

# Calculation of ship length probability distribution with two parameters

Ailan Shi, Shesheng Zhang

**Abstract**— The macro planning of waterway transportation requires the distribution of the ship's characteristic length. Using two times log transformation and parameter transformation, the expression of the Weibull distribution of ship length was converted into linear form. The theoretical distribution of ship length was replaced by its cumulative distribution, the parameters of Weibull distribution was calculated by regression analysis. And then the probability distribution of 9 types of ship length was obtained.

**Index Terms**—Weibull distribution, linear regression, Green function.

## I. INTRODUCTION

The ocean plays a very important position in the world. It is very important to develop and utilize the ocean reasonably. The main tools for the development and utilization of the ocean are all kinds of ships, and the ship will be influenced by the waves in the ocean. Therefore, it is very important to study the hydrodynamic force and the response of ship to wave. In the linear wave theory, the free surface Green function is a mathematical tool for solving wave problems. As to the study of the free surface Green function, one aspect is Green function and its partial derivative, and the other is the numerical calculation. For the latter, it is necessary to simplify the numerical calculation method, as little as possible to take up computer memory and meet certain accuracy, so that it can be applied in engineering. Zhan Chengsheng [1] proposed a new calculation method, that is, calculating Green function by partition domain. Han Ling[2] has studied the method of calculating the Green function polynomial of finite depth in the time domain. Teng Bin[3] solved Green function by using the higher order boundary element method in the integral equation. On the basis of previous studies, Shen Liang[4] proposed a new method about the node and node interpolation to make the calculation of Green function faster. Shen Jing[5] focused on viscous dissipation and proposed a new method. Liu Changfeng[6] started from the convolution theorem, the relationship between the time domain and the frequency domain of the Green function was obtained according to the nature of the Fourier transform, finally derived A new recursive formula for calculating convolution. Wang Xing[7] used the Fourier transform and inverse transform to deduce the Green function and its derivative formula, and analyzed its properties, and gave some special points on the value of Green's function. The author constructed a complex fast numerical method, and finally gave the calculation system of Green's function. In Wang Xing's[7] article, there is no

calculation of the parameters. In the formula, there are two parameters, including the length of ship and the number of wave. If the parameters range are not determined, the value of Green's function can be infinite, losing the meaning of accelerated computing.

From the above discussion, we can know that after using Fourier transform and inverse transform to get the analytic formula of Green function, the problem of parameters' range need to be solved. The parameters range of ship length can be obtained from the joint distribution of the length and the number of wave number after solving the distribution of ship length, and the confidence interval of the parameters can be given under a certain degree of confidence. However, there is no literature on the distribution of ship length, the author will combine the ship length data to discuss the ship length distribution..

## II. REGRESSION MODEL

Suppose the ship length  $X$  obey Weibull distribution, then its distribution density function is:

$$f(x) = \begin{cases} \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (1)$$

here  $a, b$  are the two parameters of the Weibull distribution,  $a > 0$  is the ratio parameter, is the shape parameters. If the probability density  $f(x_i)$  of point  $x_i$  is given,  $i=1,2,\dots,N$ . According to the above formula can get:

$$f(x_i) = \frac{b}{a} \left(\frac{x_i}{a}\right)^{b-1} e^{-\left(\frac{x_i}{a}\right)^b} \quad i = 1, 2, \dots, N$$

In the above equations, there are  $a, b$  two unknown quantities, they are nonlinear about  $a, b$ . In theory, their values can be obtained by using the nonlinear optimization method. However, in practice, the computation of two dimensional nonlinear optimization is very large and there is calculation error. Linear formula which is easy to calculate is constructed in the following.

The distribution function of the Weibull distribution from negative infinity to  $x$  integral:

$$F(x) = 1 - \exp\left(-\left(x/a\right)^b\right) \quad (2)$$

The nonlinear (2) is converted to linear form:

Ailan Shi, Department of statistics, Wuhan University of Technology, Wuhan, 430063, China.

Shesheng Zhang, Department of statistics, Wuhan University of Technology, Wuhan, 430063, China.

$$\ln[-\ln(1-F)] = -b \ln a + b \ln x \tag{3}$$

Linear function can be get by variable substitution :

$$y = c + bz \tag{4}$$

in the equation:

$$y = \ln[-\ln(1-F)], c = -b \ln a, z = \ln x$$

Using the formula and ship length data, a linear regression was performed, and the regression coefficient, was obtained, and then the parameter was calculated by using the formula, then ship length probability distribution was obtained.

### III. SHIP LENGTH DISTRIBUTION

#### 3.1 data sources and processing

The ship data mainly from the international shipping network, secondly from CNKI for supplement. There are 9 kinds of ship, respectively: liquid cargo ships, dry cargo ships, engineering ships, passenger ships, special ships, marine engineering ship, official ships, fishing boats and working boats, a total of 2040 data.

Data processing is to delete some outliers, after the logarithm transform, there are points (outliers) which will lead to the regression result not ideal, so need to delete them.

#### 3.2 overall ship length distribution

After the preliminary analysis of the data, the frequency distribution histogram of the overall ship length can be drawn, as shown in Fig. 1. L means overall ship length.

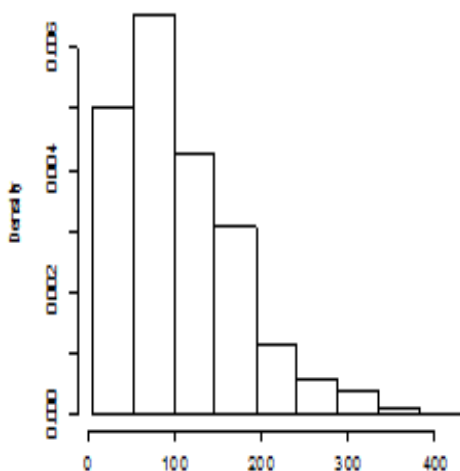


Fig. 1 frequency distribution histogram

It can be seen that the figure is similar to the distribution pattern of Weibull. Use the above model to fit, the parameters of the regression equation and related data can be got by least square method, see Fig.2 Form figure, we easy obtain he linear regression results is

$$y = -8.297153 + 1.740584 z \tag{6}$$

here P value is 0, Adjusted R square value is 0.9849, The results show that the regression equation is significant and the result is satisfactory. the ship length is approximately Weibull (117.839, 1.771) distribution.

Overall ship length distribution function:

$$F(x) = 1 - \exp[-(x/117.8216)^{1.708983}] \quad x \geq 0 \tag{5}$$

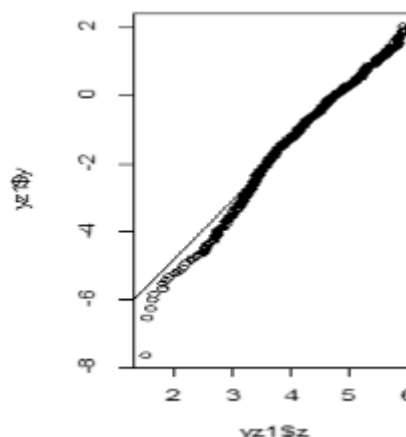


Fig. 2 ship length regression

#### 3.3 Classified ship length distribution

Similar to the overall ship length, the frequency distribution histogram of the 9 categories of ship lengths, as shown in Figure 3, L means classified ship length.

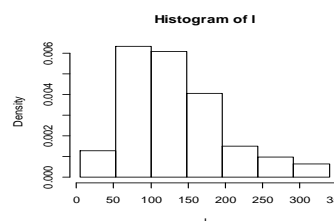


Fig. 3 Liquid cargo ship length frequency histogram

It can be seen from the figure, the ship length distribution of different ship is similar to overall ship, but different ship has different ship length distribution, which can be reflected by frequency distribution histograms. That is to say, different kind of ship has different probability on different ship length value, and intervals which correspond to maximum frequency are also different, but are still obey the Weibull distribution in general.

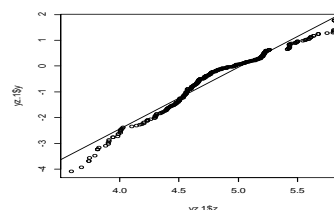


Fig. 4 Effect of Liquid cargo ship master regression

From the listed regression results, all adjusted R square value are very high, in addition to the special ship, other ships' adjusted R square value are above 0.9, the minimum is 0.9392, and special ship' s adjusted R square value is lower compare to other ships, but has also reached 0.8808, which is highly significant, indicating that transformation of the independent variables and dependent variables exist significant linear relationship, also illustrating the classified ship lengths were also obey Weibull distribution, but parameters between different categories vary. In general, the P values of all types of regression coefficients were 0, indicating that all the regression coefficients were significant, and that all the regression equations were significant. The

regression effect diagrams are shown in figure 4.

The distribution functions of the 9 types of ship length are:

(Liquid cargo ship)

$$F(x) = 1 - \exp[-(x/151.6835)^{2.3945}] \quad x \geq 0$$

(Dry cargo ship)

$$F(x) = 1 - \exp[-(x/161.7972)^{2.5805}] \quad x \geq 0$$

(Engineering ship)

$$F(x) = 1 - \exp[-(x/88.3503)^{2.1136}] \quad x \geq 0$$

(Passenger ship)

$$F(x) = 1 - \exp[-(x/64.7674)^{1.7853}] \quad x \geq 0$$

(Special ship)

$$F(x) = 1 - \exp[-(x/74.2747)^{1.2922}] \quad x \geq 0$$

(Marine engineering ship)

$$F(x) = 1 - \exp[-(x/62.2370)^{0.6982}] \quad x \geq 0$$

(Official ship)

$$F(x) = 1 - \exp[-(x/45.2288)^{1.6265}] \quad x \geq 0$$

(Fishing boat)

$$F(x) = 1 - \exp[-(x/47.1124)^{3.2001}] \quad x \geq 0$$

(Working boat)

$$F(x) = 1 - \exp[-(x/56.0162)^{2.7253}] \quad x \geq 0$$

#### IV. CONCLUSION

The probability distribution function of Weibull probability density function is obtained by integrating the Weibull probability density function and is converted into a linear form with appropriate calculation and parameter changes. Using the cumulative distribution function of the ship length data replace the theoretical distribution function, the corresponding Weibull distribution parameters a,b are obtained by linear regression. The method used by the author is not only suitable for overall ship length, but also for the classified ship length. Weibull distribution fitting for each type of ship length can achieve the ideal effect. By solving the joint distribution of the wave number and its distribution, the confidence interval of the parameters under different confidence interval can be obtained to solve the problem of Green function' s numerical calculation..

#### ACKNOWLEDGMENT

The paper is financially supported by China national natural science foundation (No.51139005).

#### REFERENCES

- [1] Zhan Chengsheng, Zou Zaojian, Zheng Weitao. Numerical analysis of far field of ship motion Green function [J]. Wuhan Science Journal of University of Technology (Transportation Science and Engineering Edition), 2002 (04): 425~428.
- [2] Han Ling, Teng Bin, Gou Ying. A method for calculating the polynomial expansion of Green function in time domain [J]. Water dynamics research and development of computing (Series A). 2004.(09).
- [3] Teng Bin, Han Ling. Application of the time domain Green function method to simulate the effect of wave on the structure in finite depth[J]. water dynamics research and development (Series A).2006 (02).

- [4] Shen Liang, Zhu Renchuan,, Liu Yingzhong, Miu Guoping, Green,, Practical numerical calculation of Green function in deep water [J]. water dynamics research and development (Series A).2007 (03).
- [5] Shen Jing. Study on the Green function in the frequency domain of the free surface with viscous dissipation [D]. Harbin Engineering University.2010.
- [6] Liu Changfeng, Gou Ying, Teng Bin. A new algorithm for time domain Green function convolution[J]. Journal of hydrodynamics (Series A).2010 (04).
- [7] Wang Xing. Research on the theory and software of Green function in the near field [D]. Wuhan University of Technology.2013.