
 論 文

大韓造船學會論文集
 第29卷 第2號 1992年 5月
 Transactions of the Society of
 Naval Architects of Korea
 Vol. 29, No. 2, May 1992

극한 파고 계산에 있어서 Type III 분포의 응용

이태일*, 권순홍**, 전영기***

Applications of the Type III Asymptotic Distribution for Extreme Sea Level Computations

by

T. I. Lee*, S. H. Kwon** and Y. K. Chon***

요 약

본 연구를 통하여 극한 파고를 계산하는 방법들을 제시하였다. Type III 분포에 근거해서 분포 함수의 파라미터 산출을 위하여 non-linear multiple regression 방법, skewness 방법, maximum likelihood 방법들을 사용하였다. 좀 더 정확한 결과를 얻기 위하여 추정된 분포 함수의 차이를 다항식을 도입하여 맞추었다. 제시한 방법을 응용하여 계산 예들을 보였다.

Abstract

The computational methods of extreme sea level are developed in this study. Based on type III asymptotic distribution, non-linear multiple regression method, skewness method and maximum likelihood method are used to evaluate the parameters of the distribution. The difference between real data and evaluated distribution function is fitted to get more desirable accuracy by employing polynomials. The numerical examples are given in the last section in order to illustrate the application of the present scheme.

1. 서 언

공학적인 설계에 있어서 확률적인 문제의 중요한 부분 중의 하나는 랜덤 변수의 최대치 또는 최소치라 할 수 있는 극한치를 통계적으로 추정하는 문제 일 것이다. 따라서 본 논문에서는 선박, 해안구조물 또는 해양구조

물의 설계에 필요한 극한 파고를 구하는 데 있어서, 관측되어진 파고의 자료를 Type III 분포에 적용시켜 50~100년 정도의 재현 기간 동안 조우할 수 있는 극한 파고를 예측하고자 한다.

Type III 분포를 통하여 이러한 예측을 하기 위해서는 이 분포 함수의 모수들의 결정이 필요한데 본 논문에서는

접수일자 : 1991년 8월 28일, 재접수일자 : 1991년 11월 25일

* 정회원, 현대선박해양연구소

** 정회원, 부산대학교 조선공학과

*** 정회원, 한국선급

大韓造船學會論文集 第29卷 2號 1992年 5月

Table 1 Data A

Significant wave height(M)	관측수
0.2-0.4	11
0.4-0.6	151
0.6-0.8	158
0.8-1.0	175
1.0-1.2	109
1.2-1.4	116
1.4-1.6	91
1.6-1.8	63
1.8-2.0	47
2.0-2.2	40
2.2-2.4	27
2.4-2.6	19
2.6-2.8	21
2.8-3.0	12
3.0-3.2	6
3.2-3.4	7
3.4-3.6	6
3.6-3.8	2
Total :	1,061

Table 2 Data B

Significant wave height(M)	관측수
0.0-0.5	144
0.5-1.0	63
1.0-1.5	48
1.5-2.0	27
2.0-2.5	15
2.5-3.0	10
3.0-3.5	5
3.5-4.0	3
4.0-4.5	1
Total :	316

모수의 결정 방법으로 Non-linear Multiple Regression Method, Skewness Method 그리고 Maximum Likelihood Method를 이용하여 모수들을 구하였다. 그러나 원래의 Type III 분포함수만으로는 만족할 만한 누적 분포 함수(Cumulative Distribution Function)를 구하기 어렵기 때문에 변형된 형태의 함수를 도입하였다.

여기서 사용된 파고의 자료는 Table 1, Table 2에 나타난 바와 같이 Type III 분포의 보편적인 적용성을 검

토하기 위하여 서로 상이한 확률 밀도를 가지는 자료들이다. Table 1은 미국 동부 해안에서 42개월간 관측된 자료이며 Table 2는 제주도 앞 해상에서 매년 2월달마다 11년간 관측된 자료이다. 이 두 자료를 통하여 각각의 모수 결정 방법에 의한 결과들을 비교해 보았다.

2. 극한치 분포

랜덤 변수의 극한치 예측에는 그 확률 밀도 함수(Probability Density Function)의 꼬리 부분과 관계하여 특정한 분포를 사용할 수 있다. 특히 극한치 예측에 많이 쓰이는 분포는 Type I 분포, Type II 분포 그리고 Type III 분포가 있다[4].

여기서 확률 밀도의 극한치 방향으로 꼬리 부분이 지수적으로 감소하는 경우 극한치는 Type I 분포를 따르고, 그 꼬리 부분의 감소가 다항식으로 표현될 때는 극한치는 Type II 분포를 따르며, 극대치나 극소치가 제한되어 있는 경우에는 Type III 분포를 사용하는 것으로 알려져 있다[2].

극한치의 분포가 위의 어떤 분포에도 속하지 않는 경우가 있으나 선박이나 해양구조물의 설계에 필요한 극한파고를 구하는 문제에 있어서 본 논문에서는 파고 자료의 해석에 가장 유용하게 사용되어 지고 있는 Type III 분포를 적용하고자 한다. 물론 문제의 해석에 필요하다고 생각되는 분포를 선택하는데 있어서 이론적 근거의 필연성은 없다. 그러나 극한 파고에 대한 장기 예측에 있어서 Type III 분포에 의한 모델링은 극한치 해석에 있어서 쉽게 응용되어 질 수 있는 잇점이 있고 또 많이 사용되어져 왔으나, 짧은 기간 동안의 자료로써 구조물 전체 수명에 대해서 예측하는 일은 관측되어진 자료의 통계적인 신뢰성과 파의 장기적인 변화성등을 통하여 신중하게 고려되어져야 할 것이다.

3. 모수 결정

Weibull 분포라고도 하는 Type III 분포의 누적 분포 함수는 다음과 같이 나타낼 수 있다.

$$F(y) = \exp\left[-\left(\frac{w-y}{w-v}\right)^k\right] \quad (1)$$

$$y \leq w, k > 0, v < \infty$$

또 확률 밀도 함수는 다음과 같다.

$$f(y) = \frac{k}{w-v} \left(\frac{w-y}{w-v}\right)^{k-1} \exp\left[-\left(\frac{w-y}{w-v}\right)^k\right] \quad (2)$$

여기서 y 는 랜덤 변수로서 유의 파고이다. (1)식에서 볼 수 있듯이 이 분포는 모수 w 에 의하여 제한되어 있기 때문에 여태까지 공학적인 문제에 있어서 많이 사용되어진 Type I 분포보다 극한치를 계산하는데 있어서 과대 예측할 가능성이 적어지게 된다.

장기 파고 예측에 있어서 실제의 자료로써 (1), (2)식의 모수를 구하는데는 통상 Non-Linear Regression Method, Skewness Method 그리고 Maximum Likelihood Method가 있다. 다음에 이들 방법을 이용하여 모수들을 구하고자 한다.

3.1 Non-Linear Regression Method

주어진 파고의 자료에 의하여 모수 k, w, v 를 구하는데 있어서 Non-Linear Regression Method를 사용할 수 있는데 이는 컴퓨터를 이용한 수치 해석적인 방법으로 가능하다. 계산을 보다 용이하게 하기 위하여 (1)식의 양변에 두번 Log를 취하면

$$\ln[-\ln F(y)] = k \times \ln\left(\frac{w-y}{w-v}\right) \quad (3)$$

여기서 좌변은 주어진 파고의 자료로부터 구할 수 있고 우변은 모수 k 에 대해서 선형이다. (3)식을 통하여 원래의 파고 자료로부터 모수 k, w, v 를 구하는데는 여러가지 수치해석적인 방법이 있는데 본 논문에서는 Grid Search Method, Gradient Search Method, Parabolic Extrapolation Method 그리고 Gradient Expansion Method, 등 4가지의 방법을 사용하여 모수들을 구하였다[3]. 자료 A와 자료 B에 대하여 Grid Search Method로 fitting한 누적 분포함수를 Fig. 1과 Fig. 2에 나타내었다. 누적 분포 함수에 있어서 자료에 의한 것과 비교해 볼 때 어느 정도는 잘 맞지만 이 모수들에 의한 Fig. 3과 Fig. 4의 확률 밀도 함수를 볼 때는 만족할 만한 결과라고 보기는 어렵다. 다른 방법에 의한 fitting 결과도 이와 비슷하였다. 따라서 (1)식으로 fitting하는데 있어서 $\ln[-\ln F(y)]$ 와 $k \times \ln[(w-y)/(w-v)]$ 의 차이를 최대한 감소시키기 위하여 $\ln[-\ln F(y)]$ 와 $k \times \ln[(w-y)/(w-v)]$ 와의 차이를 구하고 이를 다시 y 에 대한 3차 다항식으로 fitting하여 (3)식의 우변에 더하였다[1]. 즉,

$$\ln[-\ln F(y)] = k \times \ln\left(\frac{w-y}{w-v}\right) + \Delta y \quad (4)$$

$$\Delta y = a + by + cy^2 + dy^3 \quad (5)$$

자료 A와 자료 B에 의한 Δy 와 (5)식에 의하여 fitting된

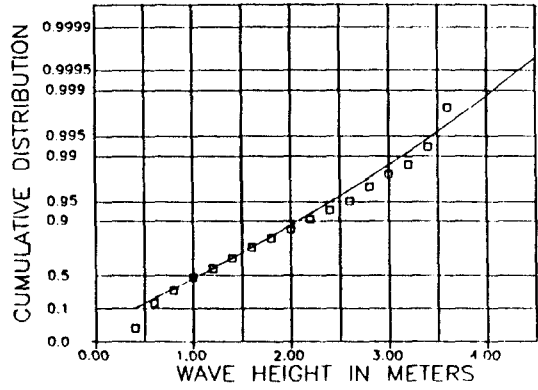


Fig. 1 CDF of data A

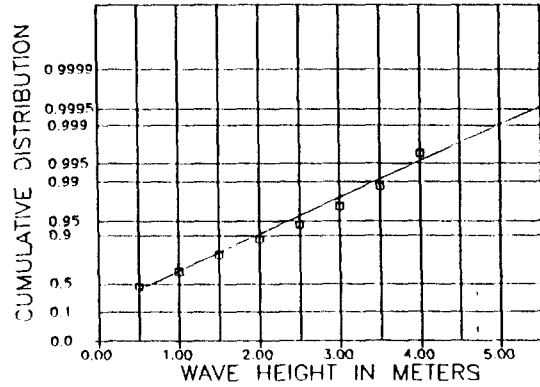


Fig. 2 CDF of data B

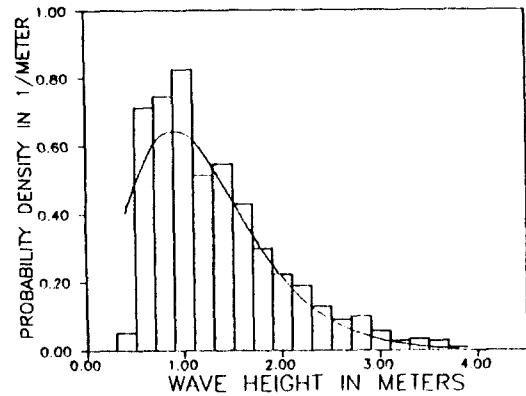


Fig. 3 PDF of data A

Δy 를 Fig. 5와 Fig. 6에 나타내었다. (5)식에 의한 Δy 를 (4)식에 대입하여 누적 분포 함수를 다음과 같이 나타내었다.

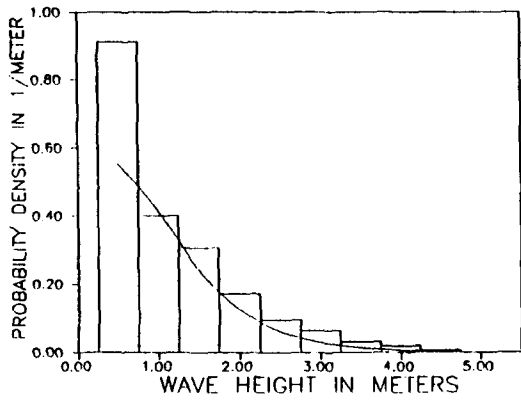


Fig. 4 PDF of data B

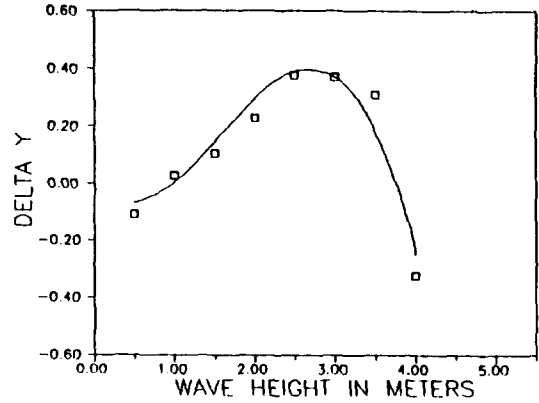


Fig. 6 Δy of data B

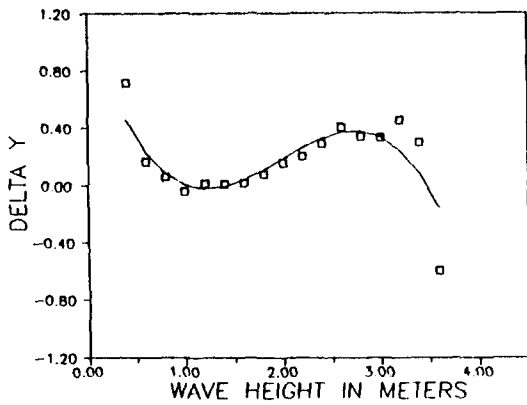


Fig. 5 Δy of data A

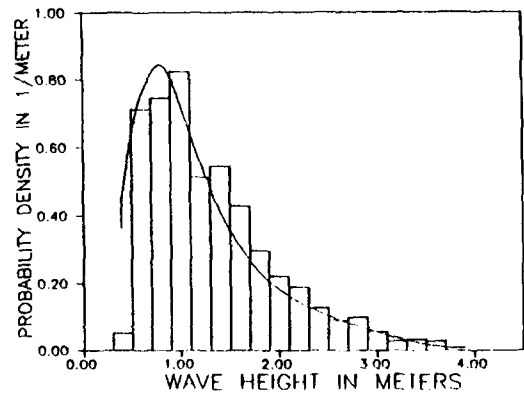


Fig. 7 PDF of data A

$$F(y) = \exp\left\{-\left(\frac{w-y}{w-v}\right)^k e^{\Delta y}\right\} \quad (6)$$

(6)식은 누적 분포 함수의 조건을 만족시키므로 Δy 는 원래의 Weibull 분포의 특성을 변화시키지 않는다. 따라서 (6)식을 보다 개선된 누적 분포함수로 취하였다. (6)식에 의한 확률 밀도 함수는 Fig. 7와 Fig. 8에 나타나 있다. 이를 앞의 결과와 비교하면 훨씬 나아졌음을 알 수가 있다. 마찬가지로 Gradient Expansion Method, Gradient Search Method 그리고 Parabolic Extrapolation Method에 대해서도 계산을 행하였다. 본 연구 목적이 누적 분포 함수에 의한 극한 파괴 예측에 있으므로 각 방법들을 비교하기 위하여 관측되어진 자료의 누적 분포와 각 방법에 의해 구해진 누적 분포 함수와의 차이를 다음과 같이 구하였다.

$$DIF = \sum_i \left\{ \ln[-\ln F(y_i)] - \left[k \times \ln\left(\frac{w-y_i}{w-v}\right) + \Delta y_i \right]^2 \right\}$$

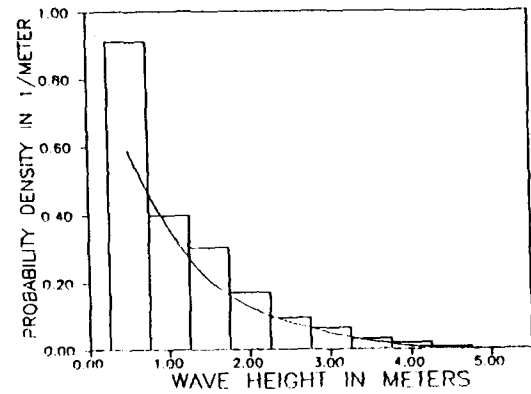


Fig. 8 PDF of data B

Table 3, Table 4에는 이들 계산 결과에 의한 모수 k, w, v, a, b, c, d 와 DIF 값을 나타내었다.

Table 3 Values of parameters and DIF from data A

parameter method	k	w	v	a	b	c	d	$\Sigma(\)^2$
Grid Search	16.23	10.13	0.869	1.168	-2.289	1.357	-0.229	0.435
Gradient Search	1100.	300.	1.07	-0.96	-0.837	1.631	-0.296	0.305
Gradient Expansion	126.93	11.790	0.996	-8.205	6.086	2.184	-0.278	0.283
Parabolic Extrapolation	90.	10.	1.	-6.499	4.362	2.141	-0.263	0.282
Skewness	679.2	351.2	1.034	0.759	-2.058	1.24	-0.214	0.034
Maximum Likelihood	5.469	5.022	1.084	1.23	-2.769	1.30	-0.177	0.670

Table 4 Values of parameters and DIF from data B

parameter method	k	w	v	a	b	c	d	$\Sigma(\)^2$
Grid Search	1817.1	1202.6	0.413	-0.009	-0.290	0.385	-0.082	0.099
Gradient Search	2000.	1600.	0.241	0.324	-0.608	0.394	-0.079	0.05
Gradient Expansion	259.2	26.	0.97	-9.27	8.269	0.5	-0.063	0.06
Parabolic Extrapolation	1000.	1600.	1	-0.02	-1.22	0.413	-0.086	0.096
Skewness	2689.3	1720.6	0.778	-0.459	-0.554	0.528	-0.099	0.074
Maximum Likelihood	55.32	9.7	1.002	-5.449	4.008	0.576	-0.033	0.098

3.2 Skewness Method

이는 Type III 분포의 모수 k,w,v를 구하는데 있어서 자료의 Skewness Coefficient를 이용하는 방법이다. 특히 Skewness Coefficient는 모수 k만의 함수로 나타낼 수 있다. 자료에 의한 Skewness Coefficient는

$$\theta = \frac{E[(y-\mu)^3]}{\sigma^3} \quad (7)$$

σ : 자료에 의한 표준 편차
 μ : 자료의 평균치

(7)식은 다시 k만의 함수로 나타낼 수 있는데

$$\theta = -\{\Gamma(1+3/k) - 3\Gamma(1+2/k)\Gamma(1+1/k) + 2\Gamma^3(1+1/k)\} \times \{1/[\Gamma(1+2/k) - \Gamma^2(1+1/k)]\}^{1/2} \quad (8)$$

(8)식에서 k를 구하고

$$\mu = w - (w-v)\Gamma(1+1/k) \quad (9)$$

$$\sigma = (w-v)^{1/2} \{\Gamma(1+2/k) - \Gamma^2(1+1/k)\} \quad (10)$$

(9), (10)식의 연립 방정식으로부터 w,v를 구할 수 있다.

Fig. 9에는 각 Skewness Coefficient에 대하여 Secant Method를 통하여 구한 k값을 나타내었다. 그림에서 볼 수

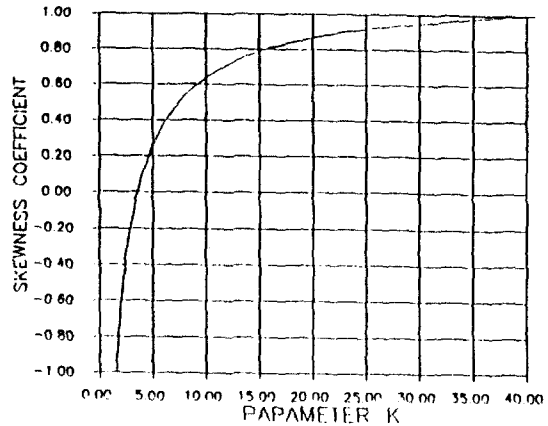


Fig. 9 Skewness coefficient and parameter K

있듯이 Skewness Coefficient가 어떤 일정치 이상 또는 이하가 될 경우 정확한 k값을 구하기가 어렵거나 불가능하게 된다. 또 k값을 구하더라도 (9), (10)식에 의하여 구해진 w,v값이 원래의 조건에 부적합하게 되는 경우도 있다. 그러나 일반적인 경우에 있어서 이 방법을 사용하는 것은 아무런 문제가 없으며 모수들의 값을 간편하게 구할 수 있다. 여기서도 Ochi[1]의 Δy 를 도입하여 모수들을 구하고 그 결과를 Table 3과 Table 4에 나타내었다.

3.3 Maximum Likelihood Method

Type III 분포의 likelihood function은 다음 식에서 구할 수 있다.

$$L(y_i | k, w, v) = \prod_{i=1}^n f(y_i) = k^n (w-v)^{-nk} \times \prod_{i=1}^n \{ (w-y_i)^{k-1} \exp\{-\sum_{i=1}^n \frac{(w-y_i)^k}{w-v}\} \} \quad (11)$$

n: 유의 파괴 구간 수

(11)식은 일의적인 증감 함수이므로 모수 k,w,v에 대한 $\ln[L(y_i | k, w, v)]$ 의 편미분한 결과들을 0으로 놓아서 모수들을 구할 수가 있다. 즉,

$$\frac{\partial}{\partial w} \ln L = -\frac{nk}{w-v} + \sum_{i=1}^n \left[k \left(\frac{w-y_i}{w-v} \right)^k \frac{v-y_i}{(w-y_i)(w-v)} + \frac{k-1}{w-y_i} \right] = 0 \quad (12)$$

$$\frac{\partial}{\partial k} \ln L = \frac{n}{k} - n \ln(w-v)$$

$$-\sum_{i=1}^n \left[\left(\frac{w-y_i}{w-v} \right)^k \ln \left(\frac{w-y_i}{w-v} \right)^k - \ln \left(\frac{w-y_i}{w-v} \right) \right] = 0$$

(12)식의 각 모수들을 수치 해석적 방법인 Newton-Raphson 방법을 사용하여 구하였다. Δy 에 의한 개선된 누적 분포 함수의 모수들과 DIF값을 Table 3과 Table 4에 나타내었다.

4. 계산 결과 및 고찰

이상에서 여러가지 방법으로 모수들을 결정해 보았고 또 그 결과들을 Table 3과 Table 4에서 비교해 보았다. Table에서 보듯이 자료 A에 대해서는 Parabolic Extrapolation Method가 가장 정도 높은 결과를 보여주고 있고 자료 B에서는 Gradient Search Method가 가장 정도 높은 결과를 보여주고 있다. 이들 각각의 누적 분포 함수와 확률 밀도 함수를 Fig. 10~Fig. 13에 각각 나타내었다.

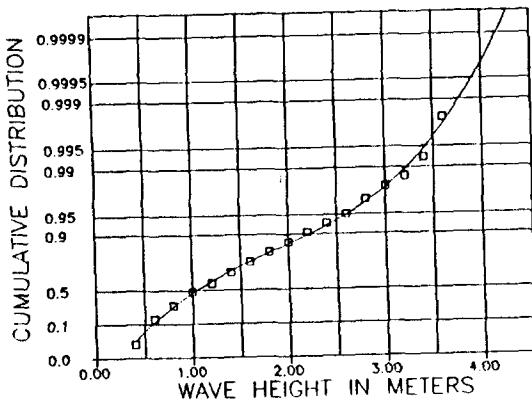


Fig. 10 CDF of data A

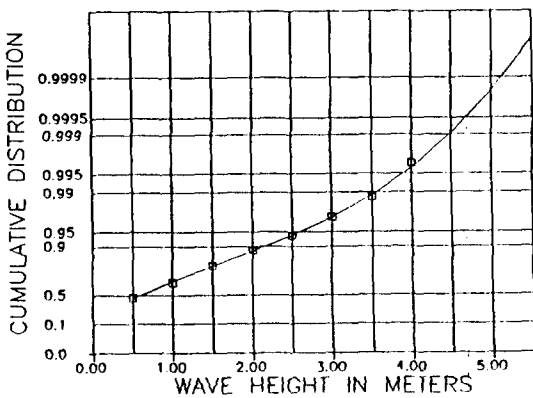


Fig. 11 CDF of data B

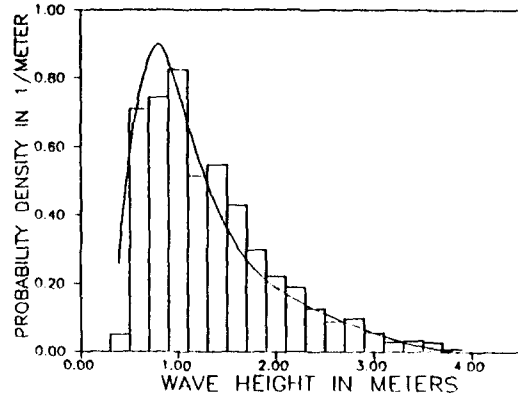


Fig. 12 PDF of data A

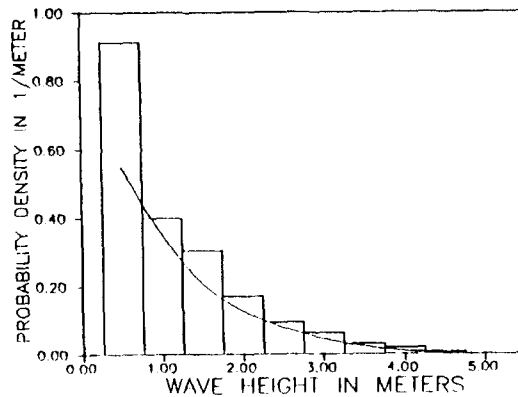


Fig. 13 PDF of data B

Table 3과 Table 4의 결과에서 보듯이 어떤 확률 밀도 형태에도 가장 우수한 방법으로 적용된다 할 만큼 뛰어난 방법은 결정하기가 어렵고, 또 극한 파고를 예측하는데 있어서 각 방법들의 차이도 미세하기 때문에 각 방법들을 적용시키는데는 별 문제가 없으리라 생각된다. 그러나 자료B에 대한 확률 밀도 함수에서 볼 수 있듯이 확률 밀도 함수를 구하는 데 있어서 Type III 분포를 적용시키기 곤란한 형태의 확률 밀도를 가지는 파고의 자료도 있다. 참고로 자료 A에 대한 Ochi의 fitting 결과를 Fig. 14에 나타내 보았고 100년의 재현 기간을 가지는 최대 파고를 예측 비교해 보았다.

Ochi : 4.3M (예측된 극한 파고)

Parabolic Extrapolation Method

: 4.31M (예측된 극한 파고)

또 자료 B에서 앞으로 100년 재현 기간동안 2월 중의 극한 파고는 4.8M로 예측할 수 있었다. 전체 방법들에 대한 극한파고 예측은 Table 5에 나타내었다.

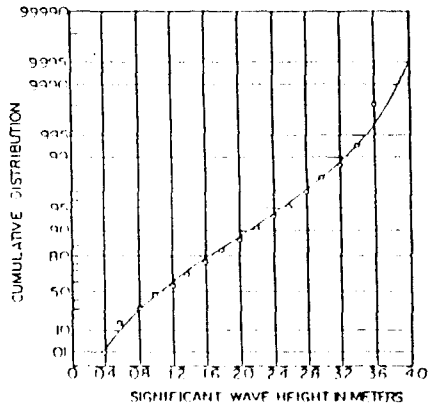


Fig. 14 CDF of data A by Ochi

Table 5 Extreme wave heights for return periods

Method	Return Period	50 Years		100 Years	
		Data A	Data B	Data A	Data B
Grid Search		4.35M	4.53M	4.46M	4.72M
Gradient Search		4.26M	4.59M	4.36M	4.80M
Gradient Expansion		4.23M	4.56M	4.32M	4.77M
Parabolic Extrapolation		4.23M	4.53M	4.31M	4.72M
Skewness		4.49M	4.55M	4.61M	4.74M
Maximum Likelihood		4.26M	4.51M	4.34M	4.67M

5. 결 론

Type III 분포의 모수를 이상에서 살펴 본 바와 같이 개선된 새로운 함수의 도입을 통하여 크게 3가지 방법으로 결정해 본 결과 다음과 같은 결론을 내릴 수 있다.

1. Non-Linear Regression Method는 전체적으로 다른 방법에 비하여 나은 결과들을 보여주고 있으나 실제 계산시 각 모수들에 대한 초기값을 계속 변화해가며 가장 나은 결과를 택해야 하는 번거로움이 있다.

2. Skewness Method는 Skewness Coefficient에 의하여 모수들을 구하는데 제한이 있을 수 있으나 일반적인 자료의 경우 별 부리 없이 적용시킬 수 있다. 그리고 Non-Linear Regression Method의 최선의 방법에 비해 정도는 조금 떨어지나 모수들을 식으로 부터 직접 구할 수 있기에 계산이 간편하고 빠른 잇점이 있다.

3. Maximum Likelihood Method도 좋은 결과를 보여주고 있으나 식이 복잡하여 계산이 까다롭고, Non-Linear Regression Method보다 더 초기값에 대하여 민감하기 때문에 실제 응용에는 힘들다고 본다.

4. 일반적으로 Skewness Method로 각 모수들을 결정한 다음 이 fitting된 누적 분포 함수와 실제 자료의 누적분포 함수와의 차이를 다시 fitting하여 최정적인 누적분포 함수를 구하는 방법이 가장 추천할 만 하다.

참 고 문 헌

[1] Ochi, Michel K. "Estimation of Extreme Sea Severity From Mmeasured Daily Maxima", 20th International Conference on Coastal Engineering in Taipei, Taiwan, 1986.
 [2] Alfredo H-S. Ang, "Probability Concept In Engineering Planning And Design", Wiley, 1984.
 [3] Bevington, Philip R. "Data Reduction and Error Analysis For The Physical Sciences", McGraw-Hill, 1969.
 [4] Gumbel, E. "Statistics of Extremes", Columbia Univ. Press, New York, 1958.