网络优先数字出版时间:2015-07-16
网络优先数字出版地址:http://www.cnki.net/kcms/detail/45.1206.G3.20150716.1435.002.html

Adaptability Research of Three-Dimensional Nested Logistic Model in Ocean Engineering*

三维双层嵌套 Logistic 模型在海洋工程中的适应性研究

CAO Xue-feng, GUO Pei-fang**, SHI Hong-yuan
曹雪峰,郭佩芳,石洪源

(Ocean University of China, Qingdao, Shandong, 266100, China)
(中国海洋大学,山东青岛 266100)

Abstract:【**Objective**】With the rapid development of massive and complicated ocean engineering, the theory and application of multivariate extreme value distribution model become more important than before.【**Methods**】This paper presents the process of the three-dimensional nested logistic model, trivariate extreme value distribution with marginals of generalized extreme value distribution, which is the unified form of three types of extreme value distribution and precludes human's disturbance from arbitrary choice of distribution type. The joint distribution of three factors, which are annual maximum water level and the corresponding wave height and wind speed, is analyzed using observed data from Shidao observation station, where these three random variables are mutually correlated. The joint probability and associated return period could be easily derived from the model.【**Results**】The model is suitable for representing the joint distribution of problems involved three correlated factors in ocean engineering.【**Conclusion**】This model is adaptable and flexible than bivariate models. It contributes significantly to assess the risk associated with different hydrological problems which involve three factors.

Key words: three-dimensional nested logistic model, joint probability, general extreme value distribution, return period

摘要:【**目的**】研究工程多因素复杂性的多元极值分布理论,探讨多元理论在实际工程的应用过程,以适应近代海洋工程规模日趋巨大的需要。【**方法**】以多元极值分布理论为基础,结合石岛湾海洋站实测的增水极值资料和同步波高、同步风速数据资料,详细阐述三维双层嵌套 Nested-Logistic 分布模型在工程实际中的应用过程,给出 3 种因素共同影响下联合重现期的推算方法。【**结果**】该模型加入了两两因素的相关性,对实际情况的刻画是科学合理的。【**结论**】模型在解决涉及 3 因素的水文工程问题方面,具有很好的灵活性和适定性,能够用于评估不同情况下的水文风险。

关键词:三维双层嵌套 Logistic 模型 联合概率 广义极值分布 重现期
中图分类号:P753 **文献标识码**:A **文章编号**:1005-9164(2015)03-0255-05

收稿日期:2015-04-02
作者简介:曹雪峰(1991-),男,博士研究生,主要从事物理海洋学研究。
*上海核工程研究设计院课题《滨海核电厂设计基准洪水概率分析研究》项目(编号:20130193)资助。
**通讯作者:郭佩芳(1955-),男,教授,主要从事物理海洋学和海洋管理学研究,E-mail:pfguo@ouc.edu.cn。

0 Introduction

【**Research significance**】In 1960, Gumbel[1,2] theoretically discussed the method of joint probability distribution about multivariate random variables which were mutually correlated. But the form of distribution function was implicit, the copular function was complex, and the calculation was large, which limited its development and application in engineering. In recent years, multivariate extreme value distribution (EVD) theory has been rapidly developed as a result of the claim of ocean engineering which has been more massive and more complicated than before. Therefore, it is necessary for human to propose one more efficient way to simplify large calculation and solve the engineering problems which involve three or more random variables. **【Achieved research progress】** Multivariate EVD theory has been successfully applied to many correlated hydrological extreme events by Coles S G, et al.[3,4], Yue S, et al.[5,6], Dong S, et al.[7,8] and Tao S, et al.[9]. Coles and Tawn[3] made use of three-dimensional nested logistic model to analyze joint probability of water level in three sites along the coast of southeast England, and demonstrated that the model was suitable. **【Current entry point】** Nonetheless, most research articles published in public discussed the engineering application of bivariate extreme value distribution model, while only a limited number of researchers have attempted to discuss three or more random variables. In consideration of Chinese region, there is much less. Shi D J, et al.[10] systematically discussed multivariate extreme value distribution theory, established the general form of three-dimensional nested logistic model for the first time, and finally estimated parameters in the method of moments (MM). That simplified large calculation and contributed significantly to the development of engineering application involved multivariables. **【Critical problem to be solved】** In this paper, 26-year consecutive data of annual extreme surge level and the corresponding wave height and wind speed from 1981 to 2004, which were observed from Shidao observation station in Shandong Province, were used to demonstrate the applicability of three-dimensional nested logistic model. Due to the limitation of statistics, the tidal level and rain-offs are ignored. Section 2 details the trivariate EVD model in a mathematical way, using marginals of generalized extreme value distribution (GEVD) which is the unified form of three types of EVD and precludes human's disturbance from arbitrary choice of distribution type. Section 3 presents a practical application, and the results of joint probability and associated return period are given at last. Section 4 concludes the usefulness of the model for representing the joint distribution of problems involved three correlated factors in ocean engineering.

1 The nested logistic model

The EVD could be divided into three types, which are Gumbel, Frechet and Weibull distribution. Theoretically, they could be unified as one form as follows:

$$F(x;\mu,\sigma,\xi)=\exp\left\{-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\},1+\xi\frac{(x-\mu)}{\sigma}>0, \quad (1)$$

where μ, σ, and ξ are the location, scale and shape parameters. Generally, shape parameter is the key to judge the EVD type of data. However, GEVD could be directly used as marginal distribution so that the disturbance of human's arbitrary choice of distribution type can be precluded.

The nested logistic model that involved three random variables was originally proposed by Mcfadden D. (1978) as follows:

$$F(x_1,x_2,x_3)=\exp\{-[(x_1^{-\frac{1}{\alpha\beta}}+x_2^{-\frac{1}{\alpha\beta}})^{\beta}+x_3^{-\frac{1}{\alpha}}]^{\alpha}\}, \quad (2)$$

where x_1, x_2 and x_3 are, respectively, the random variables of marginal Frechet distribution functions that could be expressed as follows:

$$F(x)=\exp(-x^{-1}), \quad (3)$$

where α and β $(0\leqslant\alpha,\beta\leqslant 1)$ are the parameters describing the association among the three random variables X_1, X_2 and X_3.

When the marginal distribution functions are standard Gumbel distribution, which can be expressed as follows:

$$F(x_1,x_2,x_3)=$$
$$\exp\{-[(e^{-\frac{x_1-\mu_1}{\alpha\beta\sigma_1}}+e^{-\frac{x_2-\mu_2}{\alpha\beta\sigma_2}})^{\beta}+e^{-\frac{x_3-\mu_3}{\alpha\sigma_3}}]^{\alpha}\}, \quad (4)$$

Shi, D J et al.[10] estimated association parameters in the method of moments and the estimator of $\hat{\alpha}$ and $\hat{\beta}$ are given by

$$\begin{cases} \hat{\alpha} = \dfrac{\sqrt{1-\gamma_{13}} + \sqrt{1-\gamma_{23}}}{2}, \\ \hat{\beta} = \dfrac{\sqrt{1-\gamma_{12}}}{\alpha}, \end{cases} \quad (5)$$

where $\gamma_{ij}(i < j, i, j = 1, 2, 3)$ is the product-moment correlation coefficient estimated by

$$\gamma_{ij} = \frac{E[(X_i - \mu_i)(X_j - \mu_j)]}{\sigma_i \sigma_j}, \quad (6)$$

where (μ_i, σ_i) and (μ_j, σ_j) are, respectively, the means and standard deviations of X_i and X_j. It is worth noting that γ_{12} is the maximum among them.

Where $\alpha = 1$, the trivariate distribution becomes trivariate logistic distribution. Where $\beta = 1$, the trivariate distribution becomes logistic model of mutlivariables. Where $\alpha = 1$ and $\beta = 1$, the trivariate distribution splits into the product of three marginal distributions, and becomes

$$F(x_1, x_2, x_3) = F(x_1)F(x_2)F(x_3). \quad (7)$$

This means the independent case and the product-moment correlation coefficient is equal to zero.

The joint return periods $T_{x_1, x_2, x_3}(x_1, x_2, x_3)$ exceeding certain values of variables X_1, X_2 and X_3 associated with the event ($x_1 > X_1$, $x_2 > X_2$ and $x_3 > X_3$, i. e., all of the three values are exceeded), can be represented by

$$T_{X_1 X_2 X_3}(x_1, x_2, x_3) = 1/[P(X_1 > x_1, X_2 > x_2, X_3 > x_3)] = 1/[1 - F_1(x_1) - F_2(x_2) - F_3(x_3) + F_{12}(x_1, x_2) + F_{13}(x_1, x_3) + F_{23}(x_2, x_3) - F(x_1, x_2, x_3)], \quad (8)$$

where $F_1(x_1)$, $F_2(x_2)$, $F_3(x_3)$, $F_{12}(x_1, x_2)$, $F_{13}(x_1, x_3)$ and $F_{23}(x_2, x_3)$ are respectively the values of marginal distribution functions.

The nested logistic model with GEVD marginals could be expressed as

$$F(x_1, x_2, x_3) = \exp\{-[((1+\xi_1 \frac{x_1 - \mu_1}{\sigma_1})^{-\frac{1}{\alpha\beta\xi_1}} + (1+\xi_2 \frac{x_2 - \mu_2}{\sigma_2})^{-\frac{1}{\alpha\beta\xi_2}})^{\beta} + (1+\xi_3 \frac{x_3 - \mu_3}{\sigma_3})^{-\frac{1}{\alpha\xi_3}}]^{\alpha}\}. \quad (9)$$

Before α and β are estimated, it is important that converting data series to obey the standard Gumbel distribution. The transformation formula could be represented as

$$Y_i = \frac{1}{\xi_i} \ln(1 + \xi_i \frac{X_i - \mu_i}{\sigma_i}) (i = 1, 2, 3), \quad (10)$$

where $Y_i (i = 1, 2, 3)$ are the transformed results.

2 Application

26-year consecutive data of annual extreme surge level and the corresponding wave height and wind speed from 1981 to 2004, which were observed from Shidao observation station in Shandong Province, were made use of to demonstrate the applicability of three-dimensional nested logistic model with GEVD marginals. Due to the limitation of statistics, the tidal level and rain-offs are ignored.

2.1 Parameter estimation

This paper selects GEVD as the single-variable analysis model. In practice, the location, scale and shape parameters of the GEVD can be estimated by maximum likelihood method (ML). In Section 2, the association parameters estimation using MM method has been detailed enough. The efficiency of this methodology, in which marginal and association parameters are estimated separately, has been tested by Shi D J, et al.[11]. For the purpose of comparison, the correlation coefficients between surge level and wave height, wave height and wind speed, and surge level and wind speed, are 0.096, 0.807 and −0.049, respectively. Thus, the three random variables X_1, X_2 and X_3 are wave height, wind speed and surge level, respectively. The estimated parameters of series are listed in Table 1.

Table 1 Estimated parameters of the model

Marginal variable	Location parameter	Scale parameter	Shape parameter	Association parameter
X_1: Wave height	1.5860	0.7277	0.0452	$\alpha = 0.9875$
X_2: Wind speed	6.9319	3.3520	0.1998	$\beta = 0.4448$
X_3: Surge level	54.1183	8.5568	−0.0131	

2.2 Validity of the model

The empirical non-exceedance probability can be calculated by the formula:

$$P_k = \frac{k}{N+1} \quad \text{or} \quad P_{ijk} = P(X \leqslant x_i, Y \leqslant y_j, Z \leqslant z_k) = \frac{\sum_{r=1}^{i}\sum_{s=1}^{j}\sum_{t=1}^{k} n_{rst}}{N+1}, \quad (11)$$

where N is the total number of observations ($N = 26$), and P_k or P_{ijk} is the cumulative frequency, the probability that a given value is less than the k th or n_{rst} th smallest observation in the total N observations. Fig. 1 (a), (b) and (c) show the fit of the marginal GEVDs of wave height, wind speed and surge level series, respectively. The empirical and theoretical joint probabilities are illustrated in Fig. 1 (d).

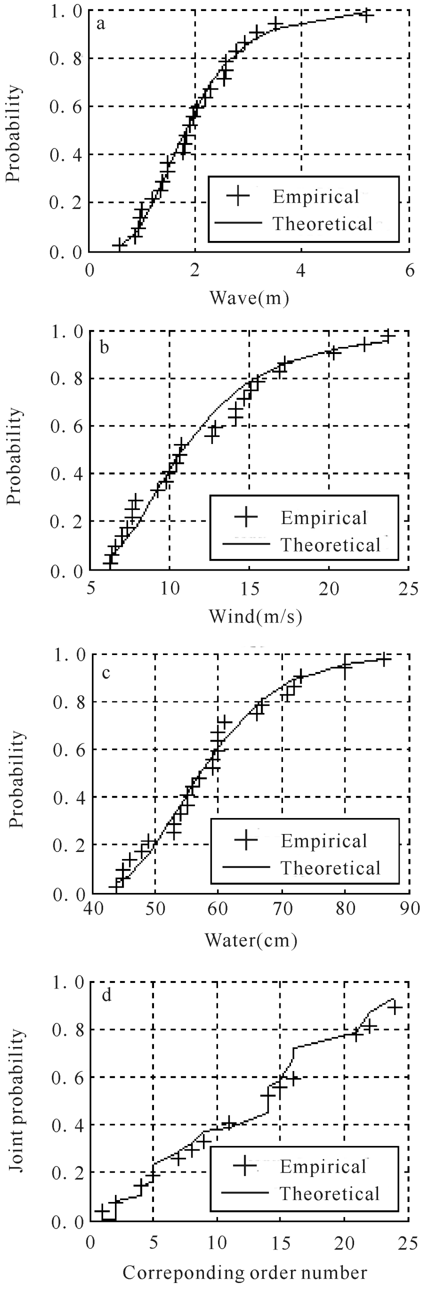

Fig. 1 Distribution of wave height series(a), wind speed series (b), surge level series(c) and joint distribution of the nested logistic model(d)

The Kolmogorov-Smirnov test is executed to test the goodness of fit of the marginal distributions and the GEV nested logistic model. The tested results show that all the distributions are accepted at the significant level 0.05. Thus, it is concluded that the model is suitable.

2.3 Joint return periods

The joint return period $T_{x_1,x_2,x_3}(x_1, x_2, x_3)$ can be calculated by Eq(8). For example, Fig. 2 illustrated the joint return periods given the values of surge level and wave height. Certainly, there are various occurrence scenarios which could be analyzed and contribute to assess the risk associated with different hydrological problems.

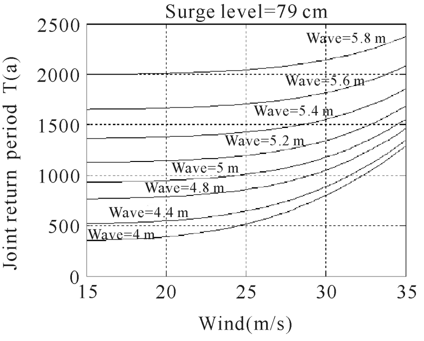

Fig. 2 Joint return period of combinations of surge level, wave height and wind speed (given surge level and wave height)

3 Conclusions

This paper proposed the detailed application process of three-dimensional nested logistic model, using 26-year consecutive data of annual extreme surge level and the corresponding wave height and wind speed from 1981 to 2004, which were observed from Shidao observation station in Shandong province. GEVD adopted as marginal distribution function is the unified form of three EVD types, which precludes human's disturbance from arbitrary choice of distribution type. The correlation coefficients among surge level, wave height and wind speed are added into the three-dimensional nested logistic model, which is more adaptable and flexible than bivariate models. In addition, the joint probability and corresponding joint return period could be obtained conveniently. Various occurrence scenarios could be analyzed and contribute to assess the risk associated with different hydrological problems. It concludes that the three-dimensional nested logistic model is suitable for representing the joint distribution of problems involved three correlated factors in ocean engineering.

Acknowledgements

This work was financially supported by the Re-

searches of Anti-flooding in Nuclear Power Plant Siting in Coastal Area. (GHME20130193).

References:

[1] Gumbel E J. Multivariate extremal distributions[J]. Bull Int Stat Inst,1960,39(2):471-475.

[2] Gumbel E J. Distributions Del Valeurs Extremes En Plusieurs Dimensions[M]. 9th Edition. Paris: Publications de L'Institute de Statistique,1960:171-173.

[3] Coles S G,Tawn J A. Statistical methods for multivariate extremes: An application to structural design[J]. Apple Statist,1994,43(1):1-48.

[4] Stuart S,Coles G. An Introduction to Statistical Modeling of Extreme Values[M]. 10th Edition. London: Springer London Ltd,2013.

[5] Yue S,Ouarda T B M J,Bobee B,et al. The Gumbel mixed model for flood frequency analysis[J]. Journal of Hydrology,1999,226(1):88-100.

[6] Yue S. The Gumbel logistic model for representing a multivariate storm event[J]. Advances in Water Resources,2001,24(1):179-185.

[7] Dong S,Hao X L. Statistical Analysis of Ocean Environmental Conditions with PTGEVD[C]. Proceedings of OMAE2004,Vancouver,Canada,2004:617-621.

[8] Dong S,Liu W,Zhang L Z. Long-term Statistical Analysis of Typhoon Wave Heights with Possion-maximum Entropy Distribution[C]. Proceedings of the ASME 2009 28th international conference on ocean, offshore and arctic engineering, OMAE2009-79278, 31 May-5 June,Hawaii,2009.

[9] Tao S S,Dong S,Wang N N,et al. Estimating storm surge intensity with Poisson bivariate maximum entropy distributions based on copulas[J]. Nat Hazards,2013,68(1):791-807.

[10] Shi D J,Sun B K. Moment estimation in a nested logistic model[J]. Systems Engineering-Theory & Practice,2001,1(1):53-60.

[11] Shi D J,Feng Y Q. Parameteric estimation by maximum likelihood and stepwise method for multivariate extreme value distribution[J]. Journal of Systems Science and Mathmatcial Sciences,1997,17(3):244-251.

(责任编辑:陆　雁)