

# 基于信息扩散原理的渤、黄海沿岸风暴潮 灾害风险分析\*

齐鹏<sup>1,2</sup>, 李明杰<sup>1,2,3</sup>, 侯一筠<sup>1,2</sup>

(1.中国科学院海洋研究所 青岛 266071; 2.中国科学院海洋环流与波动重点实验室 青岛 266071;  
3.中国科学院研究生院 北京 100049)

**提要** 根据收集的渤、黄海沿岸主要海洋站(塘沽、羊角沟、烟台、威海、乳山口、青岛和日照)1950年以来各年最高潮位资料(个别站位资料样本较小),应用能优化处理小样本的基于信息扩散原理的模糊风险评估模型,对渤、黄海沿岸各海洋站年最高潮位的多年一遇极值情况进行了超越概率计算,并与《港口工程技术规范》推荐的极值 I 型(Gumbel 曲线)方法进行比较。除羊角沟站误差 3%外,其它站误差均在 2%及以下,表明文中基于信息扩散原理的风险评估模型在兼顾小样本条件下所给结果是比较可靠的。这对于潮位等水文实测资料短缺的新建或拟建工程区进行工程水文要素设计重现期的推算将有较好的帮助。

**关键词** 风暴潮灾害, 模糊风险评估, 信息扩散, 渤、黄海沿岸

**中图分类号** P731

海洋潮位实测值一般包含两种成分:一为在天体引潮力作用下潮位的周期性变化,称为潮汐(天文潮);另一为在强烈的大气扰动(强风和气压骤变)或地震等外界强迫作用下潮位的非周期性异常变化——前者引起的沿岸海洋灾害称为风暴潮灾害,后者引起的称为海啸。在渤、黄海沿岸,由潮位的非周期性异常变化引起的海洋灾害主要是风暴潮灾害。

渤、黄海沿岸不仅时常受夏季北上台风风暴潮的侵袭,而且莱州湾和渤海湾等地区还频受主要发生在春、秋和冬三季的温带系统风暴潮的侵袭,是我国北方沿岸风暴潮灾害的频发区。1950—1990年间渤、黄海沿岸(主要集中于莱州湾和渤海湾)发生最大增水在 1.0 m 以上的温带系统风暴潮共有 521 次,其中最大增水在 2.0 m 以上的有 57 次,3.0 m 以上的有 3 次,造成严重风暴潮灾害的有 4 次,造成较大灾害的有 6 次(杨华庭等,1994)。在台风风暴潮灾害方面,根据相关文献(杨华庭等,1994;李培顺,1998;张晓慧等,2006)提供的数据,1949—2000年间影响青岛和山东半岛南

岸及至渤海沿岸的致灾台风风暴潮共计 13 次,其中造成上述地区严重灾害的有 5 次。值得注意的是在这 5 次严重级别的台风风暴潮灾害中,1980 年以后的 20 年里占了 4 次之多(其中的 8509、9216 和 9711 号台风风暴潮灾害的级别为特重),即平均约 5 年出现 1 次严重及以上级别的台风风暴潮灾害;而轻度及以上级别的台风风暴潮灾害 1980 年以后约有 10 次,即平均约 2 年出现 1 次。这表明,进入 80 年代后该地区台风风暴潮的成灾频率和灾害程度均呈明显增大趋势。究其原因主要与进入 80 年代后沿海地区国民经济开发和建设速度逐年加快,昔日荒芜的海岸和近岸海域,不断被码头、工厂、水产养殖场、旅游娱乐场所替代,即承灾体与年俱增有关(李培顺,1998)。显然,减少和防止风暴潮灾害损失对保障渤、黄海沿岸地区经济发展具有重要意义。而对于减灾防灾来说,其基础性的工作之一是进行风暴潮灾害风险评估。

本文在收集渤、黄海沿岸主要海洋站(塘沽、羊角沟、烟台、威海、乳山口、青岛和日照)1950 年以

\* 国家自然科学基金资助项目,40676018 号;山东 908 专项 SD-908-02-08 任务资助项目和国家 908 专项“海浪灾害及其对沿海地区社会经济发展影响综合评价”课题(2005—2010)资助项目。齐鹏,博士,副研究员,E-mail: pqi@ms.qdio.ac.cn

收稿日期:2010-03-10,收修改稿日期:2010-06-25

来实测的年最高潮位资料的基础上, 考虑个别站位资料样本较小, 引入能优化处理小样本的基于信息扩散原理的模糊风险评估模型(黄崇福等, 1998), 对渤、黄海沿岸上述各海洋站年最高潮位的多年一遇极值情况进行了超越概率计算。为验证基于信息扩散方法结果的可靠性, 又与《港口工程技术规范》推荐的极值 I 型(Gumbel 曲线)方法的结果进行了比较。

## 1 基于信息扩散原理的灾害模糊风险评估方法

### 1.1 关于风暴潮灾害风险的定义

以灾害学领域的观点, 自然灾害风险被认为是自然灾害事件发生的可能性及由其造成后果的严重程度(史培军, 1996)。自然灾害风险评估被定义为估计灾害事件发生的可能性大小及损失大小的过程。

对沿海某一特定区域而言, 当风暴增水使潮位越过警戒水位后, 就有可能发生风暴潮灾害。风暴潮灾害风险是指未来若干年内可能达到的灾害程度及其发生的可能性。对风暴潮灾害风险的数学描述(黄崇福, 2001)如下:

设  $X$  为风暴潮灾害指标,  $T$  年内关于  $X$  的超越概率分布定义为风暴潮灾害风险。例如, 设

$$X = \{x_1, x_2, \dots, x_n\} \quad (1)$$

又设强度超越  $x_i$  的概率为

$$p_i = P(x \geq x_i), \quad i=1, 2, \dots, n \quad (2)$$

则概率分布

$$P = \{p_1, p_2, \dots, p_n\} \quad (3)$$

称为风暴潮灾害的风险。

### 1.2 信息扩散原理和灾害风险评估模型

“信息扩散”是为了弥补信息不足而考虑优化利用样本模糊信息的一种对样本进行集值化的模糊数学处理方法。该方法可将只有一个观测值的样本, 变成一个模糊集。或者说, 是把单值(single-valued)样本, 变成集值(set-valued)样本。也就是说, 当样本不多时, 所有样本提供认识风险的信息是不完善、不完备的, 具有模糊的不确定性。因此, 不应把单个样本的信息仅看作一个确切的观测值, 而应看作一个模糊集的观测样本, 即对样本进行集值化(黄崇福等, 1998), 以部分弥补资料不足的缺陷, 达到提高风险估计精度的目的。

设所讨论的灾害现象的论域为

$$U = \{u_1, u_2, \dots, u_n\} \quad (4)$$

一个单值观测样本  $s_j$  ( $j=1, 2, \dots, m$ , 这里  $m$  为观测样本的个数)按下式可将其携带的信息扩散给  $U$  中的所有点。

$$f_j(u_i) = \frac{1}{\sqrt{2\pi}k} \exp\left[-\frac{(s_j - u_i)^2}{2k^2}\right] \quad (5)$$

式中,  $k$  为扩散系数, 可根据样本集中的最大样本值( $s_{\max}$ )、最小样本值( $s_{\min}$ )和样本个数  $m$  按下式计算得到:

$$k = \begin{cases} 1.6987(s_{\max} - s_{\min})/(m-1) & 1 < m \leq 5 \\ 1.4456(s_{\max} - s_{\min})/(m-1) & 6 \leq m \leq 7 \\ 1.4230(s_{\max} - s_{\min})/(m-1) & 8 \leq m \leq 9 \\ 1.4208(s_{\max} - s_{\min})/(m-1) & 10 \leq m \end{cases} \quad (6)$$

令

$$C_j = \sum_{i=1}^n f_j(u_i) \quad (7)$$

并构造模糊子集的隶属函数

$$\mu_{s_j}(u_i) = f_j(u_i)/C_j \quad (8)$$

从而实现将单值样本  $s_j$  变成以  $\mu_{s_j}(u_i)$  为隶属函数的模糊子集, 并称  $\mu_{s_j}(u_i)$  为样本  $s_j$  的归一化信息分布。

再令

$$q(u_i) = \sum_{j=1}^m \mu_{s_j}(u_i) \quad (9)$$

其含意为: 对观测样本集  $\{s_1, s_2, \dots, s_m\}$ , 经上述信息扩散公式推算, 在灾害现象论域范围里, 考虑所有样本代表  $s_j$  ( $j=1, 2, \dots, m$ ), 论域中  $u_i$  点上的样本个数应为  $q(u_i)$  个。显然,  $q(u_i)$  是一个不小于零的实数。

令

$$Q = \sum_{i=1}^n q(u_i) \quad (10)$$

事实上,  $Q$  就是各  $u_i$  点上的样本数的总和。从理论上讲, 必有  $Q = m$ , 但实际计算中因四舍五入带来的误差, 使得  $Q$  与  $m$  之间略有差别。于是,

$$p(u_i) = q(u_i)/Q \quad (11)$$

就是样本落在论域中  $u_i$  点上的几率, 作为概率的估计值。

显然, 超越  $u_i$  的概率应为:

$$P(u \geq u_i) = \sum_{l=i}^n p(u_l) \quad (12)$$

$P(u \geq u_i)$  就是所要求的风险估计值。

上述信息扩散原理、方法可按以下计算流程实现(图 1)。

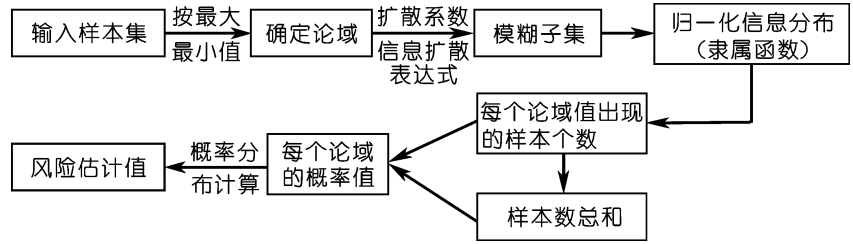


图 1 信息扩散方法计算流程

Fig.1 Flow diagram of the information diffusion computation

## 2 渤、黄海沿岸年最高潮位多年一遇极值的超越概率计算

沿渤、黄海沿岸选择塘沽、羊角沟、威海、烟台、乳山口、青岛和日照 7 个海洋站作为代表站, 收集了 1950 年以来各站实测的年最高潮位资料。所收集的各站实测资料样本数差别较大, 其中样本数较大的站有塘沽(50 个)、日照(42 个)、青岛(41 个)和羊角沟(41 个); 样本数较小的站有威海(19 个)、烟台(25 个)和乳山口(25 个)。按以往交通部《港口工程技术规范》推荐的 Gumbel 曲线方法进行概率计算, 其对样本数有较高的要求。对于资料样本不足信息不够完备的情况, 这里应用上节介绍的信息扩散原理和方法进行超越概率计算。

首先, 综合考虑各站资料样本集中的最小值和最大值提出各站统一的论域  $U$ 。由于所收集的各站实测潮位值均是对于各站当地潮位基准面而言, 因而需要首先将各站的潮位值统一到黄海面上。塘沽、羊角沟、威海、烟台、乳山口、青岛和日照各站潮位基准面的 85 黄海高程分别为 -257、-289、-122、-211、-243、-239 和 -259 cm。于是, 将各站潮位数据修正到黄海面上后它们在信息扩散方法中的统一论域  $U$  确定为 [100, 400], 并通过等间隔离散, 使之成为离散论域。为使模型计算获得较高精度, 这里离散间隔取 1 cm。论域  $U$  离散为

$$U = \{u_1, u_2, \dots, u_n\} = \{100, 101, \dots, 399, 400\} \quad (13)$$

以下为便于各站之间进行风险水平比较, 潮位水平为从黄海面起算。图 2 比较显示了信息扩散方法得到的各海洋站年最高潮位的超越概率估计曲线。由图 2 可获知, 对于横坐标上不同潮位水平各站发生风险的超越概率估计值。对图 2 中概率值低于 0.2 的部分进行放大显示, 得到图 3。由图 3 容易读出各海洋站不同重现期对应的潮位。例如, 对于 321 cm 潮位水平(按黄海面起算), 它是羊角沟站 10 年一遇的潮

位水平, 对于塘沽和日照来说它是 50 年一遇的潮位水平, 而对于其它各站其风险概率水平则是零。

## 3 信息扩散方法的验证

上节给出了基于信息扩散方法渤、黄海沿岸年最高潮位多年一遇极值情况的超越概率计算。为验证信息扩散方法结果的可靠性, 作者也应用交通部《港口工程技术规范》中采纳的 Gumbel(耿贝尔)曲线方法对文中各海洋站多年一遇潮位情况进行了概率计算。由于本节重点在于对信息扩散方法可靠性进行验证, 而不在于各站之间风险水平比较, 同时, 为便于与前人结果进行比较, 这里不对各站实测数据进行统一黄海水准面的修正, 即各站样本中的年最高潮位值仍按各站对于当地潮位基准面起算的潮高。一般地说, Gumbel 方法要求样本序列至少具备几十年的观测资料, 因而文中的塘沽、日照、青岛和羊角沟站的资料样本数均能较好地满足这一要求, 而其余各站的样本数对于 Gumbel 方法来说是不够理想的。表 1 是应用上述两种方法得到的各站不同重现期潮位的概率推算结果的比较。两种方法结果的绝对误差容易从表 1

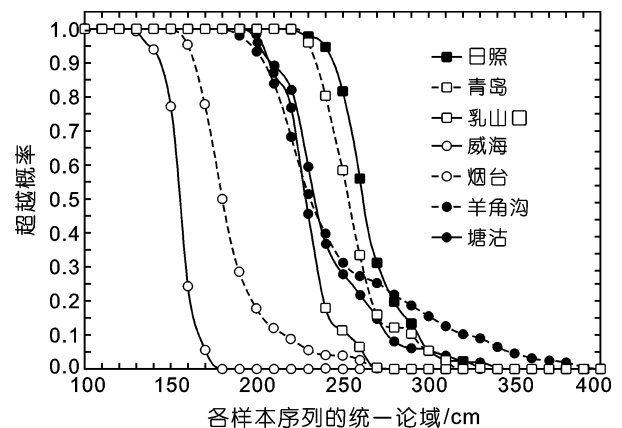


图 2 各海洋站年最高潮位发生风险的超越概率估计曲线

Fig.2 Exceeding-probability risk assessment of the highest annual water levels at each ocean station

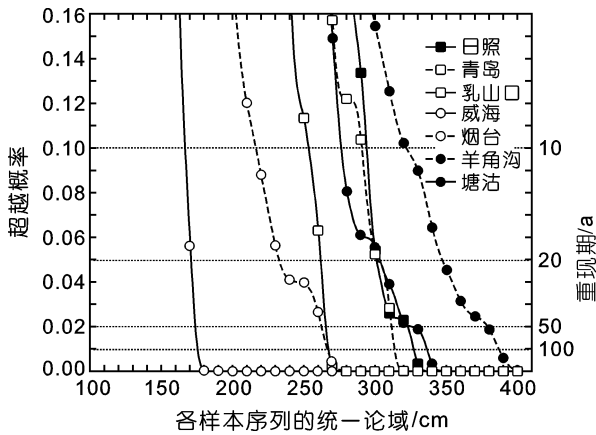


图 3 图 2 的局部放大图

Fig.3 Partial enlarged details of Fig.2

中算得, 由此计算了信息扩散结果对于 Gumbel 结果的相对误差(表 1)。容易看出, 两种方法的结果相当接近(但以信息扩散方法给出的概率估计值偏大者占多数), 二者之间相对误差控制在 3%以内, 绝对误差最大不超过 18 cm。又以青岛为例, 李培顺(1996)曾给出青岛港 10 年、20 年、50 年和 100 年一遇的潮位分别是 521 cm、532 cm、546 cm 和 557 cm, 本文信息扩散结果与其绝对误差不超过 10 cm。另外, 表 1 还给出了上述两种方法应用于资料样本个数仅为 25 的烟台站的结果, 二者之间绝对误差最大不超过 18 cm, 相对误差最大不超过 4%。可以推断, 文中信息扩散方法在兼顾小样本情形下所给各海洋站不同重现期的潮位推算结果是比较可靠的。

表 1 应用信息扩散与 Gumbel(耿贝尔)方法得到的各海洋站不同重现期潮位推算结果的比较(单位: cm)

Tab.1 Comparison of different return periods of the highest water levels at each ocean station between the information diffusion technique and Gumbel method(unit: cm)

重现期(a)	日照			青岛			烟台			羊角沟			塘沽		
	信息扩散	耿贝尔	相对误差 (%)	信息扩散	耿贝尔	相对误差 (%)	信息扩散	耿贝尔	相对误差 (%)	信息扩散	耿贝尔	相对误差 (%)	信息扩散	耿贝尔	相对误差 (%)
10	553	550	0.6	530	521	1.7	427	426	0.2	611	593	3.0	533	535	0.4
20	560	561	0.2	541	533	1.5	443	439	0.9	636	623	2.1	560	551	1.6
50	582	575	1.2	551	547	0.7	473	455	4.0	668	657	1.7	583	572	1.9
100	586	585	0.2	553	558	0.9	477	467	2.1	676	683	1.0	593	587	1.0

#### 4 结语

在收集渤、黄海沿岸主要海洋站(塘沽、羊角沟、烟台、威海、乳山口、青岛和日照)1950 年以来实测的年最高潮位资料的基础上, 考虑个别站位资料样本较小, 应用能够优化处理小样本和体现风险评估模糊不确定性的信息扩散方法, 通过对资料充足测站进行风险概率推算, 和与应用《港口工程技术规范》推荐(Gumbel 曲线)方法的结果进行比较验证, 显示了信息扩散方法所给结果对 Gumbel 方法结果的相对误差能控制在 3%以下。这一方面表明应用信息扩散方法所给推算结果是较为可靠的; 另一方面, 由于“信息扩散”是为弥补信息不足而对样本进行集值化的模糊数学处理方法, 即每一个观测样本值的信息均被“扩散”至整个论域空间, 因而, 在资料样本不足的情况下应用信息扩散方法进行不同重现期的极值推算应成为较好的解决方案。

总之, 本文应用能优化处理小样本和体现风险评估模糊不确定性的信息扩散原理和方法, 通过对

资料样本充足测站进行风险概率推算和不同方法之间的比较验证, 进而将该方法推广到资料样本不足地区的工程推算, 这对于潮位等水文实测资料短缺的新建或拟建工程区进行工程水文要素设计重现期的推算将有较好的帮助。

#### 参 考 文 献

史培军, 1996. 再论灾害研究的理论与实践. 自然灾害学报, 5(4): 6—17  
 张晓慧, 盛春雁, 邵滋和, 2006. 青岛沿海风暴潮分析. 海洋预报, 23, 增刊: 42—46  
 李培顺, 1996. 青岛地区的台风风暴潮灾度预报探讨. 海洋预报, 13(1): 53—57  
 李培顺, 1998. 青岛地区的台风风暴潮灾度预报研究. 海洋预报, 15(3): 72—78  
 杨华庭, 田素珍, 叶琳等, 1994. 中国海洋灾害四十年资料汇编(1949—1990). 北京: 海洋出版社, 1—6  
 黄崇福, 2001. 自然灾害风险分析. 北京: 北京师范大学出版社, 1—30  
 黄崇福, 刘新立, 周国贤等, 1998. 以历史灾情资料为依据的农业自然灾害风险评估方法. 自然灾害学报, 7(2): 1—9

## RISK ASSESSMENT OF STORM SURGE DISASTER IN THE COASTAL AREAS OF BOHAI SEA AND YELLOW SEA BASED ON INFORMATION DIFFUSION PRINCIPLE

QI Peng<sup>1,2</sup>, LI Ming-Jie<sup>1,2,3</sup>, HOU Yi-Jun<sup>1,2</sup>

(1. *Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, China;*

2. *Key Laboratory of Ocean Circulation and Waves(KLOCAW), Chinese Academy of Sciences, Qingdao, 266071, China;*

3. *Graduate School, Chinese Academy of Sciences, Beijing, 100049, China)*

**Abstract** The coastal areas of Bohai sea and Yellow sea are harmed by storm surges of both tropical cyclones in summer infrequently and temperate synoptic systems in spring, autumn and winter frequently. Based on the collected data of annual highest water levels in the areas since 1950, we focused on the calculation of highest water levels of multiyear return periods and risk assessments along the coast of Bohai and Yellow Sea. Due to insufficiency of the historic data we collected, a fuzzy risk assessment model was introduced based on the principle of information diffusion, which is able to optimize small samples. The results from information diffusion were compared and verified with the results from Gumbel curve method, which is recommended by code of harbor engineering. It is shown that the calculated results obtained by information diffusion are reliable and will be helpful to the calculation of highest water levels of design return periods for new engineering areas to be constructed where measured data are insufficient.

**Key words** Storm surge disaster, Fuzzy risk assessment, Information diffusion, Coastal areas of Bohai Sea and Yellow Sea