

基于空间信息格网的极端洪水灾害期望损失估测

马树建^{1,2,3}, 张丽丽³

(1. 南京工业大学安全科学与工程学院, 江苏省城市与工业安全重点实验室, 江苏 南京 210009)

(2. 南京工业大学经济与管理学院, 江苏 南京 211816)

(3. 南京工业大学数理科学学院, 江苏 南京 211816)

[摘要] 分析了我国极端洪水灾害的社会属性和自然属性, 得到受灾区域社会经济数据空间展布格网和极端洪水自然水文特性格网, 通过社会经济数据空间展布格网和极端洪水特性格网的叠加, 建立极端洪水灾害损失空间信息格网, 利用极值理论和分割多目标风险方法, 得到极端洪水风险损失的分布及其条件期望值. 选取哈尔滨市1998年的极端洪水为例进行实证分析, 给出了相应的极端洪水风险损失估计.

[关键词] 极端洪水, 空间信息格网, 分割多目标风险方法, 期望损失

[中图分类号] X43 **[文献标志码]** A **[文章编号]** 1001-4616(2015)04-0086-09

Estimation of Expected Losses for Extreme Flood Disaster Based on Spatial Information Grid

Ma Shujian^{1,2,3}, Zhang Lili³

(1. Jiangsu Key Laboratory of Urban and Industrial Safety, College of Safety Science and Engineering,

Nanjing Tech University, Nanjing, 210009, China)

(2. School of Economics and Management, Nanjing Tech University, Nanjing, 211816, China)

(3. School of Physical and Mathematical Sciences, Nanjing Tech University, Nanjing, 211816, China)

Abstract: With support of the remote sensing and GIS, the social economics data and the character of extreme flood disaster spatial information grid are given through which the loss of extreme flood spatial information grid is allowed. Through this grid, this paper gives the loss in each unit, the expected value and the distribution using the extreme value theory and partitioned multi-objective risk method (PMRM). This paper selects the example of Harbin city and gives the estimation loss of extreme flood disaster using the spatial information grid in 1998. With support of Matlab simulation, this paper gives the estimation of loss for extreme flood disaster whose distribution is proved to be generalized Pareto distribution. The parameter estimations are showed by using the method of MLE.

Key words: extreme flood, spatial information grid, PMRM, expected losses

随着全球气候的变化和人类活动的加剧,“极端洪水”的概念也随着社会变化而变化,对其无法统一地定性和定量的认识^[1-2],必须根据实际情况进行分析.人们一般从损失金额、死亡人数、损失波及的范围、发生的频率、周期长短等特点方面加以衡量、标识,以区别于小范围、小损失、高频率、短周期的一般的极端洪水灾害^[3-4].各个国家和地区根据本国实际情况在不同历史时期对极端洪水风险进行定义和划分^[5].随着社会的发展,对极端洪水的界定也随之变化.但是,从本质上分析,不论是哪个阶段,哪个社会存在的极端洪水都应该呈现出的两个最基本的特点^[6-7]:(1)损失金额相当大;(2)发生的概率非常小.虽然极端洪水事故发生的频率非常低,但极端洪水事件所造成的损失却是非常巨大的.一次极端洪水可能导致大范围内众多建筑的毁坏,造成数亿、数十亿、甚至数百亿美元的损失.统计资料表明,20世纪80年代后期开始,自然灾害和人为灾祸发生的频率加大,平均每年高达250多起.尤其在20世纪90年代的10年里,自然灾害和人为灾祸所带来的损失大约相当于70年代和80年代的总和^[8-10].据SIGMA数据统计资料显示,1992年、1999年的保险损失

收稿日期:2014-10-12.

基金项目:国家自然科学基金青年基金项目(41101509)、教育部人文社会科学研究青年基金(12YJC630290)、江苏省高校自然科学基金项目(09KJB570002)、江苏省博士后科研资助计划(1501048A).

通讯联系人:马树建,博士,副教授,研究方向:灾害保险金融与服务供应链管理. E-mail: shujianma@njtech.edu.cn

额分别高达324亿美元和286亿美元. 进入21世纪以来, 仅在2001年, 自然灾害所导致损失就达120亿美元. 而其中的极端洪水灾害造成的损失就占了很大一部分. 而我国1998年长江、嫩江和松花江特大洪水造成的直接经济损失高达16亿美元. 极端洪水风险所造成的巨大损失已经威胁到人类社会的可持续发展.

从极端洪水的自然和社会特性的分析可知, 极端洪水的损失是由于极端洪水的自然属性(如极端洪水产生的气候框架、致灾因子、淹没水深、历时等)和极端洪水的社会属性(承灾体)共同作用的结果. 极端洪水的自然属性和社会属性都存在着空间分布的特性, 即在不同区域, 极端洪水的淹没水深、历时等自然特性分布是不同的; 在不同的受灾区域内, 极端洪水的承灾体也是有差异的. 所以, 极端洪水损失也具有空间分布的特性. 本文将从极端洪水自然分布特性和社会分布特性两方面进行分析, 分别建立基于空间信息格网的极端洪水自然和社会分布特性的空间信息格网^[11-12], 然后通过两个格网的叠加, 讨论极端洪水损失的时空估测问题, 利用广义的分割多目标风险方法和极值统计方法的结合, 给出极端洪水期望损失估计值.

1 极端洪水自然和社会特性的空间信息格网

1.1 极端洪水自然特性空间信息格网

极端洪水的自然属性可以用极端洪水的发生时间、淹没范围和水深分布等指标来表示. 格网思想已经在各个领域得到广泛应用和对模型概化的优越性, 同时考虑极端洪水淹没的范围广、损失大的特点. 所以, 采用空间信息格网研究极端洪水的自然属性是比较合理的. 极端洪水发生前的行为及可能的淹没范围的确定可以通过基于GIS与水动力学模型的洪水模拟演进模型来预测, 得到极端洪水淹没信息的空间信息格网. 目前已经有一些比较成熟的模型可以应用^[13-15]. 如基于GIS的二维洪水模拟模型和基于GIS格网模型的洪水淹没分析模型都是研究的热点. 通过基于GIS的洪水模拟演进模型的计算, 可以得到极端洪水格网中极端洪水的发生频率、淹没水深、淹没的范围和流速等洪水特性参数, 为极端洪水的损失估测提供了水文特性的空间分布信息格网.

由于格网本身对模型的概化的优越性, 同时考虑到与极端洪水演进和洪涝灾害损失评估模型能够更好的结合. 所以, 提出基于空间信息格网模型的淹没分析思想. 由DEM可以生成三角形的单元格网和任意多边形格网. 如图1和图2所示.

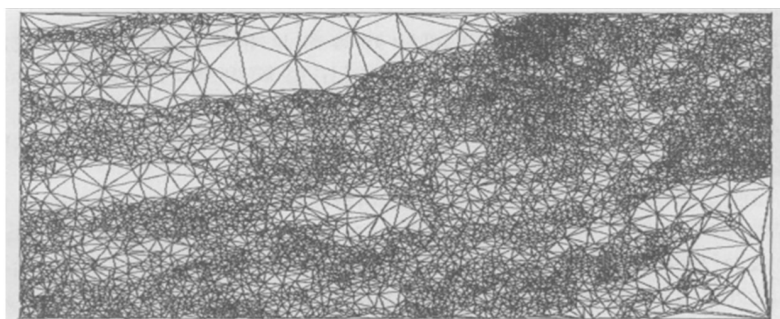


图1 三角形格网示意图

Fig.1 Triangle of spatial information grid

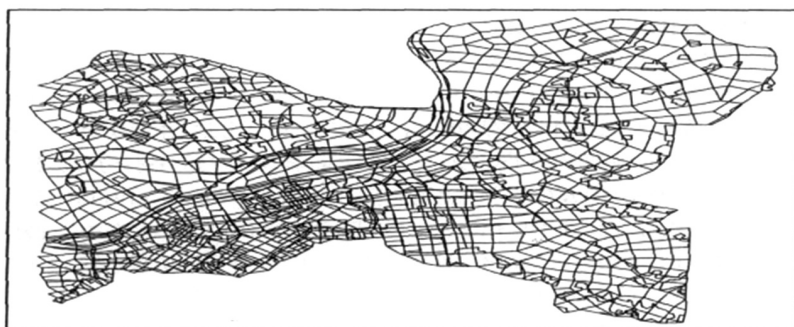


图2 任意多边形格网示意图

Fig.2 Random shape of spatial information grid

在极端洪水水文特性空间信息格网共有 M 个空间信息格网组成. 在空间区域上任意取某一格网单元 $l(l=1, 2, \dots, M)$, 此单元 l 在时间段 $[T_1, T_2]$ 内水文特性表示为 $Q_l^{[T_1, T_2]} = \{a_l^t, b_l^t, c_l^t, d_l^t\} \Big|_{t \in [T_1, T_2]}$, 其中 $a_l^t, b_l^t, c_l^t, d_l^t$

分别表示单元 l 在时间 t 极端洪水发生的频率、淹没的深度、淹没的范围和流速. 通过上述假设, 可以得到某一极端洪水发生区域 l 的极端洪水水文特性格网信息集为:

$$\{Q_l^{[T_1, T_2]}\}_{l=1,2,\dots,M} = \{Q_l^t = \{a_l^t, b_l^t, c_l^t, d_l^t\} \mid t \in [T_1, T_2]\}_{l=1,2,\dots,M}. \quad (1)$$

1.2 极端洪水社会特性空间信息格网

利用社会经济数据空间展布技术方法, 建立极端洪水社会特性的空间信息格网. 空间展布式社会经济数据库是通过收集不同区域以行政单元为单位的 socioeconomic 统计数据, 将其展布到对应的土地利用类型上, 使社会经济数据在空间分布上更加合理, 建立社会经济空间分布信息的 GIS 图层和相应的关系型数据库, 由 GIS 图层和关系型数据库构成了空间展布式社会经济数据库, 即可得到极端洪水社会特性的空间信息格网.

建立空间展布式社会经济数据库要区分极端洪水对城市和农村社会经济影响. 区域的地理位置和社会经济的发展促进了社会经济的分布特点, 使社会经济在空间上满足一定的自然规律, 从而出现了城市和农村. 社会经济分布密度较大, 防洪标准相对较高. 在极端洪水发生时, 城市是优先保护、免受灾害损失的重点区域. 极端洪水对城市和农村造成的损失具有不同的特点. 农村的建筑物主要是以农房为主, 抵御极端洪水侵袭的能力相对较差, 农村主要以农产品生产为主等等, 这些特点决定了极端洪水给农村带来的损失具有如下特点: 损失主要是以农产品受淹造成的减产损失为主, 农房损坏严重. 在城市极端洪水的损失主要是以居民和行政企业的财产损失和因洪水造成的企业停产、停工损失为主. 对于城市建筑物造成的损失相对较小. 根据上述分析, 要更加合理的估测极端洪水灾害损失, 更好地确定社会经济数据空间展布和损失率等问题, 有必要合理地农村和城市在行政单元和空间区域上给出合理的划分. 另一方面, 对于在同一个行政区域或空间区域上, 也有必要根据当地的经济水平等经济指标给出合理的划分.

根据上述分析, 社会经济统计数据是一个统计单元内经济数据的汇总或均值, 实际上社会经济数据在空间上的分布往往是不均匀的. 假设研究区域按照社会经济信息空间格网由 N 个格网单元区域组成. 在时空模型中, 社会经济指标可以看作是三维空间中的分布密度, 记为 $f=f(x, y, t)$, 满足密度函数的基本要求:

$$f(x, y, t) \geq 0, \quad \iiint_{\Omega} f(x, y, t) dx dy dt = 1.$$

$\Omega = \{(x, y, t) \mid (x, y) \in \mathbb{R}^2, t > 0\}$

式中, f 表示分布在时空上的分布密度, $f(x, y, t)$ 表示分布在时空上的社会经济指标的分布密度函数, (x, y) 表示区域所处的地理位置, t 表示时间. 所以, 在一定的格网单元 $d(d=1, 2, \dots, N)$ 和时间范围 $[T_1, T_2]$ 内, 此区域格网单元 d 的社会经济指标为:

$$S_d^{[T_1, T_2]} = \iiint_{\substack{\Omega = \{(x, y, t) \mid (x, y) \in d, \\ t \in [T_1, T_2]\}}} f(x, y, t) dx dy dt. \quad (2)$$

这里, $S_d^{[T_1, T_2]}$ 表示格网单元 d 在时间段 $[T_1, T_2]$ 的社会经济指标值.

社会经济空间分布特性主要决定于社会经济指标的分布密度函数, 要准确地确定社会经济指标的分布密度函数是不可能的. 通常的做法是通过典型区域在一定的时间范围内的社会经济指标调研抽样, 再延伸到其他类似的区域. 通过对社会经济数据的空间展布技术来获取社会经济数据的空间分布密度函数. 这是目前通常采用的技术方法.

通过上述方法, 可以得到社会经济数据的空间分布密度, 最后通过下式可以得到研究区域在某一时段的社会经济指标:

$$\{S_d^{[T_1, T_2]}\}_{d=1,2,\dots,N} = \left\{ \iiint_{\Omega = \{(x, y, t) \mid (x, y) \in d, t \in [T_1, T_2]\}} f(x, y, t) dx dy dt \right\}_{d=1,2,\dots,N}. \quad (3)$$

根据上述分析, 极端洪水灾害的社会特性信息格网即社会经济指标信息库不仅仅存在着空间上的分布差异, 在时间上也是变化的. 所以, 极端洪水损失估测需要考虑一个动态的具有时间序列的空间展布式的社会经济数据库.

1.3 极端洪水灾害损失率的计算

极端洪水灾害损失率是指社会资产受到极端洪水侵袭时遭受的毁坏程度. 它可以看作是极端洪水损失模型中的一个参数. 极端洪水灾害损失率包括分行业损失率、面上综合损失率和损失增长率等. 确定极端洪水灾害损失率的方法有很多. 目前比较流行的方法有典型区域现场调查方法、非线性多因素统计回

归模型方法等.这里,为了研究问题的方便,采用非线性多因素回归模型方法.

极端洪水损失率是由不同因素作用的结果,比较难获得的一个参数.不同的区域、行业、时间段甚至不同的致灾环境下损失率都是不同的.可以根据极端洪水损失率的典型区域的调查数据,建立适当的回归模型,得到分行业的极端洪水损失率.影响极端洪水损失率的因素主要包括极端洪水特性因子 W (包括淹没水深、流速等)和承灾体种类因子 C (包括社会资产、各种农作物等),极端洪水损失率 R 与两因子呈现指数关系.这里作适当的扩展,表达公式如下:

$$R = \beta \cdot W^\alpha \cdot C^\gamma + \varepsilon_i, \quad (4)$$

式中 α, β, γ 为待估参数, ε_i 为扰动项,满足 $E[\varepsilon_i] = 0, \text{Var}[\varepsilon_i] = \sigma^2$.

通过上述方法,对特定区域,可以预先根据极端洪水的水文特性建立不同的社会经济指标的极端洪水损失率数据库,为基于空间信息格网的极端洪水的损失估测服务,即由空间信息格网获得社会经济指标类型和极端洪水的水文特性,通过搜索极端洪水损失率数据库建立每一单元格网上各类社会经济指标的损失值.

1.4 极端洪水损失空间格网的形成和损失估测

通过上述分析,针对极端洪水的水文特性和受灾区域社会经济指标,分别建立了基于GIS的极端洪水自然和社会特性空间信息格网模型.借助空间信息格网技术,建立极端洪水空间信息格网.首先对极端洪水的淹没范围作离散化处理,得到极端洪水水文特性格网.目前比较有效的获取极端洪水淹没范围的方法有两种:一种是通过遥感影像获得;另一种是通过洪水的演进模型和洪水的淹没分析模型获取.在每一个单元里,近似地认为是淹没水深相等.可以通过社会经济空间展布的方法,解决社会经济空间数据分布不合理的问题,得到社会经济数据空间格网.将极端洪水的水文特性空间信息格网和社会经济空间格网的叠加,得到极端洪水损失空间信息格网,为更加准确计算极端洪水损失提供更好的方法.

通过极端洪水水文特性空间信息格网和社会经济空间格网的叠加得到的极端洪水损失空间信息格网与一般意义上的格网是不一样的.它是基于GIS的具有空间拓扑结构和关系属性,同时具有水文特性和社会经济指标特性的格网.它具有高度的相关性,信息格网单元可以由三角形、四边形及任意的多边形组成.

通过极端洪水空间信息格网的建立,可以得到极端洪水的水文特性,如极端洪水发生的时间、频率、淹没水深和范围、社会经济指标特征,再结合社会经济指标的种类,利用已经建立的极端洪水损失率数据库,即可估测极端洪水的损失.

假设研究区域的损失记为 D ,则由式(1)、(3)、(4)得:

$$D = \left\{ Q_i^{[r_1, r_2]} \right\}_{i=1,2,\dots,M} \oplus \left\{ S_d^{[r_1, r_2]} \right\}_{d=1,2,\dots,N} \otimes R. \quad (5)$$

其中运算符“ \oplus ”表示极端洪水的自然空间信息格网和社会信息格网的叠加,运算符“ \otimes ”表示叠加数据在极端洪水损失率数据库中选取适当的损失率作乘法.这里规定先进行“ \oplus ”运算,后进行“ \otimes ”运算.

下面将式(5)作适当的变形.基于空间信息格网的极端洪水损失评估主要考虑不同行业固定资产的损失和停工停产造成的财产损失;极端洪水水文特性主要考虑极端洪水的水深、淹没历时和预警时间体现的抗洪抢险行为的力度等.

假设下列变量和参数:

$R_{nkm} = (R_{1km}, R_{2km}, \dots, R_{nkm})^T$ 表示损失率向量,其中元素 R_{ikm} 表示第 i 种行业在第 k 级水深,第 m 级淹没历时的损失率.

$Y_{ns} = (Y_{1s}, Y_{2s}, \dots, Y_{ns})^T$ 表示损失率的折扣系数向量.其中元素 Y_{js} 表示第 j 种行业在第 s 级预警时间内损失率的折扣系数.

$A_{pm} = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{p1} & \dots & A_{pn} \end{pmatrix}$ 表示格网行业的资产数矩阵.其中 A_{ij} 表示在第 i 个格网单元内第 j 种行业资产

数量. δ_i 表示单元格网的类型,当单元格网为耕地、居民用地等已利用的土地类型时, $\delta_i = 1$;当单元格网为未利用的土地类型时, $\delta_i = 0$. e_j 表示第 j 种行业的年经济增长率, $e_j > 0$. 令 $\eta_j = (1 + e_j)^N$,其中 N 表示损失计算时间,即为到社会经济数据普查数据时间的年限.记 D_{ikm} 表示第 i 个格网单元内在第 k 级水深,第

m 级淹没历时,第 n 级预警时间内的损失,记 D 为研究区域的格网损失向量,即 $D = (D_{1km}, D_{2km}, \dots, D_{pkm})^T$,则在一定的时间,对某一研究区域,每个单元格网的极端洪水灾害经济损失评估模型如下:

$$D = \begin{bmatrix} D_{1km} \\ D_{2km} \\ \vdots \\ D_{pkm} \end{bmatrix}_{p1} = \begin{bmatrix} A_{11}\eta_1\delta_1 & A_{12}\eta_2\delta_1 & \cdots & A_{1n}\eta_n\delta_1 \\ A_{21}\eta_1\delta_2 & A_{22}\eta_2\delta_2 & \cdots & A_{2n}\eta_n\delta_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}\eta_1\delta_p & A_{p2}\eta_2\delta_p & \cdots & A_{pn}\eta_n\delta_p \end{bmatrix} \times [R_{nkm} \cdot Y_{ns}] = \begin{bmatrix} A_{11}\eta_1\delta_1 & A_{12}\eta_2\delta_1 & \cdots & A_{1n}\eta_n\delta_1 \\ A_{21}\eta_1\delta_2 & A_{22}\eta_2\delta_2 & \cdots & A_{2n}\eta_n\delta_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1}\eta_1\delta_p & A_{p2}\eta_2\delta_p & \cdots & A_{pn}\eta_n\delta_p \end{bmatrix}_{pn} \times \begin{bmatrix} R_{1km}Y_{1s} \\ R_{2km}Y_{2s} \\ \vdots \\ R_{nkm}Y_{ns} \end{bmatrix}_{n1}.$$

通过空间信息格网技术,将某一研究依据极端洪水的水温特性和社会经济的展布技术分割成若干个单元信息格网.在每个单元信息格网中,准确找出此单元信息格网所分布的社会经济展布空间信息,考虑损失率与极端洪水的淹没水深和淹没历时相关,同时考虑损失率的时间价值和由于预警时间而导致的损失率的折扣率等问题.根据上式,可以计算出在考虑研究区域的每个单元空间信息格网上极端洪水的损失值.

2 极端洪水灾害的期望损失计算

2.1 分割多目标风险方法与极值统计

分割多目标风险方法(PMRM)是一种风险分析方法,它主要来求解概率性质的多目标问题.

令 X 是 $r.v.$,代表损失值或者损失的数量.对于任意一个策略 $s_j, j=1, \dots, q$, 随机变量其边际概率密度函数为 $p_X(x, s_j)$, 满足概率密度函数的两个条件,建立损失概率与损失值或损失数量的关系.其累积分布函数为

$$P_X(x, s_j) = \int_{-\infty}^x p_X(y, s_j) dy, \quad j=1, \dots, q.$$

规定对于所有非空区间 $[x_1, x_2]$, 存在唯一反函数 p_X^{-1} .

PMRM将概率轴划分为一个 n 个区间的集合,其中 n 的选择依赖于决策问题和过程的特性.对于极端洪水的损失值,重点关注极端区间. α_i 表示概率轴上的分割点.对于每一个概率值 α_i 和每个策略 $s_j, j=1, \dots, q$, 都存在唯一的损失值 $\beta_{i,j}$, 使得 $p_X(\beta_{i,j}, s_j) = \alpha_i$. 如图3所示.

损失条件期望为:

$$g_i(s_j) = E[X | X \in (\beta_{i,j}, \beta_{i+1,j})] = \frac{\int_{\beta_{i,j}}^{\beta_{i+1,j}} xp_X(x, s_j) dx}{\int_{\beta_{i,j}}^{\beta_{i+1,j}} p_X(x, s_j) dx} = \frac{\int_{\beta_{i,j}}^{\beta_{i+1,j}} xp_X(x, s_j) dx}{\alpha_{i+1} - \alpha_i} \triangleq \frac{\theta_i}{\theta_i}.$$

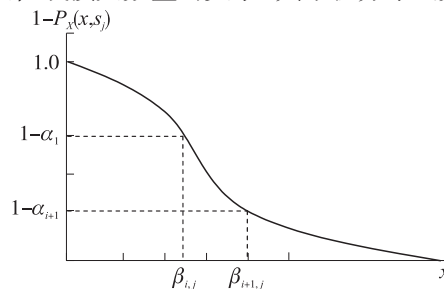


图3 概率轴的分割点在损失轴上的映射图

Fig.3 The reflection of the loss based on the probability

极值统计是测量极端条件下风险损失的一种常用的方法.它具有超样本数据的估计能力,并且可以准确刻画损失分布的尾部特征.在处理特殊条件下的极值问题时,通常采用极值分析的方法建模,在极值分析中主要有两类模型:一类是极值理论模型(EVT),极值理论模型主要是针对观察数据中的最大或最小值建模,即区间最大方法(Block Maxima Method, BMM).另外一类极值分析方法是广义 Pareto 分布模型(GPD),此模型也被称为 POT 模型(Peak-Over-Threshold Model),此模型主要是对观察值中所有超过某一个较大门限值的数据进行建模.广义 Pareto 分布模型能有效地使用有限的极端观察值,在实践中具有广泛的应用价值.

根据上述结论,阈值的选择是非常重要的,不合理的阈值的可能会造成模型的拟合产生巨大的误差.适当选择门限值 u , 可以通过两种方法确定.一种方法是 Hill 图法,一种是经验平均超出函数图法.

假设 $X_i, i=1, \dots, n$ 是取自某一分布总体的一个样本,将其按照大小顺序 $X_{(1)} \leq \dots \leq X_{(n)}$ 称 $\{X_{(1)}, \dots, X_{(n)}\}$ 为次序统计量, $X_{(1)} = \min\{X_{(1)}, \dots, X_{(n)}\}$ 为样本极小值, $X_{(n)} = \max\{X_{(1)}, \dots, X_{(n)}\}$ 为样本极大值,极大值和极小值的分布称为极值分布.

Hill 图法,采用样本的超限期望图确定阈值 u , 令 $X_{(1)} < X_{(2)} < \dots < X_{(n)}$, 样本的超限期望函数定义为:

$$e(u) = \frac{\sum_{i=k}^n (X_i - u)}{n - k - 1}, \quad k = \min\{i | X_i > u\}.$$

超限期望图为点 $(u, e(u))$ 构成的曲线,选取充分大的 u 作为阈值,使得当 $x \geq u$ 时 $e(x)$ 为近似线性函数.另外,当 $x \geq u$ 时,若超限期望图是向上倾斜,表明数据来源于参数 ξ 为正的

GPD分布;若向下倾斜,表明数据来源于尾部较短的分布;若是水平的,则表明该数据来源于指数分布.这一判断方法是根据广义 Pareto 分布在参数 $\xi < 1$ 时,它的超限期望函数 $e(m)$ 是一个线性函数得到的:

$$e(m) = E(X - m | X > m) = \frac{\sigma + \xi m}{1 + \xi}, \quad \sigma + \xi m > 0.$$

可验证极端洪水灾害超过门限值的损失数据服从 GPD 分布而且形状参数 $\xi > 0$. 对于极端条件下得到的极端洪水数据,适当选择门限值 u , 假设极端洪水损失为一个的 $r.v.X$, 则极端洪水的风险的超出门限值 u 的损失分布函数 $F_u(y)$ 可以表示为:

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(y + u) - F(u)}{1 - F(u)}.$$

这里 $0 \leq y < x_0 - u, x_0 \leq +\infty$, 表示为分布 F 的右端点. 对于极端洪水超出损失分布满足如下的极限定理.

定理 1 (Pickands (1974)) 对于一大类(几乎包括所有的常见分布)超出损失分布函数 $F_u(y)$, 存在函数 $\beta_u(y)$, 使得

$$\lim_{u \rightarrow x_0, 0 \leq y \leq x_0 - u} \sup |F_u(y) - G_{\xi, \beta(u)}(y)| = 0.$$

由定理 1 可知, 损失分布 $F_u(y)$ 的极值分布为 GPD. 其模型如下:

$$G_{\xi, \beta}(x) = \begin{cases} 1 - \left[1 + \xi \frac{x}{\beta}\right]^{-\frac{1}{\xi}}, & \xi \neq 0, \\ 1 - \exp\left(-\frac{x}{\beta}\right), & \xi = 0. \end{cases}$$

当 $\xi \geq 0$ 时, $x \geq 0$; 当 $\xi < 0$ 时, $v < x \leq -\frac{\beta}{\xi}$. 其中 $\beta > 0$, β 称为分布的刻度参数. 当 $\xi \geq 0$, GPD 表示的是重尾分布 (Heavy-tailed Distribution), 称 $\frac{1}{\xi}$ 为尾指数, 称 ξ 是形状参数.

此 GPD 模型的概率密度函数为:

$$g_{\xi, \beta}(x) = \begin{cases} \frac{1}{\beta} \left[1 + \xi \frac{x}{\beta}\right]^{-\frac{1}{\xi}-1}, & \xi \neq 0, \\ \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right), & \xi = 0. \end{cases}$$

给定样本值 x_1, \dots, x_{N_u} , 通过下式的极大似然函数可以得到参数 ξ, β 的估计.

$$L(\xi, \beta | Y) = \begin{cases} -N_u \ln \beta - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_u} \ln\left(1 + \frac{\xi}{\beta} y_i\right), & \xi \neq 0, \\ -N_u \ln \beta - \frac{1}{\beta} \sum_{i=1}^{N_u} y_i, & \xi = 0. \end{cases}$$

其中, $y_i = (x_i - u)^+$ N_u 为大于门限值 u 的样本个数.

2.2 极端洪水灾害期望损失估测分析

通过前面部分的讨论, 观察区域被空间信息格网模型分割成若干单元格网. 我们取若干单元格网中某一个典型格网进行讨论. 假设典型格网 k 的极端洪水损失为 L_k , 由式(1)、(3)、(4)可知:

$$L_k = Q_k \oplus S_k \otimes R. \quad (6)$$

其中运算符“ \oplus ”表示极端洪水的自然空间信息格网和社会信息格网的叠加, 运算符“ \otimes ”表示叠加数据在极端洪水损失率数据库中选取适当的损失率作乘法. 这里规定先进行“ \oplus ”运算, 后进行“ \otimes ”运算. 通过式(6), 可以得到研究区域每个单元信息格网的极端洪水损失.

由于极端洪水的历史数据相对较少, 对于有限的观察数据, 采用广义 Pareto 模型分析方法. 假设在第 k 个典型单元格网内, 极端洪水发生的历史损失样本为: l_k^1, \dots, l_k^S , 其中 S 表示数据的数目. 根据极端洪水的特点, 假设极端洪水历史样本数据 l_k^1, \dots, l_k^S 均大于给定的门限值 u , 由定理 1 可假设极端洪水超出损失样本 $l_k^1 - u, \dots, l_k^S - u$ 服从 GPD 分布模型. 考察极端洪水损失分布的尾部特征, 选取适当大于门限值的损失 β_i , 研究在损失 β_i 的极端洪水损失分布函数:

$$\alpha_i \triangleq F_u(\beta_i) = P(X - u \leq \beta_i | X > u) \sim \text{GPD}, \quad (7)$$

$$\beta_i \triangleq F_u^{-1}(\alpha_i). \quad (8)$$

可知,当极端洪水在典型单位格网 k 上损失介于区间 $[\beta_i, \beta_{i+1}]$ 的条件期望损失为:

$$El_k^i = E[X | X \in (\beta_i, \beta_{i+1})] = \frac{\int_{\beta_i}^{\beta_{i+1}} x g_{\xi, \beta}(x) dx}{\int_{\beta_i}^{\beta_{i+1}} g_{\xi, \beta}(x) dx} = \frac{\int_{\beta_i}^{\beta_{i+1}} x g_{\xi, \beta}(x) dx}{\alpha_{i+1} - \alpha_i} = \frac{\int_{\beta_i}^{\beta_{i+1}} x g_{\xi, \beta}(x) dx}{\theta_i}.$$

这里 $\xi > 0$. 将 GPD 分布模型的概率密度函数代入上式,得:

$$\begin{aligned} El_k^i = E[X | X \in (\beta_i, \beta_{i+1})] &\triangleq \frac{\int_{\beta_i}^{\beta_{i+1}} x g_{\xi, \beta}(x) dx}{\theta_i} = \frac{\int_{\beta_i}^{\beta_{i+1}} x \frac{1}{\beta} \left(1 + \frac{\xi}{\beta} x\right)^{-\frac{1}{\xi}-1} dx}{\theta_i} = \frac{\int_{\beta_i}^{\beta_{i+1}} \left[\left(1 + \frac{\xi}{\beta} x\right) - 1\right] \left(1 + \frac{\xi}{\beta} x\right)^{-\frac{1}{\xi}-1} dx}{\xi \theta_i} = \\ &= \frac{\int_{\beta_i}^{\beta_{i+1}} \left(1 + \frac{\xi}{\beta} x\right)^{-\frac{1}{\xi}} - \left(1 + \frac{\xi}{\beta} x\right)^{-\frac{1}{\xi}-1} dx}{\xi \theta_i} = \frac{1}{\xi \theta_i} \left[\frac{1}{1 - \frac{1}{\xi}} \frac{\beta}{\xi} \left(1 + \frac{\xi}{\beta} x\right)^{1 - \frac{1}{\xi}} \right]_{\beta_i}^{\beta_{i+1}} - \frac{1}{1 - \frac{1}{\xi} - 1} \frac{\beta}{\xi} \left(1 + \frac{\xi}{\beta} x\right)^{1 - \frac{1}{\xi} - 1} \right]_{\beta_i}^{\beta_{i+1}} = \\ &= \frac{\beta}{\xi \theta_i} \left\{ \frac{1}{\xi - 1} \left[\left(1 + \frac{\xi}{\beta} \beta_{i+1}\right)^{1 - \frac{1}{\xi}} - \left(1 + \frac{\xi}{\beta} \beta_i\right)^{1 - \frac{1}{\xi}} \right] + \left[\left(1 + \frac{\xi}{\beta} \beta_{i+1}\right)^{-\frac{1}{\xi}} - \left(1 + \frac{\xi}{\beta} \beta_i\right)^{-\frac{1}{\xi}} \right] \right\} = \\ &= \frac{\beta}{\xi \theta_i} \left[\left(\frac{\beta + \xi \beta_{i+1}}{\beta} \right)^{-\frac{1}{\xi}} \left(\frac{\xi(\beta_{i+1} + \beta)}{\beta(\xi - 1)} \right) - \left(\frac{\beta + \xi \beta_i}{\beta} \right)^{-\frac{1}{\xi}} \left(\frac{\xi(\beta_i + \beta)}{\beta(\xi - 1)} \right) \right]. \end{aligned}$$

通过上式,可以得到在研究的典型单元格网内,极端洪水损失介于给定损失区间内的期望损失,根据极端洪水的历史数据,选择一定的方法,估计 GPD 分布中的参数 β, ξ , 再根据实际厘定费率模型中对极端洪水损失的控制范围,就可以计算出极端洪水的条件期望损失.

3 实例

以哈尔滨 1998 年极端洪水的遥感数据资料为例,极端洪水损失估测主要分为 3 个步骤.

(1) 建立极端洪水自然水文特性的空间信息格网,结合哈尔滨市的地形,可以分为江南主城区、松北开发区、群力新区、江中保护区和呼兰区 5 个格网. 以江南主城区这个信息格网为例,遥感监测淹没范围和淹没水深等情形,从遥感监测影像上提取江南主城区淹没范围,将遥感监测淹没范围与该地区的地形叠加计算,可以计算得到淹没范围内的水深分布等情况,该区域极端洪水的水文特性见表 1.

(2) 根据不同频率洪水的淹没范围、水深分布情况,江南主城区的资产类别主要划分为居民财产、机关企事业单位及基础设施等经济指标,建立哈尔滨市江南主城区空间展布式社会经济数据库,根据哈尔滨市 1998 年的统计年鉴,江南主城区的社会经济状况见表 2.

(3) 不同重现期的洪水淹没损失还取决于各类资产的洪水灾害损失率. 根据哈尔滨市历史洪水资料,采用回归分析的方法,考虑到经济变化等扰动因素,可得到在不同重现期内各类财产损失率. 如表 3 所示.

表 1 1998 年哈尔滨市江南主城区极端洪水水文特性表

Table 1 Hydrological characteristics of exterm flood in river south main city area of Harbin in 1998

频率/%	峰流量/(m ³ /s)	受淹面积/km ²	最高水位/m	最大受淹历时/d	最大淹没/m
2	16 600	36.57	120.89	45	6.2

表 2 1998 年哈尔滨市江南主城区社会经济状况表

Table 2 Socioeconomic status in river south main city area of Harbin in 1998

总人口(万人)	城市居民		机关、企事业单位		
	房屋		产业单位	法人单位	
	平房户数/户	楼房一层以下户数/户	营业收入/(万元/年)	固定资产/(万元)	营业收入/(万元/年)
95.762	48 162	11 857	987 436.5	2 063 514	5 137 564

表3 不同洪水频率下哈尔滨市江南主城区各类财产损失率

Table 3 Under the different flood frequency all type of property loss ratio in river south main city area of Harbin

洪水频率	工商企事业及基础设施损失率	城市居民财产损失率
5%	10.2	3.3
2%	17.1	6.7
1%	23.1	10.5
0.5%	32.6	14.9

依据空间信息格网技术,以及数据可得性原则,计算可以得到1998年哈尔滨市江南主城区的直接经济损失为140.66亿元.应用基于空间信息格网的极端洪水灾害损失评估方法,不仅能够得到洪水淹没范围、最大淹没水深以及具体的空间范围,而且可以详细地评估受灾人口数量、空间分布、各类经济财产损失数量及其空间属性,提高了评估的客观性和准确性.

通过上述的例子,利用空间信息格网技术评估了1998年特大洪水造成哈尔滨市江南主城区的直接经济损失,该应用说明空间信息格网技术应用于极端洪水灾害损失评估是可行的.假设哈尔滨市分为江南主城区、松江区和呼兰区三个空间信息格网,根据上述例子的结果,利用蒙特卡罗方法随机模拟极端洪水风险损失原始数据.图4~6给出了模拟数据相应的平滑直线图.

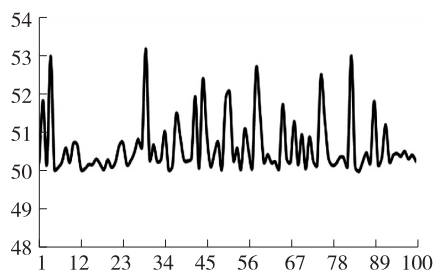


Fig.5 The smoothing linear of extreme flood losses in river north area of Harbin

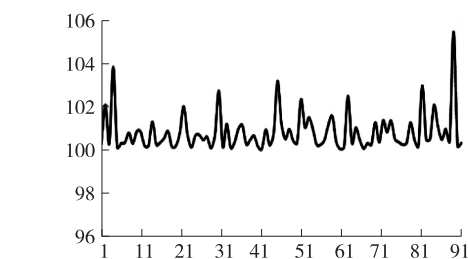


图4 江南主城区极端洪水数据的平滑直线图
Fig.4 The smoothing linear of extreme flood losses in river south main city area of Harbin

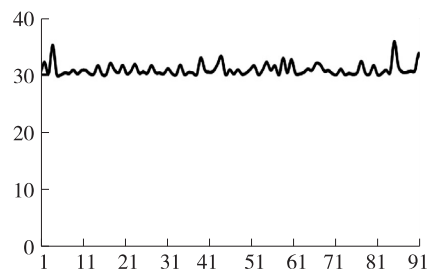


Fig.6 The smoothing linear of extreme flood losses in Hulan area of Harbin

将模拟产生的数据分别写出相对应的次序统计量根据Hill图法,采用样本的超限期望图确定阈值 u ,

令 $X_{(1)} > X_{(2)} > \dots > X_{(n)}$,样本的超限期望函数定义为: $e(u) = \frac{\sum_{i=k}^n (X_i - u)}{n - k - 1}$, $k = \min\{i | X_i > u\}$.超限期望图点 $(u, e(u))$ 构成的曲线,选取充分大的 u 作为阈值,使得当 $x \geq u$ 时 $e(x)$ 为近似线性函数.另外,当 $x \geq u$ 时,若超限期望图是向上倾斜的,表明数据来源于参数 ξ 为正的GPD分布;若向下倾斜,表明数据来源于尾部较短的分布;若是水平的,则表明该数据来源于指数分布如图7所示.

计算得三个空间格网上的阈值分别为 $u_1 = 101.2645$, $u_2 = 51.276$, $u_3 = 31.3673$.

根据图8~10可知,三个信息格网上随机产生的损失数据服从广义帕累托分布.

通过建立极大似然函数,进行参数估计,公式如下,

$$L(\xi, \beta | Y) = \begin{cases} -N_u \ln \beta - (1 + \frac{1}{\xi}) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\xi}{\beta} y_i \right), & \xi \neq 0, \\ -N_u \ln \beta - \frac{1}{\beta} \sum_{i=1}^{N_u} y_i, & \xi = 0. \end{cases}$$

Matlab编程利用搜索法可以得到GPD分布的参数:

$\xi_1 = 0.1220$, $\beta_1 = 0.6864$; $\xi_2 = 0.1726$, $\beta_2 = 0.5406$;

$\xi_3 = 0.1395$, $\beta_3 = 0.7695$.

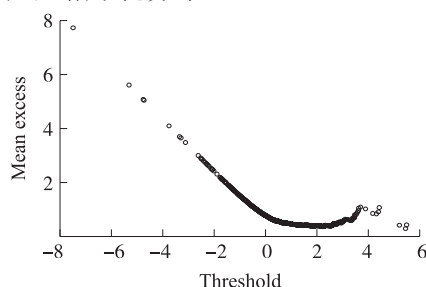


图7 超额均值函数图

Fig.7 The main excess function

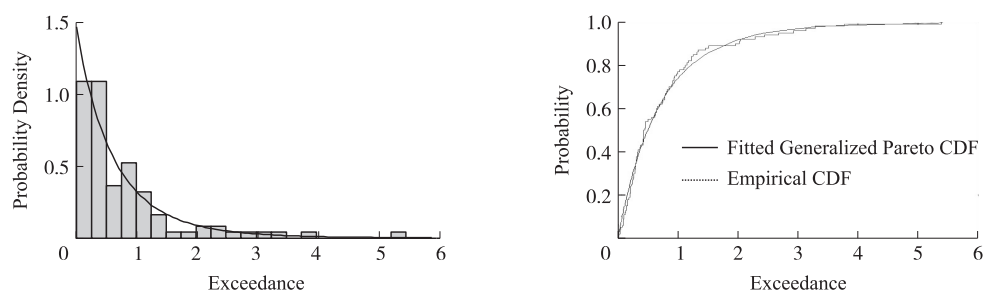


图8 江南主城区极端洪水极值分布与经验分布的比较(stdErr=0.109 3 0.101 4)

Fig.8 The comparison between the expirical distribution and extreme distribution in river south main city area of Harbin

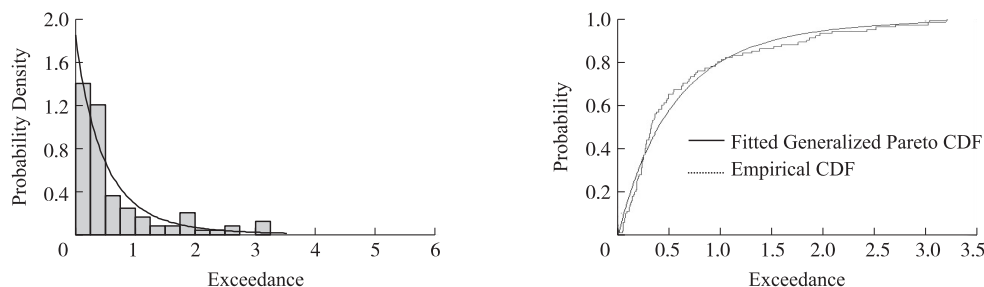


图9 松北区极端洪水极值分布与经验分布的比较(stdErr=0.126 5 0.086 7)

Fig.9 The comparison between the expirical distribution and extreme distribution in river north area of Harbin

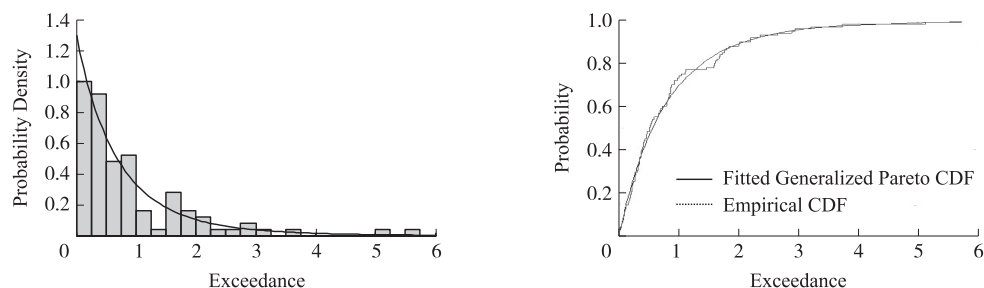


图10 呼兰区极端洪水极值分布与经验分布的比较(stdErr=0.1184 0.118 8)

Fig.10 The comparison between the expirical distribution and extreme distribution in Hulan area of Harbin

[参考文献]

- [1] 刘树坤. 国外防洪减灾发展趋势分析[J]. 水利水电科技进展, 2000, 20(1): 55-59.
- [2] 陈泮勤. 全球变暖对自然灾害的可能影响[J]. 自然灾害学报, 1996b, 5(2): 95-101.
- [3] 陈泮勤. 气候变化与自然灾害[J]. 自然灾害学报, 1996a, 5(2): 11-17.
- [4] 施雅风. 全球与中国变暖特征及未来趋势[J]. 自然灾害学报, 1996b, 5(2): 1-9.
- [5] HOUGHTON J T, DING Y H, GRIGGS D J, et al. Climate change 2001: the scientific basis [M]. Cambridge: Cambridge University Press, 2001: 100-120.
- [6] 高学杰. 中国地区极端事件预估研究[J]. 气候变化研究进展, 2007(3): 42-46.
- [7] GAO X J, ZHAO Z C, GIORGI F. Changes of extreme events in regional climate simulations over East Asia [J]. Advance atmosphere science, 2002, 19(5): 927-942.
- [8] 林蓉辉. 自然灾害与风险管理[J]. 灾害学, 1991(2): 88-91.
- [9] 魏一鸣, 范英, 余菊良, 等. 洪水灾害风险分析的系统理论[J]. 管理科学学报, 2001(2): 9-13.
- [10] 周寅康. 洪水灾害风险评价[J]. 大自然探索, 1995, 4(2): 67-72.
- [11] 邵振峰, 李德仁. 基于网格计算环境下的空间信息多级格网研究[J]. 地理信息世界, 2005(2): 35-39.
- [12] 谢正辉, 刘谦, 袁飞, 等. 基于全国 50 km×50 km 网格的大尺度陆面水文模型框架[J]. 水利学报, 2004(5): 76-82.
- [13] 金君, 李成名, 印洁, 等. 人口数据空间分布化模型研究[J]. 测绘学报, 2003(2): 92-96.
- [14] Federal Geographic Data Committee. United States national grid: FGDC-STD-011-2001 [S]. Reston, 2001.
- [15] WHITE D, KIMERLING A J, SAHR K, et al. Comparing area and shape distortion on polyhedral-based recursive partitions of the sphere [J]. International journal of geographical information science, 1998, 12(8): 805-827.

[责任编辑: 陈 庆]