

국내 연안 유의파고 관측자료의 최적 확률분포 추정 및 평가 Estimation and Assessment on the Optimal Probability Distribution of the Monitoring Wave Height Data in the Korean Coast

이옥재* · 조홍연** · 고동휘***

Uk-Jae Lee*, Hong-Yeon Cho** and Dong-Hui Ko***

요지 : 파랑 자료의 확률분포 모델 추정은 파랑 자료가 어떤 확률적 성질을 가지는지 이해하고, 공학적 설계 변수의 정량적 평가를 가능하게 한다. 이에 본 연구에서는 국내 연안 20개 정점에서 관측된 유의파고 자료를 기반으로 최적 확률분포를 추정하고, 그 자료의 확률적 특성과 꼬리 거동을 분석하였다. 확률분포모델로는 총 9개(Normal, Log-normal, Weibull-3, Gamma-3, Generalized Gamma, Generalized Beta-2, Johnson SB, Generalized Hyperbolic, Extended Generalized Inverse Gaussian)의 모델을 이용하였으며, 각 분포의 매개변수 추정법으로는 최대우도추정법을 사용하였다. 모델의 성능을 평가하기 위한 적합도 검정 방법으로는 Kolmogorov-Smirnov 검정과 Kullback-Leibler Divergence 지표를 사용하였으며, Q-Q plot을 통해 확률 분포 꼬리 영역에서의 적합성을 분석하였다. 그 결과, Generalized Hyperbolic, Johnson SB, Generalized Gamma 분포가 많은 지점에서 높은 적합도를 보였으며, 복수의 형상 매개변수를 갖는 확률분포들이 유의파고의 비대칭성과 꼬리 특성을 잘 설명하는 것으로 나타났다. 그러나, 일부 확률분포 모델에서는 극단 분위수 영역에서 유의파고를 과소 또는 과대 추정하는 경향을 보여 유의파고 자료를 보다 정확하게 설명하기 위해서는 대상 지점의 파랑 자료에 적합한 확률분포를 추정하여 사용하는 것이 적절할 것으로 판단되었다.

핵심용어 : 유의파고, 최적 확률분포, 최대우도추정법, 분포 적합도 검정, 꼬리 거동

Abstract : The estimation of probability distribution models for wave data enables an understanding of the probabilistic characteristics of wave conditions and the quantitative evaluation of engineering design variables. In this study, the optimal probability distributions were fitted using by significant wave height data observed at 20 points around the Korean coast, and the probabilistic characteristics and tail behaviors of the data were analyzed. A total of nine distribution models were considered, including the Normal, Log-normal, Weibull-3, Gamma-3, Generalized Gamma, Generalized Beta-2, Johnson SB, Generalized Hyperbolic, and Extended Generalized Inverse Gaussian distributions. The parameters for each distribution were estimated using the Maximum Likelihood Estimation (MLE) method. To evaluate the performance of each model, goodness-of-fit tests such as the Kolmogorov-Smirnov test and the Kullback-Leibler divergence metric were used, and Q-Q plots were employed to assess the fit in the tail regions of the distributions. As a result, the Generalized Hyperbolic, Johnson SB, and Generalized Gamma distributions showed high goodness-of-fit at many stations. Distributions with multiple shape parameters were found to effectively capture the asymmetry and tail characteristics of the significant wave height data. However, some distribution models tended to underestimate or overestimate the extreme quantile regions. Therefore, it is considered appropriate to select and apply the probability distribution model that best fits the wave data at each target site in order to more accurately describe the characteristics of significant wave heights.

Keywords : significant wave height, optimal probability distribution, maximum likelihood estimation, goodness-of-fit test, tail behavior

*한국해양과학기술원 해양공간개발·에너지연구부 연수연구원(Post Doctoral Scientist, Ocean Space Development and Energy Research Department, Korea Institute of Ocean Science and Technology)

**한국해양과학기술원 해양빅데이터·AI센터 책임연구원(Principal Research Scientist, Marine Bigdata·AI Center, Korea Institute of Ocean Science and Technology)

***한국해양과학기술원 해양공간개발·에너지연구부 책임연구원(Corresponding author: Dong-Hui Ko, Principal Research Scientist, Ocean Space Development and Energy Research Department, Korea Institute of Ocean Science and Technology, 385, Haeyang-ro, Yeongdo-gu, Busan 49111, Korea, Tel: +82-51-664-3530, kodh02@kiost.ac.kr)

1. 서 론

파랑은 해안 및 항만 구조물의 설계, 선박 설계 등에 있어서 매우 중요한 설계 변수로 작용하며, 시시각각 변동하는 파랑의 불확실성을 정량화하기 위한 노력이 지속적으로 요구되고 있다. 특히 해역별 파랑 특성은 지형과 기상에 따라 다르기 때문에 각 해역의 파랑 특성은 관측 자료를 기반으로 확률분포 특성을 이해하는 것은 중요하다(Young et al., 2011; Young and Ribal, 2019; Haselsteiner and Thoben, 2020; Shamshirband et al., 2020; Dehghan et al., 2022). 이러한 유의파고의 확률적 특성을 정량적으로 표현하기 위한 대표적인 방법으로는 확률분포 모델을 활용한 통계적 접근이 주로 사용되고 있으며, 이는 파랑의 발생 빈도, 크기 및 극치값 등을 포괄적으로 설명할 수 있다는 점에서 널리 사용되고 있다(Ferreira and Soares, 1999; Ferreira and Soares, 2000). 또한 다양한 분야의 국제표준에서는 대상 해역의 극치해상상태를 평가하기 위한 방법으로 I-FORM(Inverse-First Order Reliability Method) 기법을 이용한 환경 등고선 작도를 적극적으로 권장하고 있다. 이때 입력변수로 사용되는 파고, 주기 등의 해양환경변수는 서로 다른 확률분포 특성을 가지며, 이들의 분포 형태는 등고선의 위치와 형상에 결정적인 영향을 미친다. 따라서 각 변수에 대해 통계적 적합도가 검증된 확률분포를 적용하는 것은 필수적이다(Mackay and Haselsteiner, 2021; Clarindo and Soares, 2024). 최근에는 Log-normal 분포 및 Weibull-3 분포와 같이 주로 사용되는 분포보다는 보다 정확한 분포 형상 조절이 가능하며, 극치영역의 재현성이 높은 4~6개의 매개변수를 갖는 비교적 복잡한 확률분포를 적용하여 유의파고 분포를 정확하게 추정하는 연구가 활발히 이루어지고 있다(Wu et al., 2016; Soukissian, 2021). 특히 Weibull-3의 경우, 낮은 파고 영역에서의 과대 추정과 높은 파고 영역에서의 과소 추정이 빈번하게 발생한다는 단점이 있다. 이는 상대적으로 적은 매개변수의 개수로 분포의 형상을 재현함에 있어 유연성이 제한되며, 실제 유의파고 확률분포의 비대칭성과 꼬리 두께를 충분히 반영하지 못한다는 한계를 갖는다(James and Panchang, 2022; Vanem and Fazerres-Ferradosa, 2022).

한편, Soukissian(2021)은 그리스 해역 4개 지점, 미국 NCDC 부이 4개 지점을 대상으로 복수의 매개변수를 갖는 다양한 확률분포 모델을 적용하여 유의파고의 적합도를 비교하였으며, 확률분포 모델 식의 매개변수 수가 증가할수록 분포 형상에 대한 조절 능력이 향상되어 유의파고의 비대칭적 분포 특성 및 극치 영역을 보다 정확하게 재현하고 있음을 확인한 바 있다. 또한, Wu et al.(2016)은 기존의 Rayleigh 및 Weibull 분포가 수심이 얇은 해역에서 발생하는 쇄파고의 분포를 추정하기 위해 유의파고의 중심 및 꼬리 부분을 각각 Weibull 분포와 Generalized Pareto 분포로 적용하여, 두 분포를 하나의 연속된 모델로 유연하게 결합한 Weibull-Generalized

Pareto(WGP) 분포를 제안하였다. 제안된 WGP 모델은 유의파고를 기준으로 두 분포가 자연스럽게 적용되도록 설계됐으며, 기존 모델보다 극치 영역의 파고 분포를 더욱 정밀하게 재현하는 것으로 나타났다. 한편, 국내에서 수행된 유의파고 확률분포 추정과 관련된 연구를 살펴보면, Jeong et al.(2015)는 국내 동해안에 위치한 대진 연안의 장기 파랑 특성을 분석하기 위하여 약 7년간 관측된 유의파고와 침투주기를 대상으로 최적 확률분포를 추정한 바 있다. 이 과정에서 사용된 후보 분포함수는 Kernel, Generalized Gamma, Generalized Extreme Values(GEV) 분포를 사용하였으며, 최소제곱법과 최대우도추정법을 사용하여 각각의 분포함수 내 매개변수를 추정하였다. 또한, Ko et al.(2017)에서는 중규모 유럽 기상예보센터 재분석 자료를 활용하여 국내 연안 유의파고의 확률분포 특성을 검토하기 위해 Lognormal, Weibull 및 GEV 분포를 적용하였으며, 각각의 확률분포의 매개변수 추정을 위해 최대우도추정법이 사용된 바 있다. 아울러, Lee et al.(2024)의 경우, 유의파고와 침투주기 사이의 결합확률분포 분석을 위하여 Normal, Lognormal, Gamma, Weibull-3 분포를 후보 분포로 결정하여 최대우도추정법을 통해 최적 확률분포를 추정한 바 있다. 유의파고에 대한 적합 확률분포 추정은 주로 국외에서 활발하게 수행되고 있으나, 국내에서는 대부분 특정 해역을 대상으로 특정 분포를 적용하거나, 비교적 간단한 형태의 분포 모델을 적용한 경우가 대부분이다.

이에 따라 본 연구에서는 우리나라 연안 20개 지점에서 관측된 유의파고 관측자료를 활용하여 Log-normal, Weibull-3, Gamma-3 분포와 더불어 4~6개의 매개변수를 가지는 대표적인 확률분포함수를 대상으로 최적 분포함수를 추정하였으며, 매개변수는 최대우도추정법을 이용하였다. 또한 추정된 분포함수에 대한 분포 적합 검정은 KS 방법으로 수행하였으며, Kullback-Leibler Divergence 측도를 이용하여 모든 지점의 유의파고 분포에 대한 적합 수준을 정량적으로 평가하였다. 또한, 평가된 분포함수에 대하여 Q-Q(Quantile-Quantile) 플롯을 이용하여 유의파고의 극단값에 해당하는 꼬리 거동을 검토하였으며, 적합 확률분포 평가 결과를 바탕으로 지점별로 최적 또는 적합 상위 분포함수의 특성 및 순위에 대한 종합 평가를 실시하여 국내 지점 별 적합한 분포함수 순위를 제시하였다.

2. 관측자료 및 방법

2.1 국내 파랑 관측자료

본 연구에서는 전국 파랑 관측 자료 제공 시스템(WINK: Wave Information Network of Korea)에서 제공하는 해양수산부(MOF) 및 국립해양조사원(KHOA)에서 관측한 자료를 대상으로 유의파고의 분포함수 추정을 수행하였다. 해양수산부는 전국 34개 지점에서 AWAC 600, Signature ADCP 500 등의 파랑 관측 장비를 활용하여 유의파고 및 주기 등 파랑

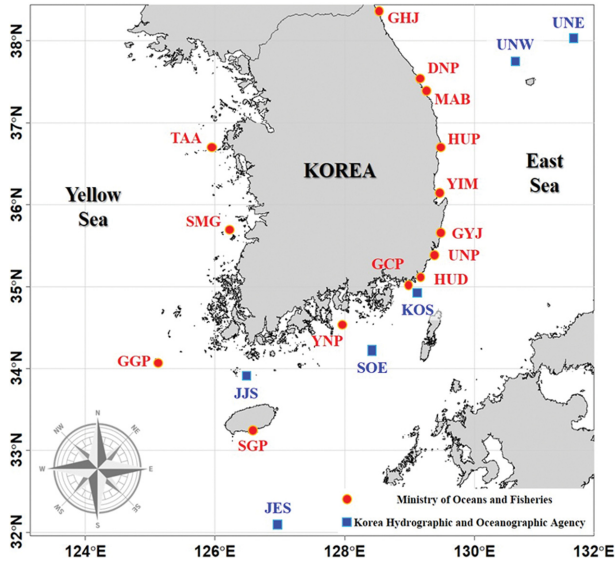


Fig. 1. Locations of significant wave height observation sites (Red circle: MOF; Blue squares: KHOA).

자료를 30분 간격으로 수집하고 있으며, 관측 기간은 최소 약 3개월에서 최대 12년 이상으로 지점에 따라 차이를 보이고 있다. 국립해양조사원은 6개의 대형 해양관측부이에서 파랑 관측을 수행하고 있으며, 파랑 관측은 GPS 방식인 Datawell 사

에서 개발한 MOSE G-1000을 통해 수면위 3.5 m에서 관측을 수행한다. 또한, 대형 해양관측부이에서 관측되는 파랑 자료의 관측 간격은 30분 간격, 최대 관측 기간은 약 12년이다.

본 연구에서 선정한 분석 대상 지점은 관측 기간이 3년 이상인 지점으로 한정하였으며, 이는 유의파고의 통계적 신뢰도를 확보하기 위해 최소 3년 이상의 관측 기간이 필요하다는 선행 연구를 근거로 결정하였다(Cho et al., 2013). 한편, 분석에 사용된 해양수산부와 국립해양조사원의 유의파고 자료는 기관별로 산출 방식에 차이가 있으며, 해양수산부에서 관측된 유의파고의 경우, 스펙트럼 분석법을 이용하고 있으며, 국립해양조사원의 경우, 파열 분석법을 이용한다. 분석 대상 지점의 위치도 및 관측 정보는 다음 Fig. 1 및 Table 1과 같다.

2.2 유의파고 확률분포 후보 함수

본 연구에서는 유의파고의 최적 확률분포를 추정하기 위해 Normal 분포, Log-normal 분포, Weibull-3 분포, Gamma-3 분포를 이용하였다. 또한, 보다 다양한 확률분포를 적용하기 위하여 4-매개변수 기반의 Johnson SB 분포, Generalized beta distribution of the 2-kind 분포 및 Generalized Gamma 분포, 5-매개변수 기반의 Generalized Hyperbolic 분포와 6-매개변수 기반의 Extended Generalized Inverse Gaussian 분

Table 1. Summary of significant wave height observation sites

No	Type	Station (Code)	Longitude, Latitude (°)	Recording Period	Depth (m)	Instrumentation Type
1	MOF	Taeam (TAA)	125.94°E, 36.72°N	2014/07/12~2024/08/04	26.2	AWAC 600
2		Saemangeum New Port (SMG)	126.24°E, 35.67°N	2020/07/29~2024/10/10	20.2	Signature ADCP 500
3		Gagudo Port (GGP)	125.13°E, 34.05°N	2020/08/17~2024/12/10	35.0	Signature ADCP 500
4		Seogwipo Port (SGP)	126.58°E, 33.24°N	2020/06/09~2024/11/12	31.0	Signature ADCP 500
5		Yeosu New Port (YNP)	127.96°E, 34.53°N	2018/12/13~2024/10/11	32.0	Signature ADCP 500
6		Gamcheon Port (GCP)	129.00°E, 35.01°N	2017/09/22~2024/10/10	36.0	Signature ADCP 500
7		Haeundae (HUD)	129.17°E, 35.12°N	2020/07/03~2024/10/11	34.0	Signature ADCP 500
8		Ulsan New Port (UNP)	129.38°E, 35.39°N	2018/11/01~2024/10/16	29.0	Signature ADCP 500
9		Gyeongju (GYJ)	129.48°E, 35.67°N	2014/07/21~2024/06/30	32.1	AWAC 600
10		Yeongilman (YIM)	129.48°E, 36.14°N	2020/06/24~2024/12/17	33.0	Signature ADCP 500
11		Hupo (HUP)	129.49°E, 36.70°N	2011/11/12~2024/06/30	31.0	AWAC 600
12		Maengbang (MAB)	129.23°E, 37.40°N	2013/09/27~2024/06/30	31.0	AWAC 600
13		Donghae New Port (DNP)	129.16°E, 37.51°N	2020/07/07~2024/09/30	35.0	Signature ADCP 500
14	KHOA	Gonghyeonjin (GHJ)	128.53°E, 38.36°N	2016/04/29~2024/06/30	32.0	AWAC 600
15		Jeju Strait (JJS)	126.49°E, 33.91°N	2012/09/25~2024/09/30	65.0	MOSE G-1000
16		Jeju south (JES)	126.97°E, 32.09°N	2012/09/08~2024/09/30	127.0	MOSE G-1000
17		South east (SOE)	128.42°E, 34.22°N	2012/09/06~2024/09/30	84.0	MOSE G-1000
18		Korea Strait (KOS)	129.12°E, 34.92°N	2012/09/05~2024/09/30	95.0	MOSE G-1000
19		Ulleungdo NW (UNW)	130.60°E, 37.74°N	2012/11/23~2024/09/30	1,300.0	MOSE G-1000
20		Ulleungdo NE (UNE)	131.55°E, 38.01°N	2012/11/24~2024/09/30	966.0	MOSE G-1000

포를 후보 분포함수로 선정하였다.

이에 각각의 분포에 대한 설명 및 확률밀도함수는 다음과 같으며, 분포 추정 대상이 되는 유의과고 변수(H)는 일반적으로 이용되는 변수기호(x , 변수집합기호 X)를 사용하고, 확률밀도함수와 누적확률분포함수의 기호는 각각 f , F 기호를 사용하였다. 또한, $dF(x)/dx = f(x)$ 조건은 두 함수의 관계를 의미하는 제약조건임을 의미한다.

1) Normal Distribution (N)

정규분포는 가장 널리 사용되는 연속 확률분포로, 평균, 분산 두 개의 매개변수를 가지며 데이터가 평균값을 중심으로 대칭적으로 분포하는 특성을 가지며, 다음 식(1)과 (2)와 같이 표현할 수 있다.

$$f(x; a, b) = \frac{1}{\sqrt{2\pi}b} \cdot \exp\left(-\frac{1}{2} \frac{(x-a)^2}{b^2}\right) \quad (1)$$

$$F(x; a, b) = \Phi\left(\frac{x-a}{b}\right) \quad (2)$$

여기서, a , b 는 각각 분포함수의 위치 매개변수(location parameter), 척도 매개변수(scale parameter)를 의미하며, 분포 함수에 종속된다. 이어지는 분포함수에서도 같은 의미의 매개변수에 해당한다. Φ 는 표준정규분포의 누적분포함수를 의미한다(Patel and Read, 1996).

2) Log-normal Distribution (LN)

로그정규분포는 변수의 로그($x' = \ln x$)가 정규분포를 따르는 분포로, 변수변환에 따른 분포추정을 적용하면 유의과고와 같이 비대칭적이고 양의 값으로만 정의되는 경우에 적합하며, 다음 식(3)과 (4)와 같이 표현할 수 있다(Athanassoulis et al., 1994).

$$f(x; a, b) = \frac{1}{bx\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(\ln x - a)^2}{b^2}\right) \quad (3)$$

$$F(x; a, b) = \Phi\left(\frac{x-a}{b}\right) \quad (4)$$

3) Gamma-3 Distribution (G3)

Gamma 분포는 변수가 양의 실수 범위($x - a > 0$)에서 정의되며, 위치, 척도, 형상 매개변수에 따라 분포의 형태가 달라지는 비대칭 분포로서, 기본 형태인 척도, 형상 매개변수에 위치매개변수를 도입하는 3개의 매개변수를 가지는 분포함수이다. G3 함수는 형상 매개변수(k , shape parameter)를 포함하는 형태로 다음 식(5)와 (6)과 같이 표현할 수 있다(Bobee and Ashkar, 1991).

$$f(x; a, b, k) = \frac{1}{\Gamma(k)b^k} (x-a)^{k-1} \cdot \exp\left(-\frac{x-a}{b}\right) \quad (5)$$

$$F(x; a, b, k) = \frac{1}{\Gamma(k)} \cdot \gamma\left(k, \frac{x-a}{b}\right) \quad (6)$$

여기서, $\Gamma(\cdot)$ 는 감마함수, $\gamma(\cdot, \cdot)$ 는 하위(lower) 불완전 감마함수를 의미한다.

4) Weibull-3 Distribution (W3)

Weibull-3 분포는 형상, 척도 및 위치 매개변수로 구성되며, 다양한 분포 형태를 표현할 수 있는 유연성을 가진다. 특히 해당 분포는 유의과고 표현에 있어 널리 사용되고 있으며, 다음과 같은 조건($x - a > 0$)에서 다음 식(7)과 (8)과 같이 표현할 수 있다(Weibull, 1951; Rao and Hamed, 2000).

$$f(x; a, b, k) = \frac{k}{b} \left(\frac{x-a}{b}\right)^{k-1} \cdot \exp\left[-\left(\frac{x-a}{b}\right)^k\right] \quad (7)$$

$$F(x; a, b, k) = 1 - \exp\left[-\left(\frac{x-a}{b}\right)^k\right] \quad (8)$$

5) Generalized Beta distribution of the 2-kind Distribution (GB)

제 2종 GB 분포의 경우, 이전의 분포와는 달리 형상 매개변수를 세분화하는, 다수의 형상매개변수를 도입하는 형태의 함수이다. 이 분포함수는 3개의 형상 매개변수와 1개의 척도 매개변수(b)를 통해 다양한 분포 형태(형상)를 비교적 정확히 표현할 수 있으며, 특히 왜도 및 첨도가 큰 분포에 적합한 장점이 있다. GB 분포는 다음과 같이 표현이 가능하며, 형상 매개변수는 k , p , q 기호로 추가하여 표현한다(식(9)와 (10) 참조).

$$f(x; b, k, p, q) = \frac{kx^{kp-1}}{b^{kp} \cdot B(p, q) \left(1 + \left(\frac{x}{b}\right)^k\right)^{p+q}} \quad (9)$$

$$F(x; b, k, p, q) = I_z(p, q) \quad (10)$$

여기서, $B(p, q)$ 는 Beta 함수, $I_z(p, q)$ 는 불완전 베타 함수를 정규화한 불완전 Beta 함수를 의미하며, 형상 매개변수를 입력으로 요구하며, 다음 식(11)과 (12)와 같이 표현된다(McDonald and Xu, 1995).

$$B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt \quad (11)$$

$$I_z(p, q) = \frac{B(z, p, q)}{B(p, q)} = \frac{1}{B(p, q)} \int_0^z t^{p-1} (1-t)^{q-1} dt \quad (12)$$

여기서, $t = (x/b)^k$, $z = t/(1+t)$ 이다.

6) Johnson SB Distribution (JSB)

Johnson SB 분포는 각각 하나의 위치 매개변수(a), 척도 매개변수(b), 그리고 두 개의 형상 매개변수(p, q)를 포함하는 4-매개변수 분포로서, 분포의 비대칭성 및 첨도 조절이 가능한

구조를 가지며, 다음 식(13)과 (14)와 같이 표현할 수 있다.

$$f(x; a, b, p, q) = \frac{ap}{(x-a)(b-(x-a))\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2}\left(q + p \ln\left(\frac{x-a}{b-(x-a)}\right)\right)^2\right\} \quad (13)$$

$$F(x; a, b, p, q) = \Phi\left(q + p \ln\left(\frac{x-a}{b-(x-a)}\right)\right) \quad (14)$$

이 분포에서 위치 매개변수인 a 는 분포의 시작점을 조절하는 매개변수로서 변수는 $a \leq x \leq a + b$ 범위로 한정된다. 형상 매개변수 중 p 는 분포의 기울기 또는 첨도를 조절하는 등 분포의 폭과 관련된 매개변수이다(Johnson, 1949).

7) Generalized Gamma Distribution (GG)

Generalized Gamma 분포는 네 개의 매개변수(위치 및 척도 매개변수와 두 개의 형상 매개변수; a, b, p, q 기호 사용)로 구성된 분포로서, 다양한 분포 형태를 표현할 수 있는 확률 분포이며, 다음 식(15)와 (16)과 같이 표현된다(Stacy, 1962; Lawless, 1980).

$$f(x; a, b, p, q) = \frac{p}{b\Gamma(p/q)}\left(\frac{x-a}{b}\right)^{pq-1} \cdot \exp\left[-\left(\frac{x-a}{b}\right)^p\right] \quad (15)$$

$$F(x; a, b, p, q) = \frac{1}{\Gamma(d)}\gamma\left(q, \left(\frac{x-a}{b}\right)^p\right) \quad (16)$$

8) Generalized Hyperbolic Distribution (GH)

Generalized Hyperbolic 분포는 총 다섯 개의 매개변수(위치, 척도 매개변수와 3개의 형상 매개변수; a, b, k, p, q)를 가지며, 분포의 형태, 비대칭성, 척도 및 위치를 조절할 수 있는 확률분포이다. 특히, 해당 분포는 Normal Inverse Gaussian, Hyperbolic, Student-t 분포 등의 특수한 형태를 포함하고 있어, 유의파고와 같이 꼬리가 두껍고 비대칭적인 관측자료를 재현하는데 적합하며, 분포함수는 다음 식(17)과 (18)과 같이 표현할 수 있다.

$$f(x; a, b, k, p, q) = \frac{\exp[q(x-a)]d^k}{\sqrt{2\pi p^{k-1/2}}bK_k(bd)} \cdot \frac{K_{k-1/2}(p\sqrt{b^2 + (x-a)^2})}{(\sqrt{b^2 + (x-a)^2})^{-(k-1/2)}} \quad (17)$$

$$F(x; a, b, k, p, q) = \int_a^x f(x'; a, b, k, p, q)dx' \quad (18)$$

여기서, $d^2 = p^2 - q^2$, $d = \sqrt{p^2 - q^2}$, $K_k(\bullet)$ 는 제2종 변형 Bessel 함수, k, p, q 는 형상 매개변수를 의미한다. 형상 매개변수 중 k 는 중심의 뾰족함, p 는 분포의 꼬리 두께, q 는 비대칭성을 결정하는 역할을 한다. 한편, 해당 분포의 누적분포

함수(CDF)는 제2종 변형 Bessel 함수 및 비선형 항의 존재로 인해 일반식이 존재하지 않으며, 일반적으로 수치적 방법으로 근사하거나, 자료 기반 경험적인 누적 분포를 활용하여 계산한다(Barndorff-Nielsen, 1977; Prause, 1999; Ardia et al., 2019).

9) Extended Generalized Inverse Gaussian Distribution (EGIG)

이 분포는 총 6개의 매개변수(위치, 척도 매개변수, a, b ; 4개의 형상매개변수, k, s, p, q)를 포함하는 확률분포로서, 유의파고와 같이 비대칭적이며, 긴 꼬리를 가지는 자료에 적합하게 설계되었다. 이 분포는 일반화된 역 감마 분포와 비슷한 형태를 가지며, Bessel 함수 기반의 정규화 상수를 포함하여 다양한 분포 형태를 재현할 수 있다. 이러한 분포는 다음 식(19)와 (20)과 같이 표현할 수 있다.

$$f(x; a, b, k, s, p, q) = \frac{q}{2b}\left(\frac{k}{s}\right)^{q/(2p)} \frac{(x-a)^{q-1}}{K_{q/p}(2\sqrt{ks})} \cdot \exp\left(-k\left(\frac{x-a}{b}\right)^p - s\left(\frac{x-a}{b}\right)^{-p}\right) \quad (19)$$

$$F(x; a, b, k, s, p, q) = \int_a^x f(x'; a, b, k, s, p, q)dx' \quad (20)$$

여기서, 형상 매개변수의 경우, k, s 는 분포 꼬리의 두께 및 비대칭을 조절하는 Shape-Scale 역할을 하며, p 는 중심 및 꼬리 부분의 곡률을 결정하고, q 는 분포의 첨도 및 분포 형태의 전반적인 형상을 결정한다. 또한, 해당 분포는 GH 분포와 마찬가지로 복잡한 함수 형태로 인해 누적분포함수 수식은 존재하지 않으며, 수치 적분 또는 경험 누적분포를 통해 근사적으로 계산하게 된다(Shakil et al., 2010; Gupta and Viles, 2011, 2012; Low et al., 2017).

2.3 최대우도추정법

확률분포 모델의 매개변수를 추정하기 위해서는 최소제곱법, 확률가중모멘트법, 최대우도추정법 등이 주로 이용된다. 최소제곱법은 추정 방법이 비교적 간단하고 직관적이나 편향(bias)이 발생하는 단점이 있으며, 확률가중모멘트법은 자료의 극단값에 덜 민감한 장점이 있다. 그러나 두 방법 모두 분포의 수학적 특성과 자료의 통계적 특성을 적절하게 반영하지 못하는 경우가 발생할 수 있으며, 특히 매개변수의 수가 증가할수록 추정의 정확도와 안정성에 한계를 보인다(Clauset et al., 2009; Katalinic and Parunov, 2020). 반면, 최대우도추정법은 주어진 자료에 대해 가능도(likelihood)를 극대화하는 통계적인 개념을 기반으로 하는 방식으로 복잡한 형태의 확률분포에 대하여 일관성있고 효율적이며 통계적으로 신뢰할 수 있는 추정을 가능하게 한다(Clauset et al., 2009; Pascal et al., 2013; Soukissian, 2021).

본 연구에서는 유의파고의 적합 확률분포를 추정하기 위해 최대우도추정법(MLE, Maximum Likelihood Estimation)을 이용하였다. 최대우도추정법은 관측자료에 가장 유사한, 관측자료가 독립이라는 조건에서 모든 자료의 발생확률을 최대화하는 조건에서 확률분포의 매개변수를 추정하는 통계적 방법이다. 특정 확률밀도함수 $f(H_i|\theta)$ 에 대하여, 관측값 $H_i(i = 1, 2, \dots, n, n = \text{자료의 개수})$ 가 독립적으로 추정되었다고 가정하면, 우도함수(L)는 다음 식(21)과 같이 정의된다.

$$L(\theta|H_i) = \prod_{i=1}^n f_i(H_i|\theta) \quad (21)$$

계산의 효율성을 위하여 일반적으로 로그우도 함수를 사용하며, 이는 다음 식(22)와 같이 표현할 수 있다.

$$\ln L(\theta|H_i) = l(\theta|H_i) = \sum_{i=1}^n \ln(f(H_i|\theta)) \quad (22)$$

최대우도추정법은 상기 로그우도함수를 최대화하는 모수 벡터($\theta = \theta_k, k = 1, 2, \dots, n_p, n_p = \text{추정 매개변수의 개수}$)를 추정하는 문제로서, 이는 다음 식(23)과 같이 정리된다.

$$\hat{\theta} = \arg \max_{\theta} l(\theta|H_i) \quad (23)$$

일반적으로 로그우도함수는 비선형함수형태이기 때문에, 해석적으로 극대값을 구할 수 있는 경우가 제한된다. 따라서 수치적 최적화 기법을 이용한 추정이 수행된다. 특히, 본 연구와 같이 다수의 매개변수를 갖는 확률분포의 경우, 매개변수의 상·하한 경계 조건을 고려할 수 있는 최적화 알고리즘인 L-BFGS-B(Limited-memory Broyden-Fletcher-Goldfarb-Shanno with Bounds) 기법을 활용하여 최대우도 추정을 수행하였다(Byrd et al., 1995; Nocedal and Wright, 2006).

2.4 적합도 검정 및 적합 측도 평가

유의파고 변수에 대한 후보 확률분포의 분포 적합 평가를 위하여 본 연구에서는 널리 이용되는 Kolmogorov-Smirnov 검정 기법을 이용하였다. 더불어, 이 검정을 통과한 후보 분포에 대해서는 Kullback-Leibler Divergence 측도를 이용하여 후보 분포함수 간의 분포 적합 정도와 순위를 결정하는데 이용하였다.

1) Kolmogorov-Smirnov 분포 적합 검정

유의파고의 확률분포 적합성을 정량적으로 평가하기 위해 Kolmogorov-Smirnov 검정(KS 검정)을 이용하였다. KS 검정은 일반적으로 분포 적합성 검정을 위한 방법으로 널리 사용되는 방법으로서, 경험 누적분포함수와 이론 누적분포함수 간의 차이를 기반으로 통계적 유의성을 평가한다(Massey, 1951; Conover, 1999; Press, 2007). 이 검정의 귀무가설은 검정 대

상이 되는 “후보 분포함수와 기준이 되는 데이터 기반 분포함수와의 유의미한 차이는 없다.” 이므로, p-value 기반 판정의 경우에는 $p < 0.05$ (유의수준 = 0.05) 조건에서 귀무가설을 기각할 수 있다. 본 연구에서는 각 후보 확률분포에 대해 KS 검정통계량을 이용하여 p-value를 산정하였으며, 5% 유의수준에서 p-value가 0.05 이상인 경우, 해당 분포가 적합하다고 판단하였다.

2) Kullback-Leibler Divergence 측도

KS 검정을 통과한 후보 확률분포에 대하여, 추가적인 적합성 비교평가를 정량적으로 수행하기 위하여 Kullback-Leibler Divergence(KL-Divergence) 측도를 이용하였다. 이 방법은 경험적 확률분포와 후보 확률분포 간의 차이를 정량적으로 평가하는 지표로서, 다음 식(24)를 이용하여 각 확률분포의 기준 분포에 대한 차이를 정량화된 수치로 계산할 수 있다. 이 수치를 이용하여, 후보 분포함수의 적합 순위를 결정하였다(Belov and Armstrong, 2011; Wu et al., 2016).

$$D_{KL}(P|Q) = \sum_{i=1}^m \left[P(x_i) \cdot \ln \frac{P(x_i)}{Q(x_i)} \right] \quad (24)$$

여기서, $P(x_i)$ 는 관측된 유의파고의 확률밀도함수, $Q(x_i)$ 는 후보 분포의 확률밀도함수를 의미하며, x_i 는 분포 비교가 수행되는 구간으로 전체 파고 범위 영역을 유한한 구간($x_i, i = 1, 2, \dots, m$)으로 구분한 유의파고를 의미한다. 또한, 보다 정밀한 적합 확률분포를 평가하기 위하여 각 후보 확률밀도함수와 유의파고 관측자료 사이의 결정계수(r^2) 및 평균제곱근오차(RMSE)를 보조적으로 산정하여 제시하였다. 아울러, 평균제곱근오차는 유의파고와 후보 확률밀도함수 사이의 꼬리 영역을 평가하기 위해 사용된 Q-Q 플롯 분석의 평가지표로 활용하였다. 결정계수 및 평균제곱근오차는 다음 식(25)~(27)을 활용하였다.

$$r^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (25)$$

여기서, x_i 는 유의파고, $\hat{x}_i$ 는 후보 확률밀도함수, $\bar{x}$ 는 유의파고의 평균, n 은 비교에 사용된 자료의 개수를 의미한다.

$$RMSE_{PDF} = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_{obs}(x_i) - f_{fit}(x_i))^2} \quad (26)$$

$$RMSE_{QQ} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - f_{fit,i})^2} \quad (27)$$

여기서, $f_{obs}(x_i)$ 는 유의파고의 확률밀도함수, $f_{fit}(x_i)$ 는 후보 확률밀도함수, n 은 비교 구간의 개수를 의미하며, $Q_{obs,i}$ 는 누적확률의 관측 분위수, $Q_{fit,i}$ 는 $Q_{obs,i}$ 에 대응되는 이론 분위수, n 은 분위수 비교 지점의 개수를 의미한다.

3. 분포함수 매개변수 추정 및 분포 적합 검정 결과

3.1 분포함수 매개변수 최적 추정 결과

본 연구에서는 우리나라 연안 20개 지점을 대상으로 유의파고 관측자료에 적합한 확률분포함수를 평가하기 위하여 다양한 매개변수를 갖는 총 9개의 후보 확률분포함수를 적용하였다. 각 분포함수에 대한 매개변수는 최대우도추정법을 통해 산정하였으며, 모든 지점에서 산정된 모든 후보 분포함수 매개변수 값은 Appendix A에 제시하였다. 아울러, 최적 추정된 분포함수에 대한 KS 분포 적합 검정 결과(p-value)와 분포함수 사이의 적합 정도 비교를 위한 KL Divergence(D_{KL}) 측도를 추정한 결과는 Appendix B에 제시하였다.

비교평가 지표를 분석한 결과, GH 분포는 전체 20개 지점 중 8개 지점(40%)에서 적합 분포로 선정되었으며, 가장 많은 지점에서 유의파고 관측자료를 잘 재현하는 것으로 나타났다. JSB 분포는 5개 지점(25%)에서 최적 분포로 도출되었으며, GG 분포는 총 4개 지점(20%)에서 적합한 결과를 보였다. 아울러, GB 분포는 2개 지점(10%), EGIG 분포는 1개 지점(5%)에서 각각 최적 분포로 평가됨에 따라 전반적으로 분포의 적합도는 해역 및 지점별로 상이하게 나타났다. 또한, 각각의 지점에서 평가된 적합 확률분포의 결정계수는 0.98 이상으로 매우 높게 나타났으며, 평균제곱근오차는 0.01~0.09 m 수준으로 매우 낮게 나타남에 따라 적합하게 평가된 확률분포의 재현성은 우수한 것으로 판단된다 (Table 2 참조).

한편, 본 연구에서는 총 20개 지점에 대해 9개의 확률밀도함수를 적용하여 적합도를 평가하였으며, 이 중 적합 확률분

포로 선정된 GH, JSB, GG, GB 및 EGIG 분포에 대하여 대표 결과를 Fig. 2에 제시하였다. 각 분포별로는 해당 분포의 적합도가 가장 높은 1개 지점을 선정하였으며, 선정된 지점은 해역별 최소 1개 지점이 포함되도록 결정하였다.

각 지점에서 최적 확률분포로 추정된 분포들은 대체로 유의파고 분포의 침투 부분을 정확하게 재현하고 있으며, 특히 중심부에서 꼬리 영역으로 확장되는 구간에 대해 전반적으로 우수한 적합도를 보이는 것으로 나타났다. 또한, Appendix B에 제시한 바와 같이 모든 지점에서 4개 이상의 매개변수를 갖는 확률분포는 대부분 p-value 값이 0.05 이상인 것으로 나타났다. 상대적으로 2, 3 매개변수 기반의 확률분포는 유의파고 확률분포 추정에 비교적 적합하지 않은 것으로 나타났다. Fig. 2와 Appendix B에서 확인할 수 있듯이, 일반적으로 유의파고 분석에 활용되는 LN 및 W3 분포는 분포의 중심부와 꼬리 영역에서의 분포 형태를 충분히 재현하지 못하여 다수의 매개변수를 갖는 분포들에 비해 전반적인 적합성이 낮은 것으로 확인됐다.

지점에 따라 평가된 적합 확률분포의 공간 분포는 다음 Fig. 3과 같다. GH 분포는 서해안과 동해안 및 남해안 일부 지점에서 주로 우세하게 나타났으며, 전반적으로 국내 연안 전역에서 광범위하게 분포하는 경향을 보였다. JSB 분포의 경우, 주로 동해안과 제주 해역 인근에서 주로 적합한 결과를 보임에 따라 지역적 집중도가 GH 분포에 비해 비교적 뚜렷한 것으로 나타났으며, GB 분포는 동해 남부 연안의 일부 지점에서만 최적 분포로 평가된 것으로 나타났다. GG 분포는 서해, 남해, 동해에서 적합한 것으로 나타나 뚜렷한 공간적 경향은 나타나지 않았으며, EGIG 분포는 대한해협 하나의 지점에서 적합 분포로 평가됐다.

Table 2. The best fitted probability distribution with the corresponding p-value of the KS test, the KL-divergence measures for the twenty sites in Korea

No	Station Code	Optimal Distribution	p-value	D_{KL}	RMSE	r^2
1	TAA	Generalized Hyperbolic	0.269	25.875	0.063	0.987
2	SMG	Generalized Hyperbolic	0.288	20.635	0.068	0.982
3	GGP	Generalized Gamma	0.126	24.048	0.035	0.997
4	SGP	Johnson SB	0.167	30.276	0.081	0.979
5	YNP	Generalized Gamma	0.107	24.922	0.052	0.985
6	GCP	Generalized Beta 2	0.209	27.235	0.062	0.988
7	HUD	Generalized Gamma	0.174	24.838	0.042	0.989
8	UNP	Generalized Beta 2	0.217	25.146	0.035	0.993
9	GYJ	Generalized Hyperbolic	0.404	22.921	0.024	0.996
10	YIM	Johnson SB	0.345	25.964	0.042	0.986
11	HUP	Generalized Hyperbolic	0.403	18.498	0.020	0.996
12	MAB	Generalized Hyperbolic	0.417	20.037	0.022	0.998
13	DNP	Johnson SB	0.231	22.071	0.030	0.991
14	GHJ	Generalized Gamma	0.066	23.771	0.071	0.968
15	JJS	Johnson SB	0.212	20.056	0.024	0.999
16	JES	Generalized Hyperbolic	0.402	15.571	0.010	0.997
17	SOE	Generalized Hyperbolic	0.418	16.778	0.015	0.996
18	KOS	Extended G.I. Gaussian	0.365	18.846	0.008	0.998
19	UNW	Generalized Hyperbolic	0.389	17.464	0.011	0.997
20	UNE	Johnson SB	0.454	18.201	0.013	0.995

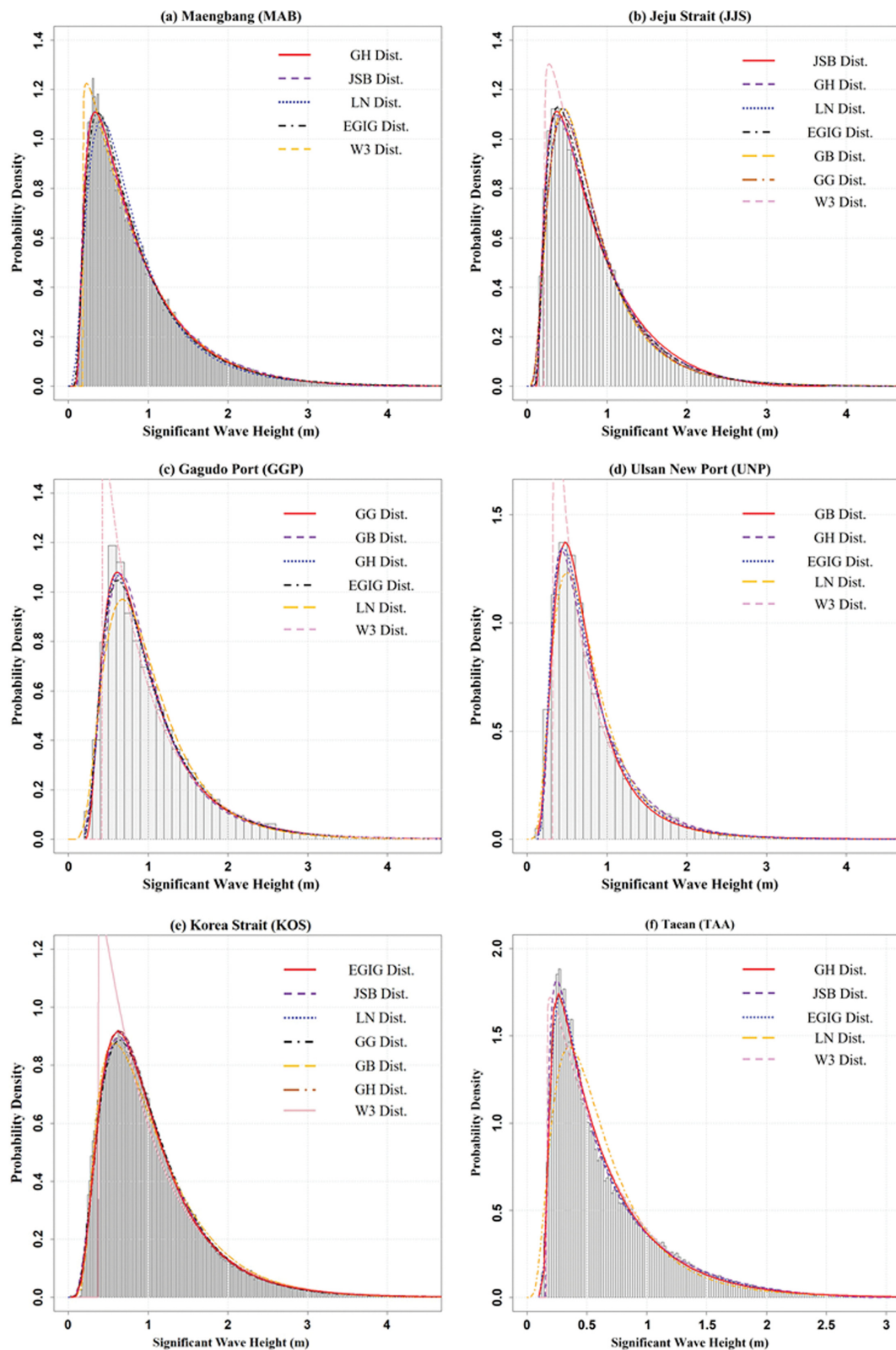


Fig. 2. Histogram and fitted probability density functions of significant wave height for representative stations.

3.2 유의파고 확률분포의 꼬리 영역 재현성 검토 결과

확률분포함수는 분포의 중심부 영역에 대해서는 일반적으로 높은 적합도를 보이지만, 꼬리 영역에 대해서는 통계적으로 불확실성이 증가함에 따라 적합도가 저하되는 한계가

존재한다. 이러한 한계를 보완하고 각 후보 분포의 꼬리 영역 재현성을 보다 정밀하게 검토하기 위하여, 본 연구에서는 분위수를 기반으로 하는 Q-Q(Quantile-Quantile) 플롯을 이용한 분석을 수행하였다. Q-Q 플롯은 후보 확률분포의 이

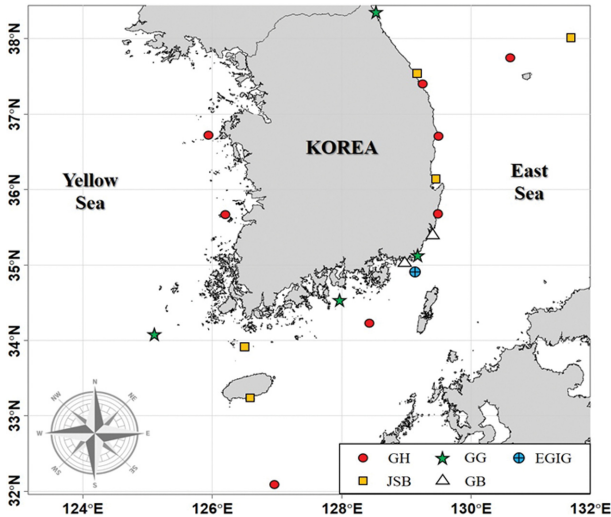


Fig. 3. Spatial distribution of optimal probability distributions of significant wave height around Korean coast.

론 분위수와 관측자료의 분위수를 비교함으로써 전반적인 분포 적합도를 직관적으로 판단할 수 있으며, 특히 극단값이 위치한 꼬리 영역에 대한 검토가 가능하여 널리 사용되고 있다(Soukissian, 2021).

아울러 Fig. 2에서 제시한 지점에 대하여 Q-Q 플롯을 가시화하였으며, 각 후보 분포 별 분위수와 관측값 분위수 사이의 RMSE를 산정하여 수치적 적합도를 정량적으로 비교하였다(Fig. 4 참조). 분석 결과, 대체로 99% 분위수 이하 구간에서는 대체로 후보 분포 간 뚜렷한 차이는 보이지 않았으나, 99% 이상의 분위수 구간에서는 분포 별 꼬리 영역의 재현성 차이가 뚜렷하게 나타나는 경향을 보였다. MAB 지점에서는 GH 분포가 RMSE 0.156 m로 가장 낮은 값을 보였으며, 전반적인 적합 성능이 우수하였으나, 99.99% 분위수 이상의 극치 영역에서는 관측값을 과대 추정하는 경향이 나타났다. JJS 지점의 경우, JSB 분포는 99.9% 분위수까지는 안정적인 적합도를 보였으나, 보다 높은 분위수 영역에서는 EGIG 분포가 극단값을 보다 정밀하게 재현하는 것으로 나타났으며, EGIG 분포의 RMSE는 0.257 m로서 JSB 분포와 유사한 수준을 보였다. GGP 지점에서는 GG 분포가 RMSE 0.280 m로 가장 낮은 값을 보이며 상대적으로 우수한 재현성을 보였으나, 모든 후보 분포에서 99.9% 이상 분위수에서는 극단값을 전반적으로 과소 추정하는 경향을 보였다. UNP 지점의 경우, GB 분포가 RMSE 0.130 m로 가장 낮은 값을 보였으며, GH와 EGIG 분포에 비해 전반적인 적합도가 우수하고, 99.99% 수준에서도 극단값을 비교적 안정적으로 재현하는 결과를 보였다. KOS 지점의 경우 EGIG 분포가 유의파고의 극단값을 가장 잘 재현하는 것으로 나타났으며, 서해안에 위치한 TAA는 GH 분포가 극단값을 잘 설명하는 것으로 나타났다. 또한, Fig. 4에 제시하지 않은 지점에 대하여 Q-Q 플롯의 RMSE 분석 결과를 Appendix C에 제시하였다.

4. 최적 분포함수 평가 및 토의

후보 분포함수를 이용한 분포 적합 성능평가 측도를 분석한 결과, 복잡한 형태의 확률 분포가 더 많은 형상 매개변수를 포함함으로써 분포의 비대칭성 및 꼬리 거동까지 유연하게 조절할 수 있기 때문에 기존 단순 분포보다 유의파고 확률분포를 정확하게 재현하는 것으로 판단된다. 한편, EGIG 분포는 6개의 매개변수를 갖는 복잡하지만 유연한 구조를 갖는 분포임에도 불구하고, 1개의 지점에서 적합한 분포로 평가됐다. 이는 과도한 유연성이 관측자료의 통계적 특성을 초과해 모델이 과적합되었거나, 상대적으로 많은 매개변수로 인해 수렴 안정성 문제가 발생했을 가능성이 있다. 또한, 매개변수의 수가 증가함에 따라 매개변수 간 상관성이 높아져 추정 과정이 복잡해지고, 최적화 과정에서 수렴이 불안정하거나 비일관적인 결과가 산출될 가능성이 있으므로, 추정된 최적 매개변수에 대한 해석에 주의가 필요하며, 보다 정확한 적합도 평가가 이루어져야 할 필요가 있다(Lubke and Campbell, 2016; Ying, 2019; Kiraly and Hangya, 2022).

한편, 유의파고의 전체 자료를 대상으로 추정된 최적 확률분포를 기반으로 정확한 극값을 추정하는 것은 자료의 독립성과 등질성의 문제로 권장되지 않는다(Goda, 2010; Vanem, 2015; Takbash and Young, 2019). 그러나, 최근 해양 구조물 설계 관점에서 요구되는 등고선 접근법은 전체 자료에 대한 최적 확률분포를 추정하여 분석하는 방법으로서 이를 정확히 재현하기 위해서는 정확한 확률밀도함수를 추정하는 것이 필요하다. 이에 따라 본 연구에서는 Q-Q 플롯 방법을 통해 확률분포 꼬리 영역의 재현성을 검토하였다. Fig. 4에 제시된 바와 같이 대체로 복수의 매개변수를 가지고 있는 후보 확률분포인 GH, JSB, GB, GG, EGIG 분포들은 분위수 99.99% 이상에서 관측치와 비교적 유사한 꼬리 거동을 보이는 것으로 나타났다. 본 연구에서 검토한 확률분포의 꼬리 거동 특성은 유의파고의 최적 확률분포를 추정하기 위한 초기적인 평가이며, 극값에 대한 정확한 분석을 위해서는 극치파랑분석을 추가로 수행하는 것을 권장한다.

본 연구에서 국내 연안 20개 지점에 대하여 평가된 적합 확률분포의 특성을 확인하기 위해 평균, 표준편차, 왜도(Skewness), 첨도(Kurtosis)를 산정하여 유의파고의 형상 특성을 검토하였으며, 그 결과는 다음 Table 3에 제시하였다. 대부분의 지점에서 유의파고의 평균은 0.65~1.40 m, 표준편차는 0.42~0.98 m 범위로 나타났으며, 대체로 수심이 깊은 지점에서 상대적으로 높은 평균값과 큰 변동성을 보였다. 왜도는 모든 지점에서 1.5 이상으로 나타나 유의파고 분포가 평균보다 작은 값에 집중되고, 일부 큰 파랑이 꼬리 영역에 존재하는 우측 왜도 형태를 보였다. 분포의 비대칭 정도를 나타내는 통계량인 왜도와 분포의 뾰족함과 꼬리 영역의 두께를 나타내는 지표인 첨도를 검토한 결과, 왜도와 첨도는 각

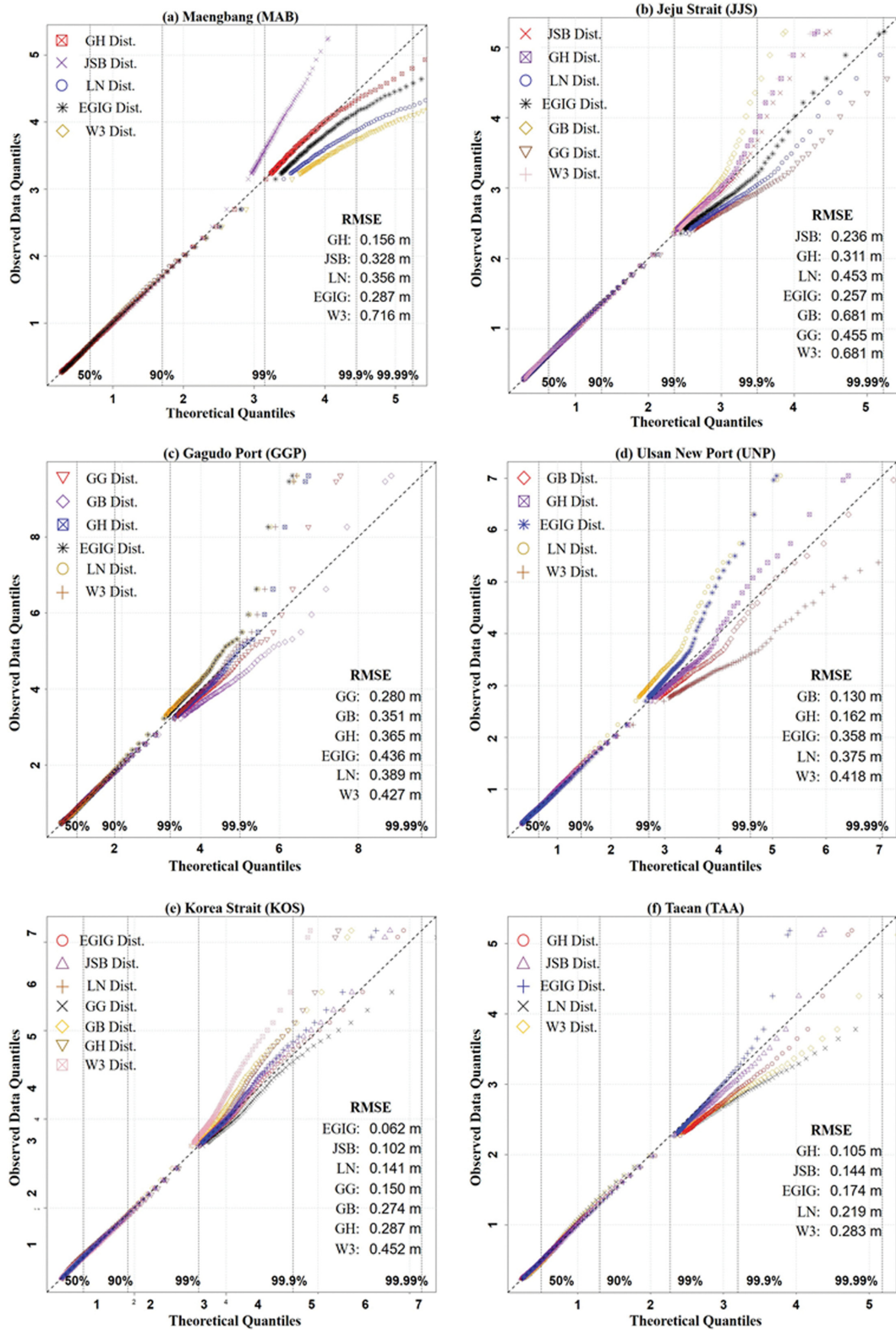


Fig. 4. Q-Q plots comparing observed and theoretical quantiles of significant wave height for representative stations using fitted probability distributions.

각 1.5~3.6, 2.8~32.2의 범위로 나타났으나, 해역 및 적합 분포에 따른 뚜렷한 경향은 확인되지 않았다. 이는 파랑 분포의 비대칭성과 침투 특성이 지역적인 해황 조건보다는 수심, 해저지형, 풍향 분포, 태풍 등 국지적 요인에 더 민감하게 반응

하기 때문인 것으로 판단된다(Annenkov and Shkira, 2014; Abroug et al., 2022). 이는 동일 해역 내에서도 해상조건의 불균질성으로 인해 확률분포의 형태가 다양하게 나타날 수 있음을 의미한다.

Table 3. Statistical characteristics of significant wave heights at 20 coastal sites around Korea

No	Station Code	Mean	SD	Skewness	Kurtosis	Optimal Distribution
1	TAA	0.6574	0.4706	1.9022	6.4058	Generalized Hyperbolic
2	SMG	0.7422	0.6533	2.2910	6.4373	Generalized Hyperbolic
3	GGP	1.0350	0.6310	2.4745	14.9010	Generalized Gamma
4	SGP	0.7407	0.5382	3.3618	22.0877	Johnson SB
5	YNP	0.7638	0.4656	3.6497	32.2423	Generalized Gamma
6	GCP	0.7343	0.4182	3.1222	23.2770	Generalized Beta 2
7	HUD	0.8125	0.4801	2.4468	14.8926	Generalized Gamma
8	UNP	0.7986	0.5205	2.5751	13.4036	Generalized Beta 2
9	GYJ	0.8189	0.5719	2.0784	7.4071	Generalized Hyperbolic
10	YIM	0.8803	0.6793	1.7608	4.1243	Johnson SB
11	HUP	0.9224	0.6110	1.7499	4.7064	Generalized Hyperbolic
12	MAB	0.8605	0.6409	1.8362	4.7188	Generalized Hyperbolic
13	DNP	0.9253	0.6323	1.7763	4.6241	Johnson SB
14	GHJ	0.8096	0.6142	2.1570	6.6695	Generalized Gamma
15	JJS	0.8329	0.5711	1.8133	4.9083	Johnson SB
16	JES	1.5026	0.9452	1.8123	6.0786	Generalized Hyperbolic
17	SOE	1.0803	0.6888	1.9696	8.9052	Generalized Hyperbolic
18	KOS	1.0468	0.6649	2.1626	9.5542	Extended G.I. Gaussian
19	UNW	1.3272	0.9073	1.5031	2.8119	Generalized Hyperbolic
20	UNE	1.4040	0.9846	1.5519	2.9803	Johnson SB

5. 결론 및 제언

유의파고의 확률분포 추정은 파랑에너지 자원량 평가, 해양 및 항만 구조물 설계 등 다양한 분야에서 활용됨에 따라 이를 정확하게 추정하고 분석하는 것은 필수적이다. 본 연구에서는 우리나라 연안을 대상으로 다양한 형태의 확률분포 모델을 적용하여 유의파고에 대한 적합 확률분포를 추정하고, 꼬리 영역의 거동 및 해역별 특성을 검토하였다.

국내에서 유의파고 확률분포 추정에 주로 사용되는 LN, W3 분포는 일부 지점을 제외하고 대부분 통계적으로 유의한 적합도를 확보하지 못했으며, 특히 분포의 중심 및 전체 영역의 재현에 한계가 있는 것으로 나타났다. 반면, 6-매개변수 기반의 EGIG 분포, 5-매개변수 기반의 GH 분포, 4-매개변수 기반의 JSB, GG, GB 등 복수의 매개변수를 가진 분포들은 높은 적합도를 보였으며, 유의파고의 중심값 및 전체 분포 특성을 보다 정밀하게 재현하는 것으로 나타났다.

그러나, 확률밀도함수만으로는 판단하기 어려운 꼬리 영역의 적합성을 정량적으로 평가하기 위해 Q-Q 플롯을 활용한 추가 분석을 수행하였다. 분석 결과, 대부분의 분포는 99% 분위수 이내에서는 관측자료와 유사한 경향을 보였으나, 분위수가 증가하여 극단값에 가까워질수록 분포 간 적합도 차이가 나타났다. 이는 매개변수가 많은 확률분포가 전반적인 분포 형태를 잘 재현하더라도, 꼬리 영역까지 일관되게 적합하기에는 한계가 있음을 의미한다. 아울러, 유의파고 확률분포 추정 시에는 하나의 분포에 의존하기보다는 다양한 매개변수를 갖는 확률분포를 충분히 고려하는 것이 필요하다고 판단되며, 확률분포 모델 선택 시 단순 매개변수 개수에 의존하는 것이 아닌 해당 분포의 구조, 안정성 등을 종합적으로 고려할 필요가 있다. 그럼에도 불구하고, 전체 유의파고 자료에

대하여 단일 확률분포를 적용해야 하는 경우에는, 본 연구에서 가장 높은 적합도를 보인 Generalized Hyperbolic(GH) 또는 Johnson SB(JSB) 분포를 사용하는 것이 가장 합리적인 선택으로 판단된다. 반면, 극치해상상태 분석이나 재현빈도 산정 등 극한 영역을 대상으로 하는 경우에는, 극치분포에 주로 활용되는 분포를 사용하는 것이 바람직할 것으로 판단된다.

본 연구에서 평가된 적합 확률분포 평가 결과는 향후 극치해상상태 분석 및 신뢰성 기반의 해양 구조물 설계, 해상작업 가능성 평가 등 다양한 실무 분야에 활용될 수 있을 것으로 기대된다. 특히, 국제표준문서에서 권장하는 등고선 접근법 분석 시, 입력변수인 유의파고에 대한 정량적 분포 정보로서 실질적인 근거를 제공할 수 있을 것으로 판단된다. 향후, 유의파고 확률분포의 특성이 관측 기간, 계절에 따라 어떤 변화를 보이는지에 대한 추가적인 검토가 필요할 것으로 보이며, 울릉도, 제주도 등 연안에 비해 상대적으로 심해에 위치한 지점에 대하여 수심 및 해저지형 등 물리적 요인을 고려한 확률분포 비교 연구가 필요할 것으로 판단되며, 확률밀도함수의 매개변수를 추정하기 위해 사용된 최대우도추정법 외에 최소제곱법, 확률가중모멘트법 및 L-Moment 법 등을 적용한 추가적인 연구가 필요할 것으로 사료된다.

감사의 글

본 연구는 산업통상자원부의 신재생에너지 핵심기술개발 사업인 “해상풍력 해상보증조사(MWS) 기술지침서 국내 표준화 개발(과제번호: RS-2025-02313959)” 과제의 일환으로 수행되었습니다. 연구비 지원에 감사드립니다.

References

- Abroug, I., Matar, R. and Abcha, N. (2022). Spatial evolution of skewness and kurtosis of unidirectional extreme waves propagating over a sloping beach. *Journal of Marine Science and Engineering*, 10(10), 1475.
- Annenkov, S.Y. and Shrira, V.I. (2014). Evaluation of skewness and kurtosis of wind waves parameterized by JONSWAP spectra. *Journal of Physical Oceanography*, 44(6), 1582-1594.
- Ardia, D., Boudt, K. and Catania, L. (2019). Generalized autoregressive score models in R: The GAS package. *Journal of Statistical Software*, 88, 1-28.
- Athanassoulis, G.A., Skarsoulis, E.K. and Belibassakis, K.A. (1994). Bivariate distributions with given marginals with an application to wave climate description. *Applied Ocean Research*, 16(1), 1-17.
- Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London A. Mathematical and Physical Sciences*, 353(1674), 401-419.
- Belov, D.I. and Armstrong, R.D. (2011). Distributions of the Kullback-Leibler divergence with applications. *British Journal of Mathematical and Statistical Psychology*, 64(2), 291-309.
- Bobee, B. and Ashkar, F. (1991). *The Gamma Family and Derived Distributions Applied in Hydrology*. Water Resources Publications, USA.
- Byrd, R.H., Lu, P., Nocedal, J. and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190-1208.
- Cho, H., Jeong, W.M. and Jun, K.C. (2013). Relationship analysis on the monitoring period and parameter estimation error of the coastal wave climate data. *Journal of Korean Society of Coastal and Ocean Engineers*, 25(1), 34-39 (in Korean).
- Clarindo, G. and Soares, C.G. (2024). Environmental contours of sea states by the I-FORM approach derived with the Burr-Log-normal statistical model. *Ocean Engineering*, 291, 116315.
- Clauset, A., Shalizi, C.R. and Newman, M.E. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661-703.
- Conover, W.J. (1999). *Practical nonparametric statistics*. John Wiley & Sons.
- Dehghan, Y., Sadrinasab, M. and Chegini, V. (2022). Probability distribution of wind speed and wave height in Nowshahr Port using the data acquired from wave scan buoy. *Ocean Engineering*, 252, 111234.
- Ferreira, J.A. and Soares, C.G. (1999). Modelling the long-term distribution of significant wave height with the Beta and Gamma models. *Ocean Engineering*, 26(8), 713-725.
- Ferreira, J.A. and Soares, C.G. (2000). Modelling distributions of significant wave height. *Coastal Engineering*, 40(4), 361-374.
- Goda, Y. (2010). *Random seas and design of maritime structures* (3rd ed). World Scientific Publishing Company, Japan.
- Gupta, R.C. and Viles, W. (2011). Roller-coaster failure rates and mean residual life functions with application to the extended generalized inverse Gaussian model. *Probability in the Engineering and Informational Sciences*, 25(1), 103-118.
- Gupta, R.C. and Viles, W. (2012). Statistical inference for the extended generalized inverse Gaussian model. *Journal of Statistical Computation and Simulation*, 82(12), 1855-1872.
- Haselsteiner, A.F. and Thoben, K.D. (2020). Predicting wave heights for marine design by prioritizing extreme events in a global model. *Renewable Energy*, 156, 1146-1157.
- James, J.P. and Panchang, V. (2022). Investigation of wave height distributions and characteristic wave periods in coastal environments. *Journal of Geophysical Research: Oceans*, 127(5), e2021JC018144.
- Jeong, W.M., Cho, H.Y. and Baek, W.D. (2015). Analysis of the long-term wave characteristics off the coast of Daejin. *Journal of Korean Society of Coastal and Ocean Engineers*, 27(2), 142-147 (in Korean).
- Johnson, N.L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, 36(1/2), 149-176.
- Katalinic, M. and Parunov, J. (2020). Uncertainties of estimating extreme significant wave height for engineering applications depending on the approach and fitting technique—Adriatic sea case study. *Journal of Marine Science and Engineering*, 8(4), 259.
- Kiraly, B. and Hangya, B. (2022). Navigating the statistical minefield of model selection and clustering in neuroscience. *Eneuro*, 9(4).
- Ko, D.H., Jeong, S.T., Cho, H.Y. and Seo, K.S. (2017). Distribution and trend analysis of the significant wave heights using KMA and ECMWF data sets in the coastal seas, Korea. *Journal of Korean Society of Coastal and Ocean Engineers*, 29(3), 129-138 (in Korean).
- Lawless, J.F. (1980). Inference in the generalized gamma and log gamma distributions. *Technometrics*, 22(3), 409-419.
- Lee, U.J., Cho, H.Y., Lee, B.W. and Ko, D.H. (2024). Joint probability distribution of significant wave height and peak wave period using gaussian copula method. *Journal of Coastal Research*, 116(SI), 96-100.
- Low, Y.C., Ong, S.H. and Gupta, R.C. (2017). Generalized Sichel distribution and associated inference. *Journal of Statistical Theory and Applications*, 16(3), 322-336.
- Lubke, G.H. and Campbell, I. (2016). Inference based on the best-fitting model can contribute to the replication crisis: Assessing model selection uncertainty using a bootstrap approach. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(4), 479-490.
- Mackay, E. and Haselsteiner, A.F. (2021). Marginal and total exceedance probabilities of environmental contours. *Marine Structures*, 75, 102863.
- Massey Jr, F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253), 68-78.
- McDonald, J.B. and Xu, Y.J. (1995). A generalization of the beta distribution with applications. *J. Econom.*, 66, 133-152.

- Nocedal, J. and Wright, S.J. (2006). Numerical optimization. Springer New York, New York.
- Pascal, F., Bombrun, L., Tourneret, J.Y. and Berthoumieu, Y. (2013). Parameter estimation for multivariate generalized Gaussian distributions. *IEEE Transactions on Signal Processing*, 61(23), 5960-5971.
- Patel, J.K. and Read, C.B. (1996). Handbook of the normal distribution. CRC Press.
- Prause, K. (1999). The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures. Ph.D. Thesis, University of Freiburg, Germany.
- Press, W.H. (2007). Numerical recipes 3rd edition: The art of scientific computing. Cambridge University Press.
- Rao, A.R. and Hamed, K.H. (2000). Flood Frequency Analysis. CRC Press LLC, New York.
- Shamshirband, S., Mosavi, A., Rabczuk, T., Nabipour, N. and Chau, K.W. (2020). Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines. *Engineering Applications of Computational Fluid Mechanics*, 14(1), 805-817.
- Shakil, M., Kibria, B.G. and Singh, J.N. (2010). A new family of distributions based on the generalized Pearson differential equation with some applications. *Austrian Journal of Statistics*, 39(3), 259-278.
- Soukissian, T.H. (2021). Probabilistic modelling of significant wave height using the extended generalized inverse Gaussian distribution. *Ocean Engineering*, 230, 109061.
- Stacy, E.W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 1187-1192.
- Takbashi, A. and Young, I.R. (2019). Global ocean extreme wave heights from spatial ensemble data. *Journal of Climate*, 32(20), 6823-6836.
- Vanem, E. (2015). Uncertainties in extreme value modelling of wave data in a climate change perspective. *Journal of Ocean Engineering and Marine Energy*, 1(4), 339-359.
- Vanem, E. and Fazeres-Ferradosa, T. (2022). A truncated, translated Weibull distribution for shallow water sea states. *Coastal Engineering*, 172, 104077.
- Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, 1168(1), 022022.
- Young, I.R., Zieger, S. and Babanin, A.V. (2011). Global trends in wind speed and wave height. *Science*, 332(6028), 451-455.
- Young, I.R. and Ribal, A. (2019). Multiplatform evaluation of global trends in wind speed and wave height. *Science*, 364(6440), 548-552.
- Weibull, W. (1951). A statistical distribution function of wide application. *Journal of Applied Mechanics*, 18, 287-293.
- Wu, Y., Randell, D., Christou, M., Ewans, K. and Jonathan, P. (2016). On the distribution of wave height in shallow water. *Coastal Engineering*, 111, 39-49.

Received 30 July, 2025

1st Revised 10 November, 2025

2nd Revised 5 December, 2025

Accepted 9 December, 2025

Appendix

Appendix A. Estimated parameters of 9 probability distribution for each observation sites

S. Code	Distributions						
	Normal		Log-normal		Gamma		
	Location	Scale	Location	Scale	Shape	Scale	Location
TAA	0.657	0.471	-0.633	0.642	1.683	0.331	0.100
SMG	0.742	0.653	-0.582	0.725	1.340	0.479	0.100
GGP	1.035	0.631	-0.112	0.528	2.123	0.389	0.210
SGP	0.741	0.538	-0.476	0.562	1.687	0.338	0.170
YNP	0.764	0.466	-0.402	0.498	2.144	0.272	0.180
GCP	0.734	0.418	-0.426	0.462	2.646	0.217	0.160
HUD	0.812	0.480	-0.348	0.520	2.340	0.283	0.150
UNP	0.799	0.520	-0.387	0.555	2.162	0.309	0.130
GYJ	0.819	0.572	-0.401	0.628	1.924	0.374	0.100
YIM	0.880	0.679	-0.387	0.721	1.519	0.514	0.100
HUP	0.922	0.611	-0.272	0.617	2.102	0.391	0.100
MAB	0.860	0.641	-0.391	0.697	2.169	0.392	0.010
DNP	0.925	0.632	-0.279	0.633	1.872	0.430	0.120
GHJ	0.810	0.614	-0.441	0.667	1.679	0.424	0.099
JES	0.833	0.571	-0.388	0.643	2.234	0.350	0.050
JJS	1.503	0.945	0.232	0.594	2.385	0.571	0.140
SOE	1.080	0.689	-0.103	0.606	2.188	0.439	0.120
KOS	1.047	0.665	-0.125	0.584	2.895	0.351	0.029
UNW	1.327	0.907	0.065	0.673	1.998	0.615	0.099
UNE	1.404	0.985	0.111	0.690	1.968	0.668	0.090

S. Code	Distributions						
	Weibull-3			Generalized Beta 2			
	Shape	Scale	Location	Shape (k)	Scale	Shape (p)	Shape (q)
TAA	1.291	0.607	0.100	1.162	0.145	10.000	2.588
SMG	1.117	0.672	0.100	1.332	0.112	10.000	1.580
GGP	1.445	0.916	0.210	1.319	0.332	9.983	3.058
SGP	1.245	0.618	0.170	2.204	0.178	10.000	1.051
YNP	1.425	0.647	0.180	1.446	0.256	10.000	2.856
GCP	1.549	0.644	0.160	1.984	0.240	10.000	1.775
HUD	1.517	0.740	0.150	1.185	0.311	10.000	4.085
UNP	1.438	0.743	0.130	1.260	0.239	10.000	3.035
GYJ	1.381	0.793	0.100	0.879	0.317	10.000	5.414
YIM	1.233	0.839	0.100	0.744	0.313	9.994	5.833
HUP	1.460	0.914	0.100	0.800	0.529	10.000	7.592
MAB	1.459	0.947	0.010	0.669	0.601	10.000	9.278
DNP	1.378	0.887	0.120	0.825	0.419	10.000	6.335
GHJ	1.280	0.771	0.100	0.906	0.231	10.000	4.248
JES	1.495	0.873	0.050	0.710	0.676	10.000	9.982
JJS	1.560	1.526	0.140	0.769	1.262	10.000	9.961
SOE	1.504	1.070	0.120	0.792	1.149	8.341	9.996
KOS	1.670	1.146	0.030	0.785	0.880	10.000	9.980
UNW	1.443	1.360	0.100	0.786	2.081	6.116	10.000
UNE	1.430	1.455	0.09	0.778	2.322	5.876	9.956

Appendix A. Continued

S. Code	Distributions							
	Johnson SB				Generalized Gamma			
	Shape (q)	Shape (p)	Location	Scale	Shape (p)	Shape (q)	Scale	Location
TAA	1.619	0.823	0.148	2.885	6.624	0.251	0.001	0.080
SMG	0.001	1.034	0.109	0.415	5.508	0.239	0.001	0.080
GGP	0.001	1.478	0.178	0.686	8.612	0.259	0.004	0.168
SGP	2.590	0.994	0.213	5.299	8.694	0.264	0.002	0.085
YNP	0.001	1.690	0.104	0.547	10.657	0.275	0.002	0.090
GCP	0.004	1.436	0.209	0.411	12.388	0.284	0.001	0.080
HUD	0.001	1.634	0.090	0.600	10.550	0.273	0.001	0.075
UNP	0.001	1.427	0.130	0.526	9.648	0.268	0.001	0.065
GYJ	2.708	1.119	0.123	6.496	8.187	0.258	0.001	0.050
YIM	1.615	0.837	0.134	4.209	6.556	0.245	0.002	0.050
HUP	2.060	1.052	0.139	4.949	8.662	0.260	0.001	0.050
MAB	1.818	0.948	0.102	4.464	7.953	0.255	0.001	0.005
DNP	1.704	0.954	0.160	4.103	8.065	0.256	0.001	0.060
GHJ	0.001	1.306	0.069	0.560	7.341	0.253	0.001	0.050
JES	1.655	0.987	0.110	3.637	8.556	0.260	0.001	0.025
JJS	2.946	1.328	0.106	11.840	9.649	0.258	0.001	0.070
SOE	2.587	1.296	0.060	7.131	7.322	0.320	0.001	0.060
KOS	0.001	1.781	-0.040	0.930	10.560	0.268	0.001	0.015
UNW	2.618	1.188	0.082	10.070	4.994	0.403	0.002	0.050
UNE	2.637	1.185	0.055	11.090	5.061	0.385	0.001	0.045

S. Code	Distributions										
	Generalized Hyperbolic					Extended Generalized I. Gaussian					
	k	p	a	b	q	k	s	p	q	a	b
TAA	0.817	0.459	0.173	0.117	0.484	0.018	0.639	−0.288	2.018	0.050	0.226
SMG	0.069	0.709	0.128	0.086	0.616	0.028	0.195	−0.647	2.025	0.028	0.487
GGP	−0.365	1.829	0.208	0.152	0.830	5.880	3.100	−0.579	0.527	0.131	1.712
SGP	−0.407	1.149	0.176	0.058	0.566	1.206	0.352	−0.861	0.999	0.143	1.369
YNP	−4.670	0.042	−0.014	0.004	0.778	22.460	11.438	−2.244	0.274	0.128	4.440
GCP	0.048	1.323	0.286	0.141	0.448	19.062	8.786	−4.274	0.308	0.141	9.015
HUD	0.088	2.032	0.139	0.104	0.677	4.150	2.780	−0.408	0.618	0.084	0.974
UNP	0.704	0.971	0.239	0.168	0.574	5.624	2.679	−0.590	0.524	0.098	1.479
GYJ	0.021	1.671	0.057	0.038	0.762	1.088	0.483	−0.160	1.090	0.040	0.969
YIM	0.892	0.399	0.165	0.129	0.715	0.166	0.064	0.073	1.820	0.061	0.741
HUP	0.876	1.192	0.14	0.122	0.784	0.242	0.544	−0.320	1.572	−0.010	0.671
MAB	0.836	0.831	0.114	0.106	0.746	1.099	0.421	−0.140	1.065	0.009	1.121
DNP	1.000	0.794	0.173	0.124	0.753	0.566	0.177	0.148	1.431	0.068	0.935
GHJ	0.107	1.234	0.092	0.083	0.718	1.009	0.205	−0.440	1.251	0.028	1.436
JES	1.279	0.540	0.161	0.138	0.679	0.821	0.364	0.405	1.160	0.049	0.713
JJS	1.094	1.526	0.155	0.130	1.347	1.329	0.932	1.165	0.892	0.107	0.805
SOE	1.448	1.024	0.163	0.146	0.918	0.745	0.565	1.561	0.934	0.119	0.328
KOS	1.430	1.002	0.200	0.185	0.854	6.055	2.576	−0.760	0.605	−0.020	2.352
UNW	1.123	1.150	0.126	0.190	1.200	0.119	0.315	−0.540	2.177	−0.210	1.246
UNE	−0.470	2.092	−0.070	0.143	1.480	0.287	0.310	−0.120	1.676	−0.110	1.275

Appendix B. Goodness-of-fit results and Kullback-Leibler divergence measures

(a) KS test p-values

S. Code	Distributions							
	LN	G3	W3	GB	JSB	GG	GH	EGIG
TAA	0.000	0.000	0.000	0.061	0.180	0.084	0.269	0.337
SMG	0.000	0.000	0.000	0.048	0.057	0.124	0.288	0.244
GGP	0.007	0.001	0.000	0.101	0.044	0.126	0.277	0.211
SGP	0.000	0.000	0.000	0.066	0.167	0.074	0.249	0.226
YNP	0.040	0.015	0.000	0.117	0.044	0.107	0.122	0.216
GCP	0.000	0.000	0.000	0.209	0.101	0.087	0.326	0.169
HUD	0.100	0.008	0.000	0.256	0.107	0.174	0.390	0.446
UNP	0.047	0.000	0.000	0.217	0.038	0.040	0.141	0.350
GYJ	0.088	0.002	0.000	0.124	0.245	0.124	0.404	0.408
YIM	0.107	0.032	0.075	0.065	0.345	0.109	0.367	0.374
HUP	0.096	0.005	0.000	0.080	0.354	0.128	0.403	0.419
MAB	0.055	0.000	0.011	0.041	0.330	0.048	0.417	0.300
DNP	0.028	0.008	0.016	0.087	0.231	0.054	0.381	0.389
GHJ	0.083	0.000	0.037	0.062	0.057	0.066	0.411	0.397
JJS	0.091	0.004	0.001	0.068	0.212	0.108	0.375	0.361
JES	0.259	0.055	0.041	0.215	0.446	0.331	0.402	0.418
SOE	0.145	0.121	0.092	0.131	0.374	0.252	0.419	0.431
KOS	0.363	0.038	0.000	0.334	0.365	0.349	0.176	0.365
UNW	0.202	0.032	0.118	0.176	0.402	0.283	0.389	0.447
UNE	0.150	0.032	0.061	0.180	0.454	0.289	0.287	0.390

(b) Kullback-Leibler divergence measures, D_{KL}

S. Code	Distributions							
	LN	G3	W3	GB	JSB	GG	GH	EGIG
TAA	41.90	85.25	125.94	51.70	26.56	33.02	25.88	36.73
SMG	52.25	139.68	178.36	46.50	29.08	39.44	20.64	31.10
GGP	36.08	86.68	151.47	29.42	58.65	24.05	25.57	28.19
SGP	86.48	216.29	183.84	50.69	30.28	89.52	32.10	38.14
YNP	32.07	131.37	140.40	36.46	27.23	24.92	35.96	29.27
GCP	66.97	174.79	234.08	27.24	27.47	65.02	32.32	47.36
HUD	28.15	94.94	143.01	28.92	23.33	24.84	29.19	24.96
UNP	37.00	113.15	172.02	25.15	41.76	36.21	35.15	29.32
GYJ	25.23	81.32	115.32	25.75	24.04	25.20	22.92	26.92
YIM	49.29	64.35	71.35	57.71	25.96	40.00	27.84	28.13
HUP	29.74	54.84	111.37	30.37	22.42	24.75	18.50	22.34
MAB	34.90	71.32	145.68	41.55	24.93	26.95	20.04	24.38
DNP	34.57	55.83	84.83	39.58	22.07	34.32	22.35	24.25
GHJ	24.87	84.92	119.14	38.00	29.81	23.77	24.20	24.01
JJS	33.41	47.00	117.12	36.98	20.06	30.70	23.17	25.19
JES	29.27	31.91	93.90	30.84	18.95	21.29	15.57	18.04
SOE	30.04	34.56	66.50	38.19	18.50	26.84	16.78	20.57
KOS	19.47	50.83	166.62	20.09	19.06	19.78	26.89	18.85
UNW	28.99	36.95	77.40	19.03	19.94	20.63	17.46	17.53
UNE	37.20	39.79	83.93	28.74	18.20	20.98	24.76	18.98

Appendix C. Quantile-Quantile RMSE analysis results

S. Code	Distributions							
	LN	G3	W3	GB	JSB	GG	GH	EGIG
TAA	0.219	2.024	0.283	0.645	0.144	0.359	0.105	0.174
SMG	2.012	2.621	0.995	0.472	0.147	0.434	0.365	0.269
GGP	0.389	1.810	0.427	0.351	0.855	0.280	0.365	0.436
SGP	0.816	1.978	0.794	1.380	0.267	0.813	0.356	0.278
YNP	0.777	1.231	0.584	0.156	0.435	0.152	0.762	0.428
GCP	0.638	0.995	0.410	0.286	0.174	0.638	0.494	0.285
HUD	0.366	1.365	0.377	0.368	0.294	0.249	0.373	0.390
UNP	0.375	0.713	0.418	0.130	0.358	1.071	0.162	0.251
GYJ	0.146	1.102	0.110	0.834	0.251	0.060	0.032	0.074
YIM	0.851	1.915	0.389	0.910	0.058	0.694	0.265	0.090
HUP	0.423	0.885	0.063	0.808	0.285	0.264	0.024	0.042
MAB	0.280	1.015	0.716	0.905	0.133	0.412	0.098	0.327
DNP	0.493	1.251	0.070	1.079	0.066	0.468	0.337	0.084
GHJ	0.343	1.655	0.235	1.004	0.803	0.110	0.141	0.106
JJS	0.311	1.408	0.681	0.453	0.236	0.455	0.257	0.356
JES	0.523	0.846	0.095	0.646	0.469	0.243	0.085	0.090
SOE	0.404	1.144	0.440	0.391	0.619	0.354	0.259	0.406
KOS	0.141	1.221	0.452	0.274	0.102	0.150	0.287	0.062
UNW	1.277	1.462	0.725	0.918	0.210	0.344	0.090	0.173
UNE	1.650	1.033	0.784	1.395	0.219	0.813	0.201	0.170