始年不确定较大.

5 华北地震区历史破坏性地震完整性分析

特别地, 由于华北地震区历史地震记录较为丰富且地震活动具有周期性. 鉴于华北地区重要的地理位置, 又采用变点分析法分析了华北地震区历史破坏性地震 ($M \geq 4^{3/4}$) 的完整性及活动周期性.

变点分析广泛应用于以时间排序数据的分析, 如股票变化、历年赤字变化情况等 (Taylor, 2001). 变点分析法具有探测多变点问题的能力, 并且其能够提供变点发生的详细信息, 如变点发生的显著性水平及变点发生的置信区间. 将这一方法应用到华北地震区破坏性地震序列的分析中.

采用累计求和 (Cumulative sum, CUSUM) (Pettitt, 1980) 及自举程序法 (bootstrapping) (Efron and Tibshirani, 1993) 分析了华北地震区破坏性地震的周期性其完整性.

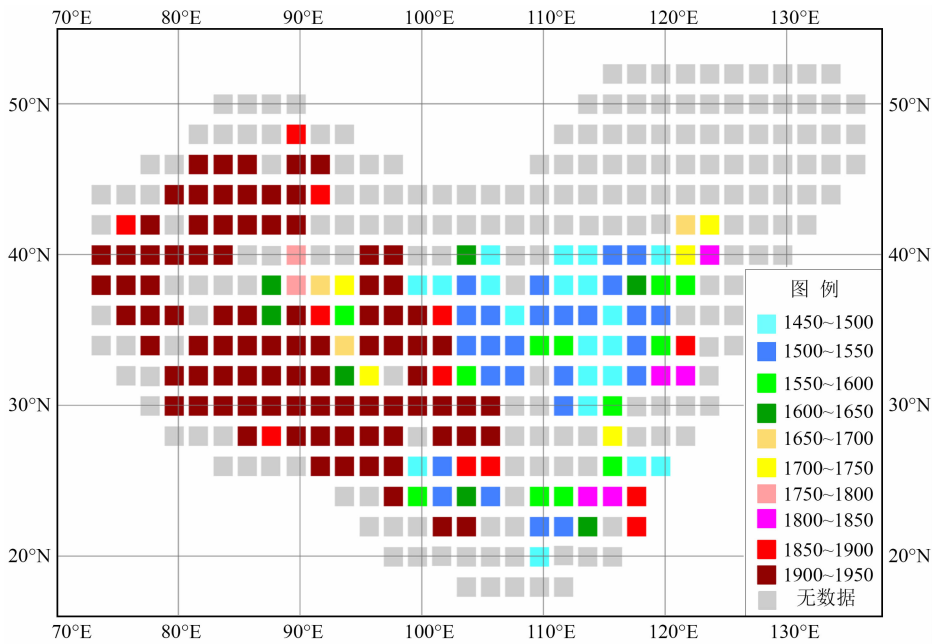


图 5 中国陆地历史地震 ($M \geq 5.0$) 完整性空间分布图

Fig. 5 Spatial distribution map of the completeness time T_c for the magnitude greater or equal to 5.0 of China mainland catalog

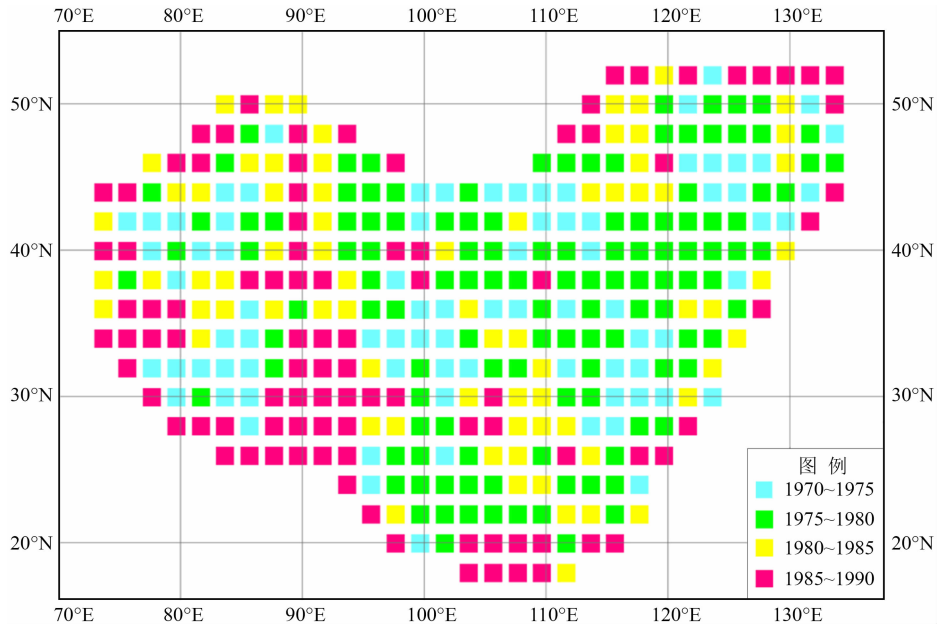


图 6 中国陆地仪器记录地震 ($M \geq 3.0$) 完整性空间分布图

Fig. 6 Spatial distribution map of the completeness time T_c for the magnitude greater or equal to 3.0 of China mainland catalog

首先根据华北地震区的历史强震,统计了公元 1200 年以来每 10 年尺度的地震发生频度,设为 R_1, R_2, \dots, R_N . 则累计和求取步骤为

$$\text{计算平均值 } \bar{R} = \frac{R_1 + R_2 + \dots + R_N}{N},$$

设零点处的累积和 $S_0 = 0$.

将前一个点的累积和加上该点值与平均值的差值取得当前点的累积和:

$$S_i = S_{i-1} + (R_i - \bar{R}) \quad i = 1, 2, \dots, N.$$

图 7b 为华北地震区 10 年尺度地震发生频度的累积和图,从图中可以看出地震发生频度至少有 3 个突变点. 将图 7a 和图 7b 对比发现每个突变点

表 2 各地震区不同起始震级目录完整
起始时间 T_c 及不确定性 T_u 和 T_l

Table 2 Completeness time T_c values and uncertainty T_u
and T_l values for seismic zones

地区	参数	$M=3.0$	$M=4.0$	$M=5.0$	$M=6.0$	$M=7.0$
东北	T_c	1976	1973	1930	1869	—
	T_l	1971	1970	1921	1813	—
	T_u	1983	1978	1941	1911	—
华北	T_c	1975	1973	1480	1410	1260
	T_l	1973	1971	1455	1350	1150
	T_u	1982	1982	1545	1500	1340
华南	T_c	1976	1974	1570	1500	1100
	T_l	1974	1972	1500	1420	910
	T_u	1985	1980	1680	1600	1380
新疆	T_c	1980	1975	1945	1935	1840
	T_l	1977	1973	1935	1910	1790
	T_u	1985	1978	1955	1950	1900
青藏高原	T_c	1984	1976	1940	1920	1900
	T_l	1978	1974	1930	1900	1870
	T_u	1987	1980	1960	1950	1930
川滇	T_c	1976	1975	1880	—	—
	T_l	1971	1970	1640	—	—
	T_u	1980	1979	1930	—	—

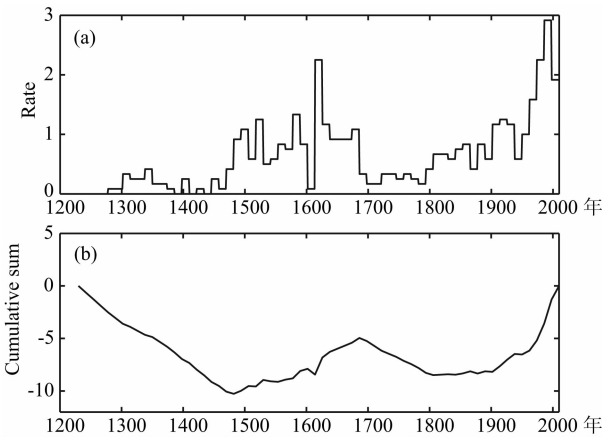


图 7 (a)华北地震区地震发生率
随时间变化曲线及(b)累积和图

Fig. 7 (a) Seismic rate as a function of time and (b) Cumulative
sum as a function of time for Northern China catalog

对应地震活动活跃期的结束或开始。

为了确认突变点确实存在,用自举程序法来求
取变点存在的置信水平. 首先将累积和变化程度的估
计值定义为 $S_{diff} = S_{max} - S_{min}$, 其中 $S_{max} = \max_{i=0, \dots, N} S_i$,
 $S_{min} = \min_{i=0, \dots, N} S_i$. 下面用自举程序法来确定这一变

化程度在多大的置信范围内. 对于自举程序法每次
抽样的步骤为

从原始的 $R_i, i = 1, 2, \dots, N$ 中有放回地随机抽
取 N 个值, 记为 $R_1^b, R_2^b, \dots, R_N^b$.

根据抽取的样本计算自举累积和, 记为 $S_1^b, S_2^b, \dots,$
 S_N^b . 计算累积和的最大值、最小值及其二者的差值
 $S_{max}^b, S_{min}^b, S_{diff}^b$. 查看求取的 S_{diff}^b 是否小于采用原始
数据计算得到的 S_{diff} .

本文中抽样次数为 1000 次. 采用自举程序法抽
取的样本可以用来代表随机记录的数据, 从而计算
在数据没有变点的情况下的累计求和值 S_{diff}^b , 通过
大量的自举样本可以计算出大量的 S_{diff}^b , 观察 S_{diff}^b
的分布情况. 若原始数据中没有变点, 那么根据原始
数据计算得到的 S_{diff} 与采用自举样本计算得到的
 S_{diff}^b 应该具有相当大小的值. 若原始数据具有变点,
则根据原始数据计算得到的 S_{diff} 应大于绝大多数自举
样本计算得到的 S_{diff}^b .

通过上述的 1000 次抽样可求得 1000 个 S_{diff}^b ,
那么变点存在的置信水平可用下式来计算:

$$\text{Confidence Level} = 100 \frac{n(S_{diff}^b \leq S_{diff})}{1000} \%$$

图 8 为 1000 个 S_{diff}^b 的概率分布图, 图中黑色竖
线为采用原始数据计算得到的 S_{diff} , 可以看出采用
自举程序法计算得到的 S_{diff}^b 绝大部分是小于 S_{diff}
的, 根据公式计算得到的置信水平为 99.6%. 这是
变点存在的一个很强的证据.

在确定了突变点的存在后, 是找准这个突变点
的位置. 文中采用计算平均均方误差的方法来判断
变点位置. 平均均方误差计算公式为

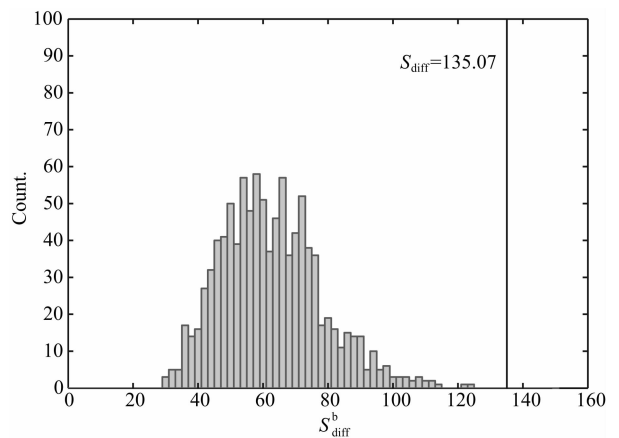


图 8 1000 次自举抽样求得的 S_{diff}^b 概率密度分布图
Fig. 8 Probability Density distribution of S_{diff}^b obtained
after the analysis of 1000 bootstrapping samples

$$MSE(m) = \sum_{i=1}^m (R_i - \bar{R}_1)^2 + \sum_{i=m+1}^N (R_i - \bar{R}_2)^2.$$

其中, $\bar{R}_1 = \frac{\sum_{i=1}^m R_i}{m}$, $\bar{R}_2 = \frac{\sum_{i=m+1}^N R_i}{N-m}$, m 为变点, 不断变化变点, 当使 MSE 值最小的 m 即为变点位置.

图 9 为华北地震区 MSE 值随时间变化的曲线, 在 $T=1480$ 年左右 MSE 取最小值. 这说明 $T=1480$ 年时华北地震区破坏性地震发生频度一个变点位置. 在寻找到一个变点后, 观测数据便以变点为分界点被分成两个部分. 对于每一部分数据可以重复上述步骤, 计算每一部分的变点位置及置信水平和置信区间. 表 3 为根据变点分析法求得的 3 个变点位置. 比较图中的 7(a, b), 可以发现 3 个变点分别对应华北地震区第一活跃期的起点、第一活跃期的结束点(平静期的起始点)及第二活跃期的起始点.

根据分析得到的变点, 比较了华北地震区第一活跃期与第二活跃期地震活动率. 图 10 为两个活跃期地震活动率的分布图, 采用 Mann-Whitney U-test (Weiss and Hassett, 1982) 分别检验了两个活跃期地震活动率中值和均值的差异性, 零假设为两个活跃期的地震活动率是相同的. 计算得到的 $p=0.53$, 这

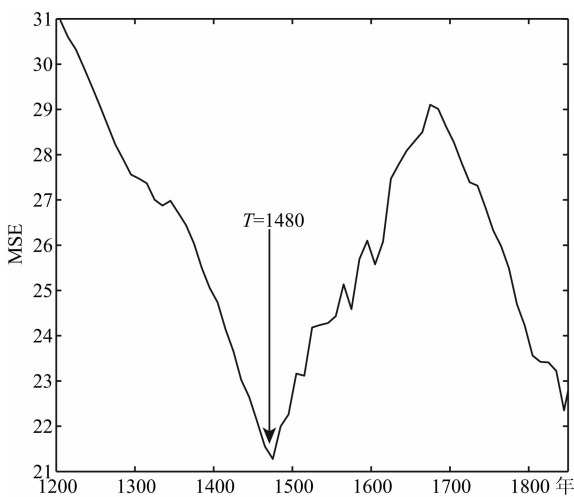


图 9 均方误差随时间的变化曲线
Fig. 9 MSE as a function of time

表 3 变点位置及其置信区间和置信水平

Table 3 The locations of changepoints and confidence interval, confidence level

序号	变点位置	置信区间	置信水平
1	1480	[1475—1490]	99.6%
2	1685	[1680—1690]	98.0%
3	1825	[1820—1830]	96.0%

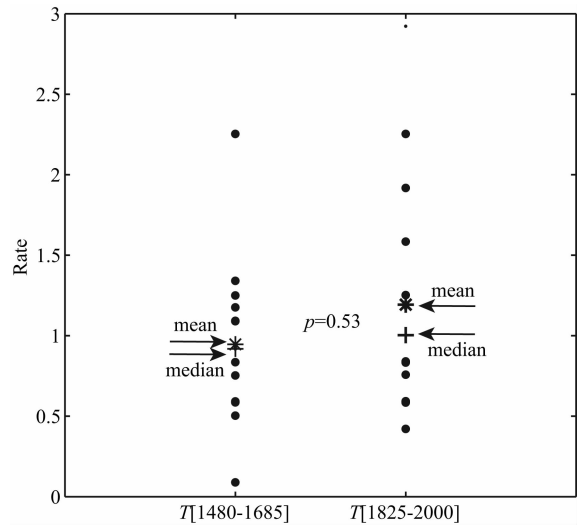


图 10 华北地震区两个活跃期历史地震地震活动率差异检验

Fig. 10 Statistical test for different of rates of two periods earthquake in Northern China

说明两个活跃期的地震活动率中值在显著性水平 $\alpha=0.05$ 下无明显差异, 或者说其差异无统计意义, 即无法拒绝零假设. 此外, 从图 10 中可以看出两个活跃期地震活动率的平均值差异要大于中值, 经检验 $p=0.06$, 在显著性水平为 0.05 下, 我们具有较弱的理由拒绝均值具有差异性. 然而通过观察两个活跃期地震活动率的分布情况, 发现采用中值进行比较比采用均值更加合理. 因此可以认为华北地震区两个活跃期地震活动率是没有差异的, 这也再次证明了华北地震区从第一活跃起来地震记录是基本完整的, 即公元 1480 年以来华北地震区破坏性地震目录是基本完整的.

6 结论与讨论

通过上述的研究认为, 对于仪器记录 ($M \geq 3.0$) 的地震目录: 中国东部大陆地区 1970—1975 年以后是基本完整的; 中国西部地区 1975—1980 年后基本完整. 对于历史地震 ($M \geq 5.0$) 地震目录: 中国大陆东部 1500—1550 年后可以认为是完整的, 这与黄玮琼等 (1994a) 结果是基本一致的; 西部 1900—1950 年后可以认为是完整的. 由于统计方法的不确定性受样本量的影响, 因此对于样本量大的仪器记录地震目录完整起始时间的不确定性较小, 而历史地震目录由于地震样本量较小, 不确定性则大一些.

考虑到华北地震区拥有丰富的历史地震目录, 并且华北地震区地震活动的非平稳性, 又采用突变

点法分析了华北地震区历史地震目录的完整起始时间. 结果表明华北地震区历史破坏性地震目录($M \geq 4\frac{3}{4}$) 在 1480 年左右以后是基本完整, 这与采用 Albarello 等(2001)方法得到的结果是一致的. 并且通过突变点分析对华北地震区地震活动的周期性有了定量的认识.

另外, 从图 5 和图 6 中可以看出我国地震目录完整起始时间的分布具有明显的空间差异, 对于仪器地震目录来说这可能跟每个地区地震仪器布设密度有关, 在地震仪器布设密度较大的地区, 地震监测能力强, 地震的完整起始时间就稍早一些, 反之则稍晚一些. 此外与仪器性能、人员素质等也有关. 对于历史地震目录, 在地理条件较好, 经济较发达人口较多的地区(如我国东部地区), 发生的地震能够被人们感知, 得以记录并保存下来, 因此地震目录的完整起始时间就早. 在地理环境恶劣的地区(如我国西部地区), 人口稀少, 社会发展较为落后, 发生的地震往往无法完全记录, 因此历史地震在 20 世纪中期以后才能完整记录. 此外战争、自然灾害等事件也会导致历史地震记录的丢失或者漏记.

文中以 $400 \text{ km} \times 400 \text{ km}$ 的网格和地震区为数据集单元, 分析了两个尺度上地震目录的完整性. 单纯从数学上来说, 数据集单元(研究区域)尺度的变化导致的是地震目录样本量的变化, 在样本量达到文中所要求的 40 的前提下, 数据集范围的变化不会导致结果有较大差异. 但是地震目录的完整性是具有空间差异的, 若选取的数据集范围地震目录记录能力空间差异较大, 则将对最终分析结果有明显影响. 因此文中为了表现地震目录完整性的空间差异, 采用的是 $400 \text{ km} \times 400 \text{ km}$ 的网格进行精细分析. 同时为了满足地震学和工程地震学中对地震区中地震目录完整性的需求, 又分析了地震区的地震目录完整性. 地震带中地震目录的完整性和其所处地震区的地震目录完整性基本相同, 差异在地震区地震目录完整性分析结果的误差范围内. 研究认为结合 $400 \text{ km} \times 400 \text{ km}$ 网格的地震目录完整性和各个地震区的地震目录完整性, 人们可基本确认所感兴趣的地区地震目录的完整起始时间.

地震目录的收集和编目具有较大的不确定性, 采用统计方法对地震目录的完整性进行研究是一种很有效的手段. 文中采用的统计分析方法不要求地震目录满足何种分布, 不依赖地震目录的 $G-R$ 关系拟合优度, 不需要用到记录地震事件的台站位置. 并且该方法对于分析历史地震目录的完整性尤其适

用. 本文的研究得到了中国大陆地震目录的完整起始时间及其不确定性. 这些结果对于地震活动规律研究和工程应用具有重要意义. 这些结果也可以供其他学者进行对比研究使用, 得到的完整的地震目录可以作为地震学研究的基础资料.

References

- Albarello D, Camassi R, Rebez A. 2001. Detection of space and time heterogeneity in the completeness of a seismic catalog by a statistical approach: An application to the Italian area. *Bull. Seism. Soc. Am.*, 91(6): 1694-1703.
- Bath M. 1983. Earthquake data analysis: an example from Sweden. *Earth Sci. Rev.*, 19(3): 181-303.
- Caputo M, Postpischl D. 1974. Contour mapping of seismic areas by numerical filtering and geological implications. *Ann. Geofis.* 27: 619-640.
- Department of Earthquake Disaster Prevention, National Earthquake Administration. 1995. The Catalogue of Chinese Historical Strong Earthquakes (23th century BC—1911 AD) (in Chinese). Beijing: Seismological Press.
- Department of Earthquake Disaster Prevention, China Earthquake Administration. 1999. The Catalogue of Modern Earthquakes in China (A. D. 1912~A. D. 1990, $M_s \geq 4.7$) (in Chinese). Beijing: China Science and Technology Publishing House
- Efron B, Tibshirani R J. 1993. An Introduction to the Bootstrap. New York: Chapman & Hall.
- Frankel A, Mueller C, Barnhard T, et al. 1996. National Seismic-Hazard Maps: Documentation June 1996. U. S. Geol. Surv. Open-File Rept, 96-532
- Gardner J K, Knopoff L. 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed, poissonian? *Bull. Seism. Soc. Am.*, 64(5): 1363-1367.
- Gutenberg B, Richter C F. 1944. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.*, 34(4): 185-188.
- Huang W Q, Li W X, Cao X F, et al. 1994a. Completeness analysis of earthquake catalogs in China Main Land I. *Acta Seismologica Sinica* (in Chinese), 16(3): 273-280.
- Huang W Q, Li W X, Cao X F. 1994b. Completeness analysis for earthquake catalogs in China Main Land II. *Acta Seismologica Sinica* (in Chinese), 16(4): 423-432.
- Kafka A L, Levin S Z. 2000. Does the spatial distribution of smaller earthquakes delineate areas where larger earthquakes are likely to occur? *Bull. Seism. Soc. Am.*, 90(3): 724-738, doi: 10.1785/0119990017.
- Knopoff L, Gardner J K. 1972. Higher seismic activity during local night on the raw worldwide earthquake catalogue. *Geophysical Journal of the Royal Astronomical Society*, 28(3): 311-313.
- Lee W H K, Brillinger D R. 1979. On Chinese earthquake history: An attempt to model an incomplete data set by point process analysis. *Pure Appl. Geophys.*, 117(6): 1229-1257.

- Liu J, Chen Q F, Chen Y. 1996. Completeness analysis of the seismic catalog in North China Region. *Earthquake* (in Chinese), 16(1): 59-67.
- Mignan A, Jiang C, Zechar J D, et al. 2013. Completeness of the mainland china earthquake catalog and implications for the setup of the china earthquake forecast testing center. *Bulletin of the Seismological Society of America*, 103(2A): 845-859.
- Mulargia F, Tinti S. 1985. Seismic sample areas defined from incomplete catalogues; An application to the Italian territory. *Phys. Earth Planet. Interiors*, 40(4): 273-300.
- Mulargia F, Gasperini P, Tinti S. 1987a. A procedure to identify objectively active seismotectonic structures. *Boll. Geofis. Teor. Appl.*, 29(114): 147-164.
- Mulargia F, Gasperini P, Tinti S. 1987b. Contour mapping of Italian seismicity. *Tectonophysics*, 142(2-4): 203-216.
- Pettitt A N. 1980. A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika*, 67(1): 79-84.
- Ren X M. 2011. Some problems of b value in seismic zoning. [Ph. D. thesis]. Beijing: Institute of Geophysics, China Earthquake Administration.
- Rotondi R, Meroni F, Zonno G. 1994. A different intensity recording for reducing the uncertainty in its assessment: An application to the completeness analysis of earthquake catalogues. *Nat. Haz.*, 10(1): 47-58.
- Rydelek P A, Sacks I S. 1989. Testing the completeness of earthquake catalogues and the hypothesis of self-similarity. *Nature*, 337(6204): 251-253.
- Stapp J C. 1971. An investigation of earthquake risk in the Puget Sound area by use of the type I distribution of largest extreme [Ph. D. Thesis]. Pennsylvania State University, 131.
- Su Y J, Li Y L, Li Z H, et al. 2003. Analysis of minimum complete magnitude of earthquake catalog in Sichuan-Yunnan region. *Journal of Seismological Research* (in Chinese), 26(Sup): 10-16.
- Taylor W A. 2001. A pattern test for distinguishing between autoregressive and mean-shift data. *Journal of Quality Technologies*, 31: 256-268.
- Tinti S, Mulargia F. 1985a. Completeness analysis of a Seismic Catalog. *Ann. Geophys.*, 3: 407-414.
- Tinti S, Mulargia F. 1985b. An improved method for the analysis of the completeness of a Seismic Catalogue. *Lettere Al Nuovo Cimento Series 2*, 42(1): 21-27.
- Wang H T, Li Y Z, Tu H W. 2006. Analysis on integrity of Xinjiang historical earthquake catalogue. *Inland Earthquake* (in Chinese), 20(1): 10-17.
- Wang X Q, Pan J Y. 2002. Completeness analysis of the earthquake catalog for East China region. *Journal of Seismology* (in Chinese), 22(3): 1-4.
- Weiss N A, Hassett M J. 1982. *Introductory Statistics*. Reading, MA: Addison-Wesley Publishing Co.
- Wiemer S, Wyss M. 2000. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the Western United States, and Japan. *Bull. Seism. Soc. Am.*, 90(4): 859-869.
- Woessner J, Wiemer S. 2005. Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. *Bull. Seism. Soc. Am.*, 95(2): 684-698.
- Xu W J, Gao M T. 2012a. Seismic hazard estimate using spatially smoothed seismicity model as spatial distribution function. *Acta Seismologica Sinica* (in Chinese), 34(4): 526-536.
- Xu W J, Gao M T. 2012b. Calculation of upper limit earthquake magnitude for northeast seismic region of China based on truncated G-R model. *Chinese J. Geophys.* (in Chinese), 55(5): 1710-1717.

附中文参考文献

- 国家地震局震害防御司. 1995. 中国历史强震目录(公元前 23 世纪—公元 1911 年). 北京: 地震出版社.
- 黄玮琼, 李文香, 曹学锋等. 1994a. 中国大陆地震资料完整性研究之一: 以华北地区为例. *地震学报*, 16(3): 273-280.
- 黄玮琼, 李文香, 曹学锋. 1994b. 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. *地震学报*, 16(4): 423-432.
- 刘杰, 陈棋福, 陈颢. 1996. 华北地区地震目录完全性分析. *地震*, 16(1): 59-67.
- 任雪梅. 2011 地震区划中 b 值统计的若干问题研究[博士学位论文]. 北京: 中国地震局地球物理研究所.
- 苏有锦, 李永莉, 李忠华等. 2003. 川滇地区区域地震目录完整性最小震级分析. *地震研究*, 26(增刊): 10-16.
- 王海涛, 李莹甄, 屠泓为. 2006. 新疆历史地震目录完整性分析. *内陆地震*, 20(1): 10-17.
- 汪雪泉, 潘坚云. 2002. 华东地区地震目录的完全性分析. *地震学刊*, 22(3): 1-4.
- 徐伟进, 高孟潭. 2012a. 以空间光滑的地震活动性模型为空间分布函数的地震危险性分析方法. *地震学报*, 34(4): 526-536.
- 徐伟进, 高孟潭. 2012b. 根据截断的 G-R 模型计算东北地震区震级上限. *地球物理学报*, 55(5): 1710-1717.
- 中国地震局震害防御司. 1999. 中国近代地震目录(公元 1912—公元 1990 年). 北京: 科学技术出版社.

(本文编辑 汪海英)