

동해 연안의 고파, 돌연 고파의 발생 빈도 분포 추정 및 특성 분석 Distribution Estimation and Characteristic Analysis of the Number of the High Waves and Sudden High Waves in the East Coastal Seas

조홍연* · 류경호** · 정예찬*** · 정원무****

Hong-Yeon Cho*, Kyong-Ho Ryu**, Ye-Chan Jeong*** and Weon-Mu Jeong****

요지 : 동해 연안의 7개 지점, 17년간의 파랑관측자료에서 추출한 고파, 돌연 고파의 월별 발생 빈도 정보를 이용하여 분포 추정, 지점 차이, 월별 차이 등 다양한 통계적인 특성 분석에 중점을 두고 수행하였다. 모든 지점에서 전반적으로 Poisson 분포가 적합한 분포로 파악되었으며, 묵호 지점은 Zero-Inflated Poisson 분포가 보다 더 적합한 분포로 파악되었다. 지점별 차이는 다른 지점과 뚜렷한 차이를 보이는 묵호와 고파 발생 빈도가 상대적으로 높은 그룹, 낮은 그룹으로 구분되고 있는 것으로 파악되었다. 반면 월별 발생 빈도는 뚜렷한 차이가 나타나고 있으나, 통계적인 신뢰구간은 상당한 크기로 추정되었다. 모든 지점을 통합하는 대 표본 자료를 이용하여 신뢰구간의 크기를 저감하여 분석한 결과, 5~8월은 평균 이하의 발생 빈도를 보였으며, 4월, 9월은 평균 수준, 나머지 기간은 평균 이상의 발생빈율을 보이는 것으로 파악되었다.

핵심용어 : 고파, 돌연 고파, 발생 빈도, Poisson 분포, Zero-Inflated Poisson 분포, 발생 비율

Abstract : Diverse statistical characteristic analyses, including distribution estimation, location differences, and monthly differences, were performed with emphasis using monthly frequency information for high waves and sudden high waves extracted from 17 years of wave observation data from seven stations along the East Sea coast. Overall, the Poisson distribution was identified as the most appropriate distribution for all locations, while the Zero-Inflated Poisson distribution was found to be a more suitable distribution for the Mukho station. Station differences were identified as distinct from other stations, with Mukho showing a clear distinction and being categorized into groups with relatively high and low high wave frequency. On the other hand, while distinct differences were observed in monthly frequency, the statistical confidence intervals were estimated to be substantial. As a result of analyzing the data by reducing the size of the confidence interval using a large sample dataset that integrates all locations, it was found that the frequency was below the average from May to August, April and September were at the average level, and the remaining periods showed an above-average frequency.

Keywords : high waves, sudden high waves, occurrence frequency, Poisson distribution, Zero-Inflated Poisson distribution, occurrence proportion

1. 서 론

연안에서 발생하는 폭풍 파랑은 대표적인 재해 요인이며, 관련 재해저감 대책 수립에 필요한 기본 정보이다. 폭풍 파랑의 발생 빈도는 지점별, 시기별로 매우 큰 변동성을 보이고 있으며, 관측자료를 이용한 빈도 추정은 기본적인 빈도 추정뿐만 아니라 불확실성 평가를 위한 구간 추정도 요구되는 상황이다. 우리나라 연안에서의 강한 바람으로 인한 폭풍파 또는 고파(high waves) 발생 연구는 지속적으로 수행되어 왔

으며, 연구 유형은 현장 관측 데이터를 이용한 빈도 분석, 수치모의, 장기 후측자료를 이용한 분석 등으로 구분된다. 기상 특보 및 현장 데이터를 이용한 연구로는 Seol(2017)이 월간 기상정보를 이용하여 2010~2014년 동안의 우리나라의 풍랑 특보 분포 특성에 대하여 해역별, 시기별 분석을 수행한 바 있다. Kang and Seol(2022)은 파고 관측자료를 이용하여 풍랑특보의 적합성을 분석한 바 있다. 또한 Jeong et al.(2016)은 동해 연안 6개 지점의 파랑 관측자료를 이용하여 고파 출현회수 및 추세 분석을 수행한 바 있다. 또한 Jeong et al.

*한국해양과학기술원 책임연구원(Principal Researcher, Korea Institute of Ocean Science and Technology)

**한국해양과학기술원 연구원(Research Scientist, Maritime ICT & Mobility Research Department, Korea Institute of Ocean Science and Technology)

*** (주)해안해양기술 사원(Engineer, Coast and Ocean Technology Research Institute)

**** (주)해안해양기술 고문(Corresponding author: Weon-Mu Jeong, Advisor, Coast and Ocean Technology Research Institute, #2407, A-dong, Hyundai Knowledge Industrial Center, 70 Dusan-ro, Geumcheon-gu, Seoul 08584, Korea, Tel: +82-2-2104-6606, wmj1960@cotri.co.kr)

(2026)은 동해 중부 연안의 돌연 고파 발생 빈도 특성에 대하여 분석을 수행한 바 있다. KOGAS(2026)에서는 관측 파고자료를 이용하여 너울성 고파 발생 여부를 판단한 장기자료(2018.11~2020.3)를 공공 데이터로 제공하고 있다. 기상청에서도 풍랑주의보, 풍랑특보를 발효하여 위험을 경고하고 있으며, 기상특보 자료로 제공하고 있다(KMA, 2026). 이처럼 동해 연안에서는 너울성 고파, 폭풍 파랑, 돌연 고파 등의 위험이 상존하는 상황에서 이에 대한 관심이 지속되어 왔다. 한편, 분포함수와 관련하여, 폭풍파랑은 아니지만 Kim and Kim(2016)은 집중호우 사상에 대한 Poisson 분포, Zero-Inflated Poisson 분포, Zero-Inflated Generalized Poisson 분포 모델의 적합에 대한 비교 평가를 수행한 바 있다. Kim et al.(2019)은 태풍에 의한 폭풍해일의 발생 빈도로 한정하여 극치해석을 Poisson-Generalized Pareto 분포함수를 이용하여 수행하고, 추세 분석 및 불확실성 분석도 수행하였다. 한편 Oh et al.(2010)는 동해안에서 빈번하게 발생하는 너울성 고파의 발생 원인 및 변화 양상을 분석한 바 있다. 반면, Kim and Jang(2020)는 40년 동안(1979~2018)의 고해상도 동해 파랑기후 후측 자료의 파고(제99백분위수), 주기 데이터를 대상으로 장기 추세를 공간적으로 분석한 바 있다. Kim et al.(2023)은 기후 변화에 의한 태풍 강도 변화, 내습 파랑의 변화 양상을 분석하고, 기존 설계파고의 제고 필요성을 제시한 바 있다. Chun et al.(2014)은 동계 폭풍파랑의 한 지점을 대상으로 통계적 특성을 분석하고, 수치모델을 이용한 재현을 시도한 바 있다. Son and Do(2022)는 태풍에 의한 폭풍 파랑 발생을 수치모델기법으로 재현하고, 폭풍 파랑 특성 분석을 수행한 바 있다.

그러나 어떤 폭풍 사상, 돌연 고파 사상 등을 확률변수로 대응하여, 발생 빈도 분포를 추정하고, 한정된 기간의 표본에 의한 추정으로부터 발생하는 불확실성에 대한 통계적인 검토 분석을 수행한 사례는 매우 제한된다.

따라서 본 연구에서는 동해 연안 7개 지점, 17년 동안의 관측자료를 이용하여 고파 발생 빈도와 예상하기 어려운 돌연 고파(sudden high waves)의 발생 빈도를 추정하고, 지점별, 시기별 변화 양상을 분석하였다. 또한 관측 자료가 가지는 표본으로서의 정량적인 한계에 대한 불확실성 평가를 위하여 구간 추정을 수행·분석하였다. 먼저 2장에서 동해 연안에서 관측된 파랑 자료에 대한 설명과 고파 및 돌연 고파의 개수 산

정(counting) 방법, 발생 빈도 자료의 분포 추정 방법을 소개한다. 3장에서는 발생 빈도에 대한 기본적인 통계정보 추정 및 발생 빈도 데이터의 독립성 검정(independence test), 분포 함수 추정 및 적합 검정, 월별 발생비율의 신뢰구간 등을 추정하고, 그 결과를 바탕으로 공간적인 차이, 계절 변화 양상에 대한 분석을 수행한다. 공간적인 차이는 두 지점의 상관계수를 이용하여 분석하였다. 마지막으로 결론 및 제언에서는 본 연구의 주요 결론과 본 연구의 한계를 극복하기 위한 제언들을 기술한다.

2. 재료 및 방법

2.1 분석에 사용한 파랑 자료

본 연구에서는 동해 연안의 7개소(대진, 속초, 묵호, 죽변, 후포, 양포, 진하)에서 장기간 수집된 파랑관측자료를 분석하였다. 이 7개소에서의 파랑 관측은 여건에 따라 빠른 곳에서는 2003년 3월, 상대적으로 늦은 곳에서는 2007년 10월에 관측이 시작되었으며 모두 2025년 6~7월까지 자료가 수집되었다. 파랑의 연별 혹은 월별 추세 등을 검토할 때는 시간의 동시성(concurrency)을 위하여 자료의 관측기간을 통일할 필요가 있다. 이를 감안하여 본 연구에서는 전 지점에서 1년 단위 규모의 관측이 수행된 17년 동안의 7개소의 자료 기간을 2008년 1월 1일부터 2024년 12월 31일로 통일하였다. 어떤 특정 지점의 특성 파악은 가능한 모든 기간의 자료 사용이 가능하지만, 두 지점 이상의 공간 차이 비교는 동일 기간의 비교가 적절하다고 판단하여 가용자료에서 공통 기간으로 시기를 제한하여 사용하였다. 한편 이 통합기간에 포함되는 결측구간의 자료는 월별 기준 결측비율을 70%로 설정하고, 결측비율이 이 기준을 상회하는 경우에는 그 자료를 이용하여 추정된 개수 자료를 이용하였다. 그러나 결측비율이 기준 이하인 경우에는 실질적으로 그 시기에 발생한 고파, 돌연 고파가 개수 추정에서 누락되었을 것으로 판단하여, 다른 연도의 같은 월 개수 자료의 중위수(median)로 보충하였다. 이 보충 자료는 본 연구의 통계분석에서 예상되는 편향(bias) 저감효과를 기대할 수 있을 것으로 판단된다.

관측 정점의 위치 및 관측 기간, 수심 정보는 Table 1에 제시하였으며, 추가로 Fig. 1에 관측 정점의 위치를 도시하여 제시하였다. 그림에서 볼 수 있는 바와 같이 파랑관측소의 위

Table 1. Information of the seven wave measurement stations

Station code	Observation period	Longitude	Latitude	Water depth (m)
Daejin	2008.01.01~2024.12.31	128°25'37.60"E	38°31'17.90"N	17.5
Sokcho	2008.01.01~2024.12.31	128°36'59.73"E	38°12'28.02"N	18.5
Mukho	2008.01.01~2024.12.31	129°07'30.89"E	37°32'52.94"N	15.0
Jukbyeon	2008.01.01~2024.12.31	129°26'00.20"E	37°03'37.20"N	18.0
Hupo	2008.01.01~2024.12.31	129°29'03.78"E	36°41'58.81"N	17.5
Yangpo	2008.01.01~2024.12.31	129°32'20.20"E	35°53'09.08"N	17.8
Jinha	2008.01.01~2024.12.31	129°21'32.36"E	35°23'13.48"N	18.0

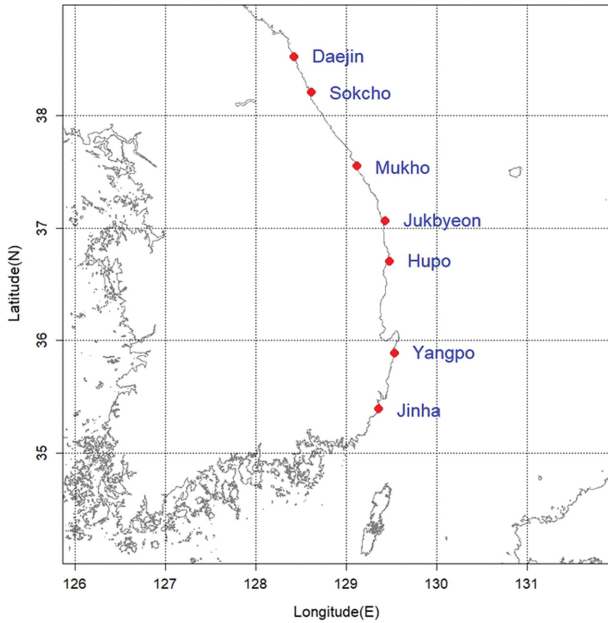


Fig. 1. Location map of the wave monitoring station in East coastal waters.

치가 동해 북단에서 울산 연안까지 동해 연안에 골고루 분포되어, 동해 연안을 공간적으로 대표할 수 있다고 판단된다.

동해 연안 7개소의 파랑 자료들은 모두 수압식 파고계(WTG; Wave and Tide Gauge)를 사용하여 취득되었으며, 스펙트럼법(spectral method)과 파별분석법을 모두 사용하여 분석하였다. 먼저 수압 스펙트럼을 구하고 수심과 주기별로 미리 구해진 수압-해수면 변위 간의 전달함수를 여기에 곱하여 해수면 변위에 대한 스펙트럼을 구하게 된다. 본 연구에서는 전체 자료에서 30분마다 2,048개의 0.5초 간격의 연속된 자료를 선택한 후 다음의 식(1)을 이용하여 유의파고를 산정하였다.

$$H_s = 4\sqrt{m_0} \quad (1)$$

여기서, $H_s (= H_{m0})$ 는 유의파고(m), m_0 는 수면 변위 스펙트럼 함수에 대한 0차 모멘트이다. 일반적으로 n 차 모멘트 m_n 에 대한 정의는 식(2)로 표현된다.

$$m_n = \int_{f_1}^{f_2} f^n \cdot S(f) df \quad (2)$$

여기서, f 는 주파수(Hertz), $S(f)$ 는 주파수 f 에 대한 수면변동에 대한 스펙트럼 밀도함수, f_1 과 f_2 는 각각 하한 및 상한 절단 주파수(cutoff frequency)이다. 일반적으로 약 95% 이상의 파 에너지가 포함되도록 상한 및 하한 절단 주파수(cutoff frequency)로 각각 $0.6f_p$, $3.0f_p$ (f_p = 첨두주파수)가 제안되고 있으나, 첨두주파수가 시간에 따라 변하는 파랑 특성으로 가변적인 적분 영역을 가지게 되는 단점이 있다. 본 연구에서는 Cho et al.(2012)의 연구 결과를 참조하여 고정 주파수 영역설정 방법을 사용하였으며 f_1 을 5/128 Hz(25.6초), f_2 를 64/

128 Hz(2초)로 설정하였다. 한편 동일한 자료에 대해 영점상향교차법을 사용하여 유의파주기($T_{1/3}$)도 구하였다. 그리고 Kim et al.(2016), Jeong et al.(2026)은 돌연 고파의 강도를 산정할 때 파고 기울기와 파랑 에너지($\tilde{E} = H^2 L$)의 기울기를 검토하였다. 파랑 에너지를 구할 때 예를 들어 쌍봉형 스펙트럼과 같은 파의 경우에는 스펙트럼법에 의한 첨두주기보다 파별분석법에 의한 유의파주기를 사용하는 것이 보다 타당하므로 본 연구에서는 고파와 돌연 고파의 분석에 유의파고(H_s)와 유의파주기($T_{1/3}$)를 사용하였다. 그리고 Fig. 2에는 7개 관측지점의 유의파고(H_s)와 유의파주기($T_{1/3}$)의 시계열을 각각 도시하였으며, 관측 간격은 30분이다. 이들 자료에서 고파, 돌연 고파를 추출하고, 개수를 산정(count)하여 분석에 이용하였다.

2.2 고파와 돌연 고파의 개수 산정(counting)

해면의 이상 상태를 표현하는 용어로는 폭풍파, 고파, 너울성 고파, 이상 고파, 돌연 고파 등과 같이 여러 가지가 사용되고 있다. 여기서는 본 연구에서 사용하는 고파와 돌연 고파에 대해서만 협의의 정의를 내리고자 한다. 기상청에서는 풍랑주의보 발효 기준으로 유의파고 3 m를 사용하고 있으며 Jeong et al.(2026)도 고파 기준으로 유의파고 3 m를 사용하였다. 어떤 지역에서 발생하는 돌연 고파의 출현 확률을 평가하려는 경우 특정 기준을 초과하는 고파, 예를 들면 유의파고 3 m를 초과하는 고파 중에서 돌연 고파가 차지하는 비율을 검토할 필요가 있다. 이 경우 고파를 규정하기 위해서는 파고뿐만 아니라 고파가 연속되는 경우 고파 유지기간의 설정 방법에 대한 검토가 필요하다.

Fig. 3에는 2021년 3월 1~10일까지 묵호항 정점에서 관측된 유의파고와 유의파주기의 시계열을 도시하였는데 이 그림에서 붉은색 원은 유의파고가 3 m 이상인 경우, 그리고 노란색으로 표시된 박스는 유의파고가 1.5 m 이상인 경우를 표시한 것이다. 이를 살펴보면 3월 1~2일의 고파는 3일을 지나면서 약화되어 4~5일에는 유의파고가 0.5 m 내외로 매우 낮아졌으므로 6일 발생한 고파는 1~2일의 고파와는 별개의 독립적인 고파 이벤트라고 쉽게 판정할 수 있다.

한편, Fig. 4에는 보다 복잡한 경우들을 제시하였다. 먼저 그림의 상부에 위치한 2018년 2월 4~7일에 걸친 고파의 경우 4일의 고파가 조금 약화되어 1.8일 정도 지속되다가 6일 저녁에 다시 강화되어 고파 요건을 만족하고 있는데 두 고파 사이의 간격(유의파고가 3 m 이상인 시각들 사이의 간격)이 1.8일 정도로 다소 길기는 하지만 그동안의 유의파고는 모두 1.7 m 이상이어서 둘은 하나의 고파 이벤트로 간주해야 할 것이다. 이런 경우는 낮은 파고의 긴 휴지기가 없는 경우로 표현되기도 한다. 가운데의 그림은 2020년 1월 24~31일의 고파 기록인데 24일, 27일, 그리고 30~31일의 3회에 걸쳐 유의파고가 3 m 이상인 고파가 연속적으로 나타났다. 이 경우 24일의 고파가 25일부터 약화되어 26일 새벽에는 유의파고

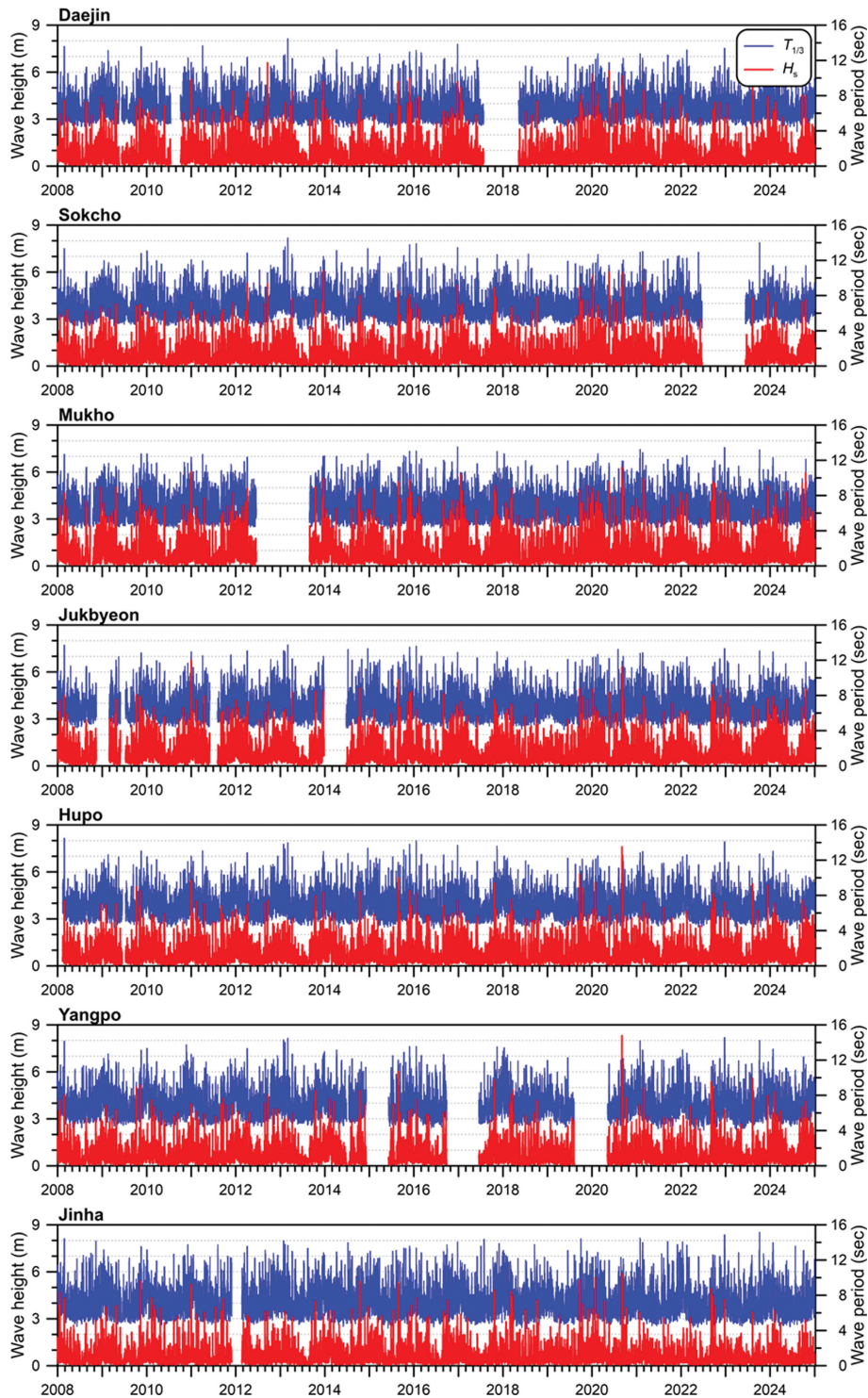


Fig. 2. Time series of wave data measured at seven wave measurement stations.

가 0.7 m 정도까지 낮아졌으며 그 이후에 파고가 증가되기 시작하여 27일 밤에 다시 유의파고가 3 m를 초과하였으므로 이들은 별개의 고파 이벤트들이라고 판단된다. 27일 밤의 고파는 다소 약화된 상태로 약 2.3일 지속되다가 30일 오전부터 유의파고가 3 m 이상인 고파가 되었다. 이 경우 두 고파 사이의 간격이 2.3일 정도로 길기는 하지만 그 동안의 유의파고는 모두 1.8 m 이상이어서 둘은 동일한 고파로 간주해

야 할 것으로 보여진다. 이 경우의 후자도 낮은 파고의 긴 휴지기가 없는 상태로 표현된다. 하부에 위치한 2022년 10월 4~7일의 고파 이벤트는 4일의 고파가 5일 오전부터 파고가 낮아지기 시작해서 1.45 m까지 낮아졌다가 다시 증가하기 시작하여 7일 오전에는 다시 유의파고가 3 m 이상이 되었는데 고파 사이의 간격은 약 1.7일이었다. 이처럼 고파가 발생하였다가 시간의 경과에 따라 완전히 약화되지 않고 다시 강화

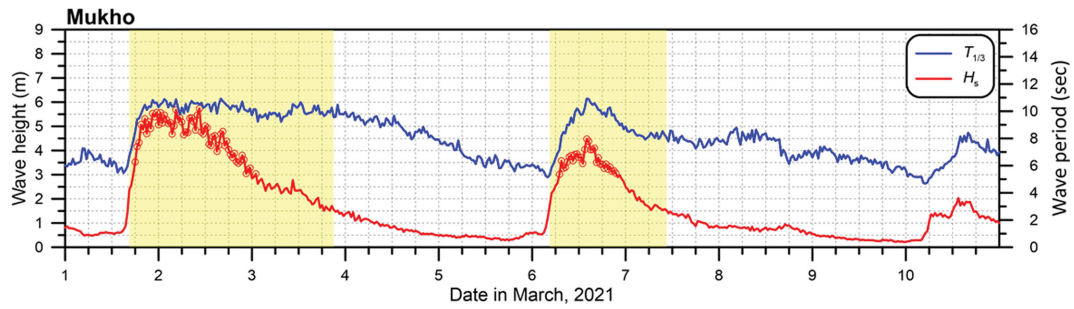


Fig. 3. Example of distinct consecutive storm wave records.

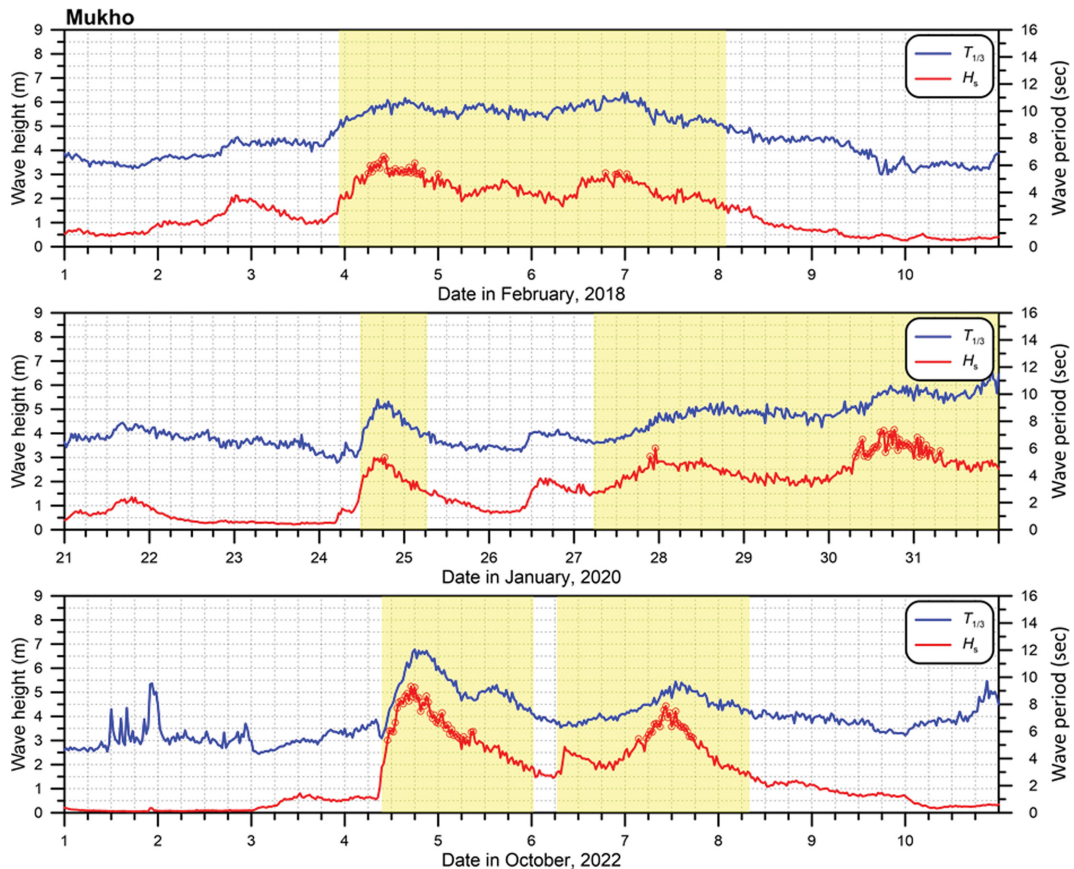


Fig. 4. Example of hard to distinguish consecutive storm wave records.

되는 경우 연결된 하나의 고파로 볼 것인지 별개의 고파로 분리해야 할 것인지에 대한 기준 설정이 필요한 것으로 생각되었다. 이 경우는 첫 번째 폭풍이 완전히 끝나기도 전에 두 번째 폭풍이 도달하여 기록 상에는 하나의 긴 폭풍처럼 보이는 경우라고 볼 수 있을 것이다. 본 연구에서는 유의파고가 3 m 이상인 연속된 고파 이벤트의 간격이 2일(48시간) 이내이고 그동안의 최소 유의파고가 1.5 m(풍랑주의보 기준인 3 m의 절반에 해당되며 Figs. 3-4에서 그 이상인 기간을 노란색 박스로 표시하였음) 이상이면 단일 고파 이벤트로 간주하는 것으로 하였다. 약 20년간의 목호항 관측 자료에 대해 고파 간격 24시간, 최소 유의파고 1.5 m로 설정한 경우 268회, 고파 간격 72시간, 최소 유의파고 1.0 m로 설정한 경우 252회로 고파의 수효가 각각 제시되어 양자의 차이가 크지는 않았다.

본 연구에서 제시한 기준[48시간, 1.5 m]을 적용한 경우 목호항에서는 고파의 수효가 260회로 제시되었으며, 다른 정점들에 대해서도 동일한 기준을 적용하여 고파 수효를 산출하였다. 목호항의 경우 적용한 기준[48시간, 1.5 m]에 대해 확장된 개념의 기준[72시간, 1.0 m]과 축소된 개념의 기준[24시간, 1.5 m]을 적용한 경우들은 모두 $\pm 3.1\%$ 의 차이를 보였다.

한편 Kim et al.(2016)은 파랑 에너지($\tilde{E} = H^2L$)의 시간변화율($\Delta\tilde{E}/\Delta t$, $t = \text{time}$)의 누적백분율을 기준으로 그 값들의 상위 20%에 해당되는 값인 $88.6 \text{ m}^3/\text{hr}$ 을 초과하는 이벤트를 돌연 고파 이벤트로 결정하였다. 최근 Jeong et al.(2026)은 돌연 고파의 시점 기준 파고를 1.25 m, 고파 기준으로 풍랑주의보 발효 기준인 3 m, 정온유지 시간(t_c)으로 3시간, 돌연출현 시간(t_s)으로 12시간을 설정하고 24시간 이내에 부여된

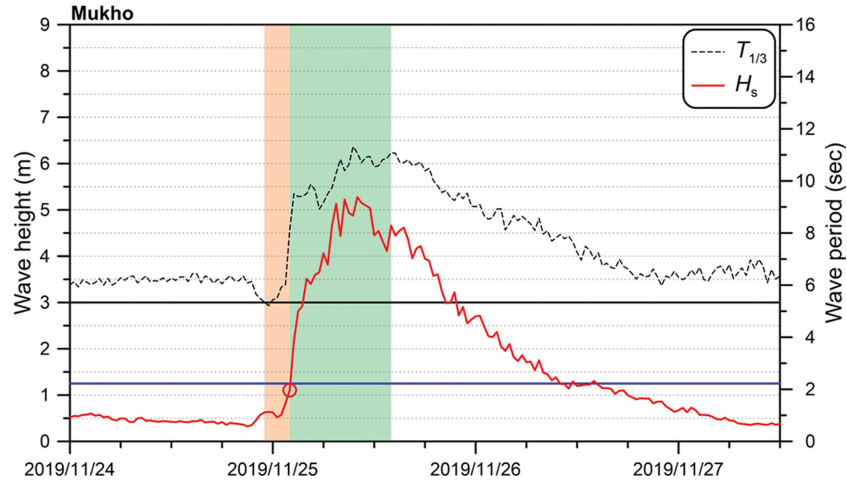


Fig. 5. Illustration of sudden high wave event.

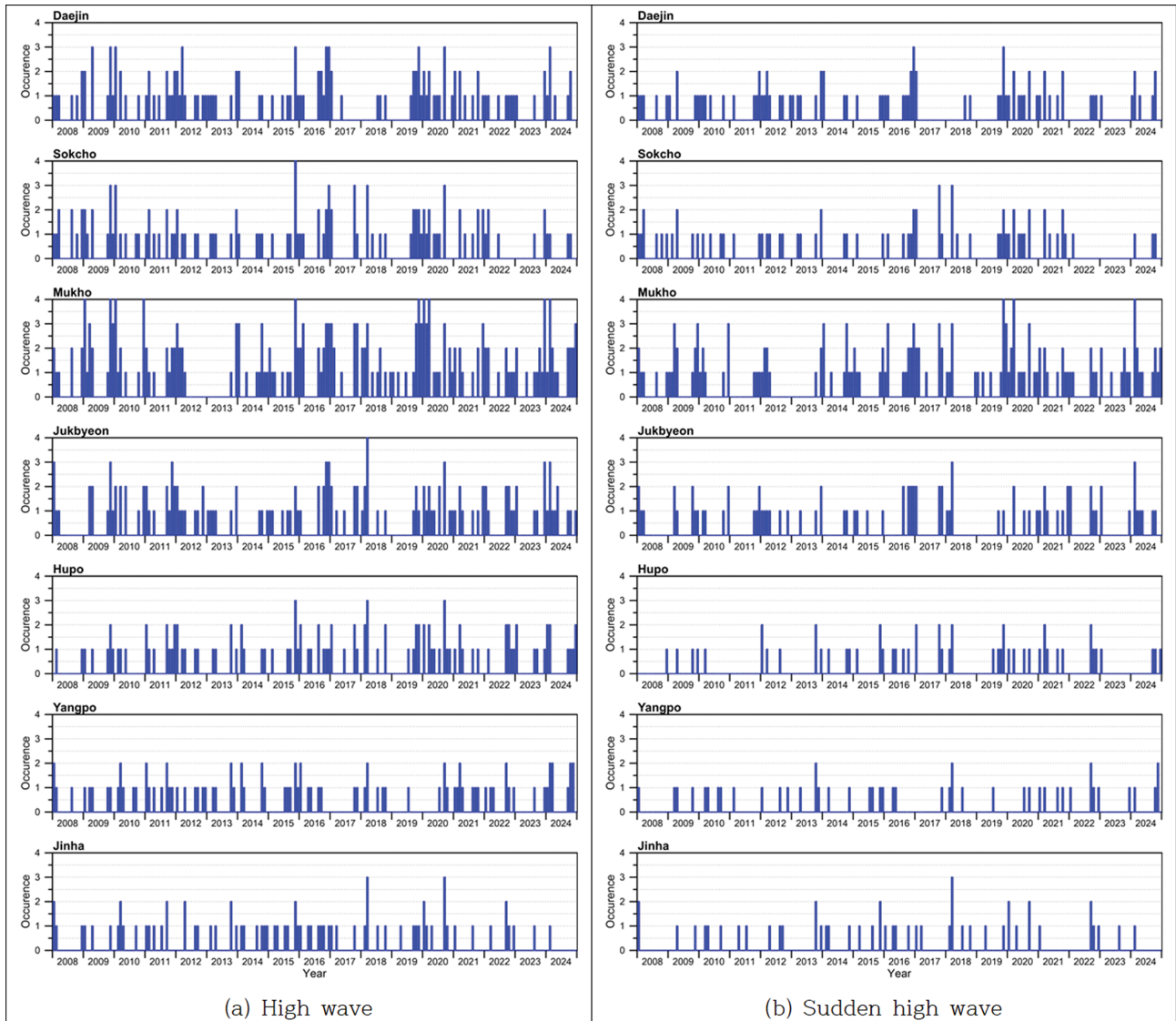


Fig. 6. Monthly occurrence numbers of the high and sudden high waves in the monitoring stations.

조건을 만족시키는 고파 이벤트를 돌연 고파로 결정하는 방법을 제안하였다. 두 방법은 고파 기준(전자는 최대 유의파

고, 후자는 3 m), 정온유지 시간의 길이(전자는 0.5시간, 후자는 3시간) 및 돌연 고파 기준(전자는 누적백분율의 상위

20% 이상, 후자는 t_c 와 t_s 를 만족하면서 3 m를 초과할 것)에서 구분되었다. Fig. 5에는 본 연구에서 적용한 돌연 고과의 선정에 사용되는 시간과 사고들을 표시하였으며 돌연 고과를 선정하는 방법은 다음과 같다. 사고 시계열을 대상으로 특정 시각에서 유의사고가 1.25 m(그림에서 청색 실선으로 표시) 이하인 경우 돌연출현 시간($t_s = 12$ 시간, 그림에서 연두색 박스로 표시) 이내에 유의사고가 3 m(그림에서 흑색 실선으로 표시) 이상인 경우를 찾으며 없는 경우에는 나타날 때까지 계속 자료를 탐색한다. 이 조건을 만족하는 시각이 나타나면 다시 정온유지 시간($t_c = 3$ 시간, 그림에서 적색 원의 시각부터 이전의 3시간을 나타내며 주황색 박스로 표시) 동안 유의사고가 1.25 m 이하를 만족하는지 탐색하고 양자를 만족하는 경우 그 시각은 돌연 고과 이벤트의 시점에 포함된다.

Fig. 6의 막대 그래프는 7개소에서 출현한 고과와 돌연 고과 이벤트들을 월별로 정리한 것이다. 그림에서 볼 수 있는 바와 같이 목호 지점이 뚜렷하게 발생 빈도가 크게 나타나고 있다. 목호 지점을 제외하면 대진, 속초, 죽변이 다소 높은 빈도로 발생하고 있으며, 후포, 양포, 진하는 상대적으로 낮은 빈도로 발생하고 있음을 알 수 있다. 그러나 보다 명확한 발생 빈도는 분포 추정을 통한 정량적인 평가 및 분포적합 검증 과정이 요구된다. 그리고 돌연 고과는 변수의 관점에서 고과의 부분집합에 해당하기 때문에 그 빈도는 전체 발생 빈도에 대한 비율 평가로 수행할 수도 있으나, 본 연구에서는 발생 빈도가 매우 작기 때문에 절대적인 발생 빈도로 분석하였다. 발생 빈도는 불연속적인 변수에 해당하기 때문에 Fig. 6에 도시된 막대 그래프로는 변화양상 파악에 제약이 따른다. 본 연구에서는 가용한 모든 기간 발생 빈도 자료를 월별로 통합하여 시간적인 변화 양상을 분석하였다.

2.3 발생 빈도 자료의 분포 추정 – Poisson 분포 적합

고과, 돌연 고과의 발생 빈도는 Poisson 분포함수를 이용하여 적합하였다. 이 분포함수는 어떤 사건이 한정된 시간, 공간영역 조건에서 발생하는 빈도를 나타내는 분포함수로 널리 이용되고 있으며, 확률밀도함수는 다음과 같다. 여기서, k 는 사건의 발생 회수이고, 매개변수는 λ 이다(Weisstein, 2026; NIST, 2012; Heckert and Filliben, 2003).

$$f_X(k; \lambda) = P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, 2, \dots, \lambda > 0 \quad (3)$$

Poisson 분포함수의 매개변수는 우도(likelihood) 함수를 이용하는 최우도법을 이용하여 간단하게 유도된다. 우도함수는 다음과 같이 정의되고, 로그 우도함수를 이용한다.

$$L(\lambda) = \prod_{i=1}^n f_X(k_i; \lambda) = \prod_{i=1}^n \frac{\lambda^{k_i} \cdot e^{-\lambda}}{(k_i)!}$$

$$\ln L(\lambda) = \sum_{i=1}^n (k_i \cdot \ln \lambda - \lambda - \ln((k_i)!)) \quad (4)$$

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = \sum_{i=1}^n \left(\frac{k_i}{\lambda} - 1 \right) = \frac{\sum_{i=1}^n k_i}{\lambda} - n = 0 \rightarrow \hat{\lambda}_{opt} = \frac{1}{n} \sum_{i=1}^n k_i$$

고과, 돌연 고과 발생 빈도 자료는 연별, 월별 자료를 모두 이용하였으며, 지점에 따른 변화 양상(발생 빈도 등) 차이분석에 이용하였다. 한편 다른 지점에 비하여 절대적인 월별 발생 빈도가 평균 1회 이상이고, Poisson 분포함수와 작은 발생 빈도(발생 회수 0, 1회)에서 상대적으로 큰 시각적인 차이를 보이는 목호 지점에 대해서는 Zero-Inflated Poisson 분포함수를 이용하여 분포 적합을 추가로 수행하였다. Zero-Inflated Poisson 분포함수는 두 개의 매개변수(z_0, λ)로 구성되며, 다음과 같이 표현된다. 매개변수 추정은 Dencks et al.(2020)가 제시한 “closed form” 공식을 이용하였다.

$$f_X(k=0; \lambda) = P(X=0) = z_0 + (1-z_0) \cdot e^{-\lambda}, k=0$$

$$f_X(k; \lambda) = P(X=k) = (1-z_0) \cdot \frac{\lambda^k e^{-\lambda}}{k!}, k=1, 2, \dots \quad (5)$$

$$\hat{\lambda} = W_0(-se^{-s}) + s$$

또는 $g(\lambda) = \bar{k}(1 - e^{-\lambda}) - \lambda(1 - (n_0/n)) = 0$ 로 제시되는 비선형 방정식을 반복 계산하여 매개변수를 추정할 수도 있으며, 초기 가정(initial guess)은 다음 $\lambda^0 = \bar{k}$ 조건을 활용할 수 있다. 나머지 하나의 매개변수(z_0)는 다음 식을 이용하여 간단하게 계산할 수 있다.

$$z_0 = 1 - \frac{\bar{k}}{\hat{\lambda}}, 0 \leq z_0 \leq 1 \quad (6)$$

여기서, $s = \bar{k} / (1 - r_0) = n\bar{k} / (n - n_0)$, $r_0 = n_0 / n$ = 전체 자료에서 zero-빈도 자료가 차지하는 비율, $\bar{k} = (1/n) \sum k_i$, n = 전체 자료의 개수, n_0 = 전체 자료에서 발생 빈도 = 0 조건에 해당하는 자료 개수, W_0 = Lambert W 함수의 “principal branch”(Corless et al., 1996)이며, R 프로그램 {pracma}에서 제공하는 lambertWp() 함수를 이용하였다(Borchers, 2023). Kim and Kim(2016)은 돌발강우 사상에 대하여 모멘트법에 의한 ZIP 매개변수 추정공식을 제안한 바 있으며 논문에서 기술한 바와 같이 식(7)과 같이 간단하게 유도된다.

$$z_0 = \frac{s_k^2 - \bar{k}}{s_k^2 + (\bar{k})^2 - \bar{k}}, \hat{\lambda} = \frac{s_k^2 + (\bar{k})^2 - \bar{k}}{\bar{k}} \quad (7)$$

분포 적합 검정은 주로 불연속적인(discrete) 변수의 적합 검정에 널리 이용되는 pmf(probability mass function) 기반의 χ^2 -검정기법을 이용하여 수행하였으며, 보조적으로 연속적인 변수에 이용되는 CDF(Cumulative Distribution Function) 기반의 Cramer-von Mises, Anderson-Darling 검정 기법을 이용하여 수행하였다(R {gofest} 패키지에서 제공하는 cvm.test(), ad.test() 함수 활용). 이 검정 방법은 KS(Kolmogorov-Smirnov) 검정 방법을 개선한 방법이다(NIST, 2012; Faraway et al.,

2021; Arnold and Emerson, 2011). 어떤 통계적인 검정 기법을 적용하는 경우, p-value 기반의 해석은 제1종 오류만을 기준으로 판정하기 때문에, 실질적인 또는 시각적인 효과 크기(effect size)를 고려하여 검정 결과를 해석하여야 한다. 자료의 개수와 유의미한 효과 크기에 따라 제2종 오류도 반영하여 통계적인 검정 결과를 해석하여야 한다(Cho and Lee, 2023). 그러나 어떤 통계량에 대한 실질적인 효과 크기는 이론적인 근거보다는 경험적인 크기 또는 목표에 따라 허용 가능한 정도로 결정되기 때문에 적절한 수준의 정량적인 설정이 요구된다.

한편 계절에 따른 발생 빈도 양상 변화는 연간 발생 빈도에 대한 월별 발생 빈도의 비율로 파악하였으며, 비율에 대한 신뢰구간 추정 결과를 이용하여 각각의 지점에서 월별 차이에 대한 검정을 수행하였다. 발생비율에 대한 신뢰구간 추정은 R 프로그램 prop.test() 함수를 이용하였으며, 신뢰구간은 95% 조건에서 상한-하한 비율을 추정하였다.

3. 결과 및 토의

3.1 고파 발생 빈도 통계정보 및 독립, 분포적합 검정

고파, 돌연 고파 발생 빈도 자료는 통계적인 변수이기 때문에 엄밀한 의미에서 독립 조건을 만족하여야 한다. 따라서, 고파 발생 빈도 자료를 기준으로 모든 지점에서 연속되는 자료의 독립검정(independence test)에 널리 이용되는 ‘runs’ 검정과 Ljung-Box 검정(Ljung and Box, 1978) 기법을 선택하였으며, R 프로그램에서 기본으로 제공하는 Box.test() 함수(option, type=“Ljung-Box”), {tseries} 패키지(Trapletti and Hornik, 2023)의 runs.test() 함수를 이용하여 5% 유의수준($\alpha = 0.05$) 조건에서 수행하였다.

독립 검정 수행 결과, 각각의 기법에 대한 p-value 정보는 다음과 같다(Table 2 참조). 여기서 귀무가설(H_0)은 “이 데이터는 독립 조건을 만족한다”이다. 따라서, $p > 0.05$ 조건에서는 귀무가설을 기각할 수 없다. 따라서, Table 2에서 보면, 연간 고파와 돌연 고파 발생 개수 자료를 이용하는 경우, 모든 정점에서 $p > 0.05$ 조건을 만족하므로, 연안 발생 개수 자료는 독립 조건을 만족한다는 귀무가설을 기각할 수 없다. 반

면, 월별 고파와 돌연 고파 발생 개수 자료를 이용하는 경우에는 돌연 고파의 경우, 속초, 진하 지점, 고파의 경우, 속초, 후포, 양포, 진하 지점을 제외하고는 $p < 0.05$ 조건에 해당하기 때문에 전반적으로 월별 발생 개수 자료의 경우, 독립 조건을 만족한다는 귀무가설이 기각된다. $p > 0.05$ 조건의 경우, 통계적인 측면에서 귀무가설을 기각할 수 없으나, 전반적으로 그리고 실질적으로 뚜렷하게 나타나는 계절 특성이 고파, 돌연 고파 등의 파랑 개수 변수 범위 등을 고려하면 통계적인 검정 결과와는 별도로 경험적인 판단에 기반을 두는 경우, 월별 자료는 독립 조건 가정에 한계가 있다고 볼 수 있다. 따라서, 월별 자료는 고파, 돌연 고파의 월별 발생 개수의 분포 비교와 분포 적합만을 위한 목적으로 한정하여 분석한다. 따라서 이를 기반으로 하는 어떤 통계적인 추론 등에 활용하는 경우에는 독립성 가정을 만족하지 못하는 자료임을 염두에 두어야 한다.

동해 연안의 파랑 관측 자료에서, 고파, 돌연 고파 발생 빈도 정보를 추정한 결과, 연간 고파 발생 빈도는 4.5~13.42, 돌연 고파 발생 빈도는 2.7~8.8개 정도이다. 동해 연안을 포함하는 동해 중부, 동해 남부, 남해 동부 해역의 평균 풍랑 특보 발효일수는 연간 66~78일 정도(KMA, 2026)로 보고되고 있다. 발효일수는 동해 모든 해역을 포함하고, 일수로 계산되기 때문에 본 연구에서와 같이 특정 관측지점의 빈도에 비하여 매우 크게 나타나고 있음을 알 수 있다. 풍랑특보는 예방의 측면에서 발효되기 때문에 관측 자료보다는 예보 자료를 이용하기 때문에 보다 불확실성이 큰 자료로 판단된다. 더불어 발효일수는 빈도와는 다르기 때문에 발효 일수(평균 72일 기준) 대비 고파, 돌연 고파의 발생비율을 추정하면, 각각 6~19%, 4~12% 정도로 추정된다.

고파와 돌연 고파의 발생 빈도 자료의 기본 정보와 Poisson 분포 적합 검정 결과는 Fig. 7, Table 3에 제시한 바와 같다. 우선 연간 발생 개수 자료를 이용하는 경우에 대한 분포 적합의 시각적인 평가를 위한 도시자료는 다음과 같다. 그리고 분포 적합 검정을 위한 p-value 산정 결과는 다음과 같다. 우선 전체 자료의 개수는 연간 자료는 17개, 월별 자료는 204(= 17 × 12)개가 된다. 연간 자료에 대한 Poisson 분포 적합에 대한 Cramer-von Mises 검정, Anderson-Darling 검

Table 2. The p-values of the independence tests

Stations	Annual count data				Monthly count data			
	Sudden high wave		High wave		Sudden high wave		High wave	
	Box test	Runs test	Box test	Runs test	Box test	Runs test	Box test	Runs test
Daejin	0.797	0.790	0.752	0.330	0.008	0.005	0.000	0.000
Sokcho	0.097	0.460	0.118	0.208	0.132	0.028	0.454	0.240
Mukho	0.329	0.495	0.727	0.790	0.000	0.018	0.000	0.000
Jukbyeon	0.364	0.528	0.821	0.692	0.003	0.003	0.007	0.002
Hupo	0.428	0.971	0.241	0.214	0.032	0.014	0.470	0.520
Yangpo	0.433	0.692	0.108	0.203	0.002	0.006	0.210	0.425
Jinha	0.876	0.442	0.127	0.565	0.389	0.367	0.621	0.549

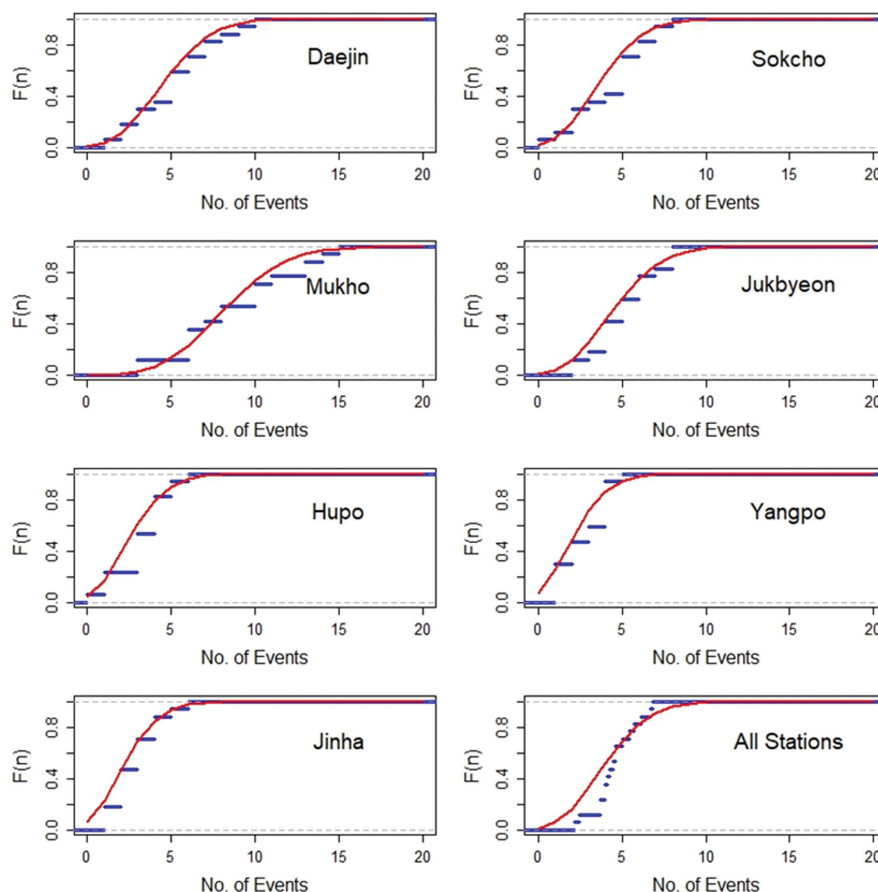


Fig. 7. CDF plots of the annual counts of the sudden high wave events.

Table 3. Basic statistical information and p-values of the goodness-of-fit test

(a) High waves								
St.	n	n_E	$\bar{k}$	$\bar{k}_L$	$\bar{k}_U$	s_k^2	D. index	p_{χ^2}
Daejin	204	133	0.652	0.535	0.769	0.730	0.892	1.000
Sokcho	204	119	0.583	0.469	0.698	0.697	0.836	0.937
Mukho	204	228	1.118	0.958	1.277	1.346	0.831	0.367
Jukbyeon	204	143	0.701	0.578	0.824	0.802	0.874	0.327
Hupo	204	111	0.544	0.440	0.648	0.574	0.947	0.556
Yangpo	204	93	0.456	0.367	0.544	0.417	1.094	1.000
Jinha	204	77	0.377	0.294	0.461	0.374	1.009	0.870
(b) Sudden high waves								
St.	n	n_E	$\bar{k}$	$\bar{k}_L$	$\bar{k}_U$	s_k^2	D. index	p_{χ^2}
Daejin	204	88	0.431	0.341	0.522	0.434	0.995	0.531
Sokcho	204	73	0.358	0.272	0.443	0.389	0.921	0.931
Mukho	204	149	0.730	0.598	0.863	0.937	0.780	0.335
Jukbyeon	204	87	0.426	0.330	0.523	0.492	0.867	0.889
Hupo	204	54	0.265	0.192	0.338	0.284	0.931	0.619
Yangpo	204	46	0.225	0.162	0.289	0.215	1.049	0.807
Jinha	204	48	0.235	0.164	0.307	0.269	0.873	1.000

Ref. n = total number of the monthly event count data, n_E = monthly numbers of the high or sudden high waves count data during 17 years, $\bar{k}$ = mean of the number of events, $\bar{k}_L$, $\bar{k}_U$ = the lower and upper limits of the mean, s_k^2 = variance of the number of the events, and p_{χ^2} = the p-values of the chi-squared goodness of fit test.

정, χ^2 검정 기법을 이용하여 산정된 p-value 범위는 각각 0.15~0.58, 0.17~0.59, 1.0 수준으로 모두 $p > 0.05$ 조건을 만족하기 때문에 “이 데이터는 Poisson 분포를 따른다”는 귀

무가설을 기각할 수 없다고 판정이 내려진다. Fig. 7에서도 볼 수 있는 바와 같이, 시각적으로 큰 차이를 보이는 구간이 다 수 보여짐에도 불구하고, 통계적인 판정은 기각할 수 없다는

판정을 내리고 있으며, 이는 소 표본에 해당하는 적은 개수 ($n = 17$)의 표본 영향으로 판단된다. “귀무가설을 기각할 만한 뚜렷한 차이로 판정하기에는 데이터가 충분하지 않다” 정도로 해석하는 것이 적절하다. 결론적으로 독립 조건을 만족하는 연간 발생 빈도 데이터는 모든 검정기법에서 “Poisson 분포를 따른다”는 귀무가설을 기각할 수 없다고 판정한다.

반면 월별 발생 개수 자료를 사용하는 경우, Table 3에서 볼 수 있는 바와 같이, chi-squared 검정기법을 적용한 경우 $p > 0.05$ 조건을 만족하기 때문에, Poisson 분포를 따른다는 귀무가설은 기각할 수 없다. 반면 Cramer-von Mises 검정, Anderson-Darling 검정 기법을 적용한 경우, 목호의 고파조건을 제외하고 모두 $p < 0.05$ 조건을 만족하는 결과를 보여 Poisson 분포를 따른다는 귀무가설은 기각된다. 그러나, 시각적인 분포 적합 그림을 보면 매우 높은 수준의 일치치를 보이고 있음을 알 수 있다(Figs. 8-9 참조). 이 경우, 통계적인 검

정 결과는 도식적인 검정 결과와는 상반되는 판정을 내리고 있음을 알 수 있다. 통계적인 검정의 한계도 있지만, 이 경우의 해석은 많은 자료를 사용하는 경우(자료 개수 204개) 어떤 통계적인 검정은 안정된 결과를 보이지만, 실질적으로는 아주 작은 차이, 그러나 실질적으로는 과도한 작은 차이를 판별해 내는 검정 능력 증가 영향으로 간주된다. 이 경우 본 연구에서 사용한 Cramer-von Mises, Anderson-Darling 검정기법은 꼬리 영역에 민감한 특성, 연속변수 대상으로 개발된 검정기법, 그리고 자료의 개수가 증가하는 경우 작은 차이도 판별하는 능력은 우수하지만, 실질적인 크기 차이가 아닌 통계적으로 유의미한 작은 차이를 판정하기 때문에 해석에 주의가 요구된다(Cho and Lee, 2023). Poisson 분포를 따르는 가상의 난수를 다수 발생하여 Cramer-von Mises, Anderson-Darling 검정을 수행하는 경우라 할지라도, 자료의 개수가 100개 이상으로 증가하는 경우에는 Poisson 분포 적합 검정

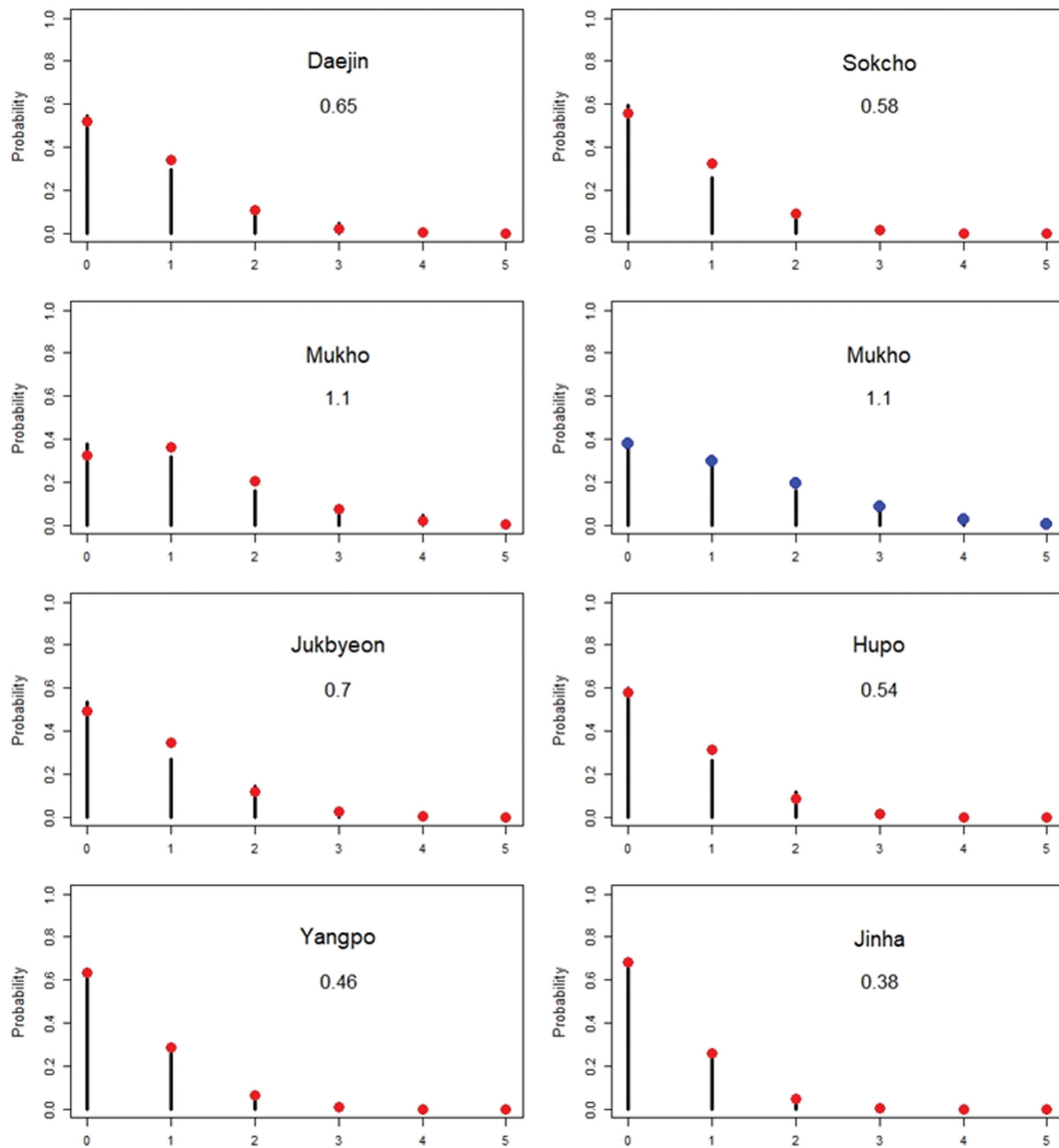


Fig. 8. Poisson distribution fitting of the high wave occurrence numbers (vertical bar = observed, solid circle = estimated (fitted) occurrence numbers) (added zero-inflated distribution fitting occurrence in Mukho station).

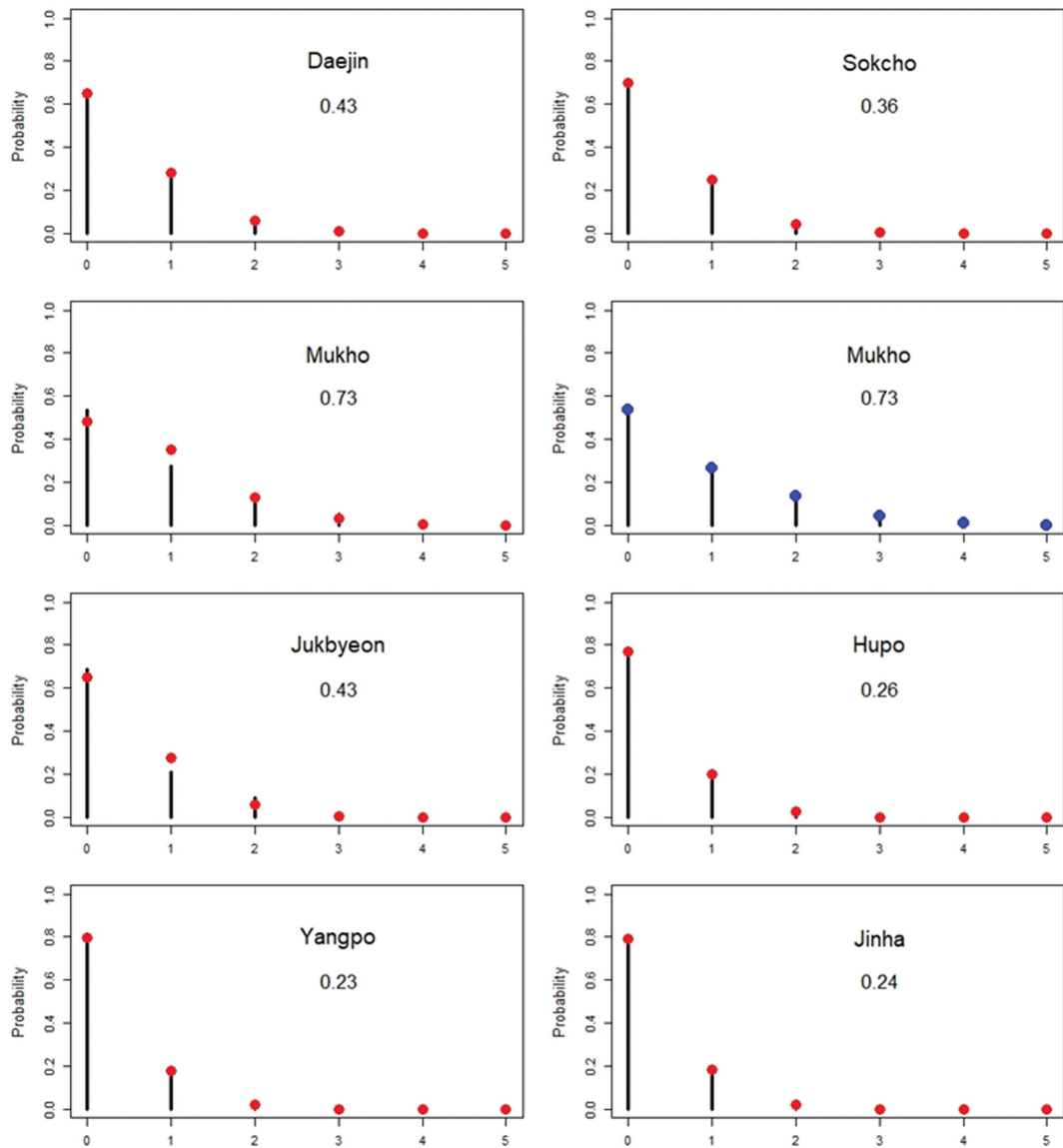


Fig. 9. Poisson distribution fitting of the sudden high wave occurrence numbers (vertical bar = observed, solid circle = estimated (fitted) occurrence numbers) (added zero-inflated distribution fitting occurrence in Mukho station).

이 기각되는 $p < 0.05$ 조건에 해당하는 결과를 제시하는 것으로 파악되었다. 따라서, 본 연구에서는 불연속 자료의 분포 적합 검정에 널리 이용되는 χ^2 검정기법을 이용하여 분포적합 검정을 수행한 결과와 비교하였다. 이 경우, $p > 0.05$ 조건으로 귀무가설을 기각할 수 없는 상반된 판정을 내리고 있음을 알 수 있다. 상반된 판정은 실질적으로 적절한 검정기법의 선택 문제가 대두되지만, 절대적인 기준이 없기 때문에 경험적인 판단에 의존하게 된다.

따라서 고파와 돌연 고파 개수의 분포 적합은 시각적인 결과와 χ^2 분포검정 결과를 이용하여 판정함이 적절하다고 판단된다. Poisson 분포 적합은 모든 지점에서 채택이 되었으나, 묵호 지점에서는 Poisson 분포의 특성에 해당하는 Poisson 분포가 대표적인 특성에 해당하는 평균 = 분산 일치 조건에서 벗어나고 있음을 알 수 있다. 그래서 묵호 지점에 대해서는 Zero-Inflated Poisson 분포 적합을 추가로 수행하

여 분석 평가에 활용하였다. 이 과정은 보다 더 적합한 분포를 비교평가하여 최적의 분포를 선택하는 과정이 아니라, Zero-Inflated Poisson 분포를 이용한 적합 수준의 개선 정도를 살펴보는 수준으로 한정한다.

Poisson 분포함수는 이론적으로 평균 = 분산 조건을 만족하여야 한다. 그러나 실제 표본 데이터에서는 두 값이 차이를 보이는 경우가 빈번하다. 그 차이는 다음과 같이 정의되는 분산 지수(dispersion index, D)를 이용하여 평가한다. 분산 지수와 유사한 변동계수는 무차원 변수이지만, 분산계수는 자료의 단위와 동일하다.

- Dispersion index(Index of dispersion), $D = \frac{s_k^2}{k}$

- Coefficient of Variation, $CV = \frac{s_k}{k}$

$D > 1.0$ 인 경우는 과분산 또는 고분산(over-dispersed),

Table 4. Parameters of the Zero-Inflated Poisson distribution at Mukho station

Waves		n	n_0	$\bar{k}$	σ_k^2	z_0	λ
High waves	MLE	204	110	1.118	1.346	0.148	1.312
	MOM	204	110	1.118	1.346	0.154	1.322
Sudden high waves	MLE	204	77	0.730	0.937	0.273	1.005
	MOM	204	77	0.730	0.937	0.279	1.013

Ref. MLE = maximum likelihood method, MOM = method of moments

$D < 1.0$ 인 경우는 저분산(under-dispersed) 조건으로 간주한다. 그러나 통계적인 관점에서, 발생 빈도 데이터로 추정한 평균의 신뢰구간을 고려하여 $D \approx 0$ (동분산, equi-dispersed) 범위 조건을 고려하면 묵호 지점을 제외하고는 모두 동분산 조건으로 간주할 수 있는 것으로 파악되었다. 평균의 90% 신뢰구간에 대한 하한, 상한은 정규분포를 이용하여 추정하였다(Cho and Lee, 2023). 분산이 이 구간에 포함되는 경우에는 동분산 조건으로 간주하였다(Table 3, k_L , k_U 컬럼 범위 참조). 동분산 조건은 수식으로 표현하면 $D = 1$, $\bar{k} = s_k^2$ 조건이지만, 통계적인 관점에서는 표본이 가지는 불확실성을 가지는 수식으로 표현하여야 한다. 이 수식을 평균의 구간추정 개념을 이용하여 표현하면 다음과 같이 표현할 수 있으며, 이 수식을 이용하여 동분산, 고분산, 저분산 조건의 판정에 이용하였다.

$$k_L \leq \mu_k \leq k_U, \quad k_L \leq \bar{k} = s_k^2 \leq k_U \quad - \text{동분산 조건}$$

$$k_U \leq s_k^2 \quad - \text{고분산 조건}$$

$$s_k^2 \leq k_L \quad - \text{저분산 조건}$$

따라서, 동분산 조건에서 벗어나는 묵호 지점에 대해서는 Zero-Inflated Poisson 분포를 이용하여 매개변수를 추정하였으며, 기존의 Poisson 분포 적합에 대한 시각적인 정도가 보다 개선된 적합 결과를 보이고 있는 것으로 파악되었다. Zero-Inflated Poisson 분포에 대한 매개변수 추정 결과는 다음과 같다. Table 4에서 볼 수 있는 바와 같이, 간단한 Moment 기법을 이용한 추정 결과와 최우도법을 이용한 결과가 큰 차이를 보이고 있지 않는 것으로 파악되었다.

3.2 Poisson, Zero-Inflated Poisson 분포 적합

Figs. 8-9는 동해 연안 7개 지점의 고파와 돌연 고파의 월별 발생 빈도와 Poisson 분포 적합을 수행한 결과이다. 그림에서 볼 수 있는 바와 같이, 그리고 χ^2 분포 적합 검정 결과에서 판단할 수 있는 바와 같이 Poisson 분포가 적절하게 발생 빈도 분포를 추정하고 있는 것으로 파악되었다. 반면 묵호 지점은 통계적인 기준에 따른 분포 적합은 통과하지만, 돌연 고파 또는 동분산 조건을 벗어나는 것으로 파악되어 Zero-Inflated Poisson 분포 적합을 추가로 수행하였다. 그 결과 Zero-빈도 영역의 적합 수준이 크게 개선되고 있음을 알 수 있다. 그리고 고파, 돌연 고파의 경우 Zero-발생 빈도에 대한 기여 비율은 각각 15%, 27% 정도로 파악되었다. 돌연 고파의 경우, Zero-발생 빈도 영향이 묵호 지점의 경우에 2배 정도 나타나고 있음을 알 수 있다.

엄밀한 의미에서 보면, 독립 조건을 만족하는 연간 발생 빈도 자료를 사용하는 경우, 전체 가용한 자료 개수가 한 지점에서 17개로 소표본에 해당하고, 각각의 지점에 대하여 Poisson 분포 적합 검정을 수행하였으며, 모든 지점이 통과하였다. 그러나 표본 개수가 작아서, 제2종 오류가 증가할 수 있으며, Power 감소가 예상된다. 따라서 연간(annual) 시간 규모를 월간(monthly) 단위 규모로 축소하여 분석을 수행하였다. 월별 발생 빈도 자료를 이용하는 경우, 자료의 독립 조건을 만족하지 못하지만, 최소한의 빈도 자료를 확보하기 위하여 발생 빈도에 대한 분포 적합을 수행하였다. 그러나 보다 장기간의 자료를 확보하는 경우 또는 장기 후측자료를 이용한 발생 빈도 추정이 통계적인 관점에서는 바람직하다고 판단된다. 그러나 자료 가용기간이 증가하는 경우, 기후변화에 의한 추세 성분 반영이 요구될 수 있으므로, 추세분석을 포함한 분석이 요구된다고 할 수 있다(Kim and Jang, 2020; Kim et al., 2023). 최근 수치모의기법을 이용한 장기 파랑 자료의 산출-제공이 활발한 상황에서 연안 안전, 재해 예방(저감)에 영향을 미치는 고파 발생 빈도 분석으로의 활용이 요구된다.

3.3 월별 변화 양상

다음은 뚜렷하게 드러나는 계절 영향을 파악하기 위하여 연간 발생 회수를 기준으로 월별 발생 빈도에 대한 비율과 그 비율에 대한 구간 추정을 수행하였다(Figs. 10-11 참조). 구간 추정은 이항분포 공식을 이용하였으며, R 프로그램에서 제공하는 prop.test() 함수를 사용하였다. 더불어 연평균 동일한 확률로 발생하는 조건에 해당하는 기준선을 이용하여 월별 발생 빈도의 평균 이상 또는 이하의 판단이 가능하도록 도시하였다. 이 기준선은 어떤 지점의 절대적인 연간 전체 발생 빈도에 대한 월별 평균 빈도로 간주된다. 그림에서 볼 수 있는 바와 같이, 월별 발생 빈도의 95% 신뢰구간은 자료 개수의 영향을 크게 받기 때문에 고파에 비하여 돌연 고파의 신뢰구간이 상당한 크기로 나타나고 있음을 알 수 있다. 모든 지점을 통합한 전체 지점에 대한 월별 발생비율은 통합으로 인한 자료개수 증가로 신뢰구간이 크게 감소하여 동해 연안 전체의 전반적인 월별 발생 빈도의 보다 정확하고, 정량적인 변화 양상 파악에 적합하다고 할 수 있다.

전반적으로 평균을 기준으로 계절변화 양상이 뚜렷하게 드러나고 있지만, 신뢰구간(추정 오차구간)을 고려하면 통계적으로 유의미한 수준으로 뚜렷하게 평균 이하의 발생 빈도를

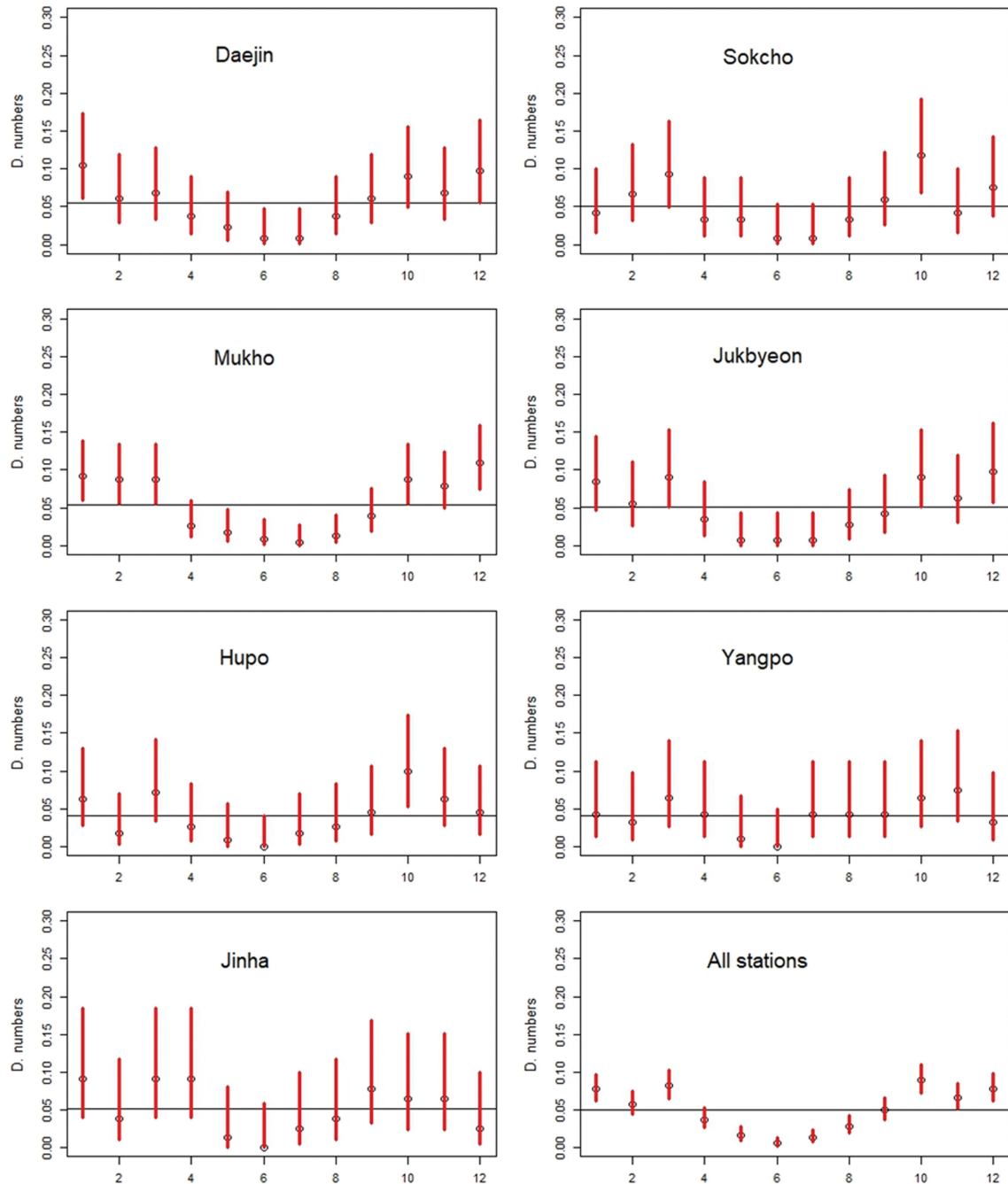


Fig. 10. Monthly variation and 95% confidence interval of the high wave occurrence proportion (horizontal solid line = mean proportion).

보이는 시기는 지점에 따라 차이를 보이고 있음을 알 수 있다. 대진은 6~7월, 묵호는 5~8월, 죽변은 5~7월, 후포는 6월로 한정된다. 반면 돌연 고파 조건에서는 대진, 속초는 6~7월, 묵호는 5~8월, 죽변은 5~7월, 후포는 6월로 제한된다. 따라서 모든 지점에서 고파, 돌연 고파가 평균 이하로 발생하는 공통 시기는 6월로 제한되고 있음을 알 수 있다. 그러나 전체 자료를 통합하여 계절변화 양상을 분석하는 경우, 고파, 돌연 고파 모두 5~8월 기간에 평균 이하 빈도를 보이고 있으며, 2월을 제외한 1, 3, 10~12월 기간은 평균 이상으로 발생하는 것으로 파악되었다. 2월도 전체적인 구간이 차지하는 부분이 평균에서 다소 높은 점을 감안하면 전체 6개월 동안

평균보다 발생 빈도가 높은 시기이고, 평균 구간을 차지하는 시기는 4월, 9월로 각각 발생 빈도의 저하, 상승 시기로 간주할 수 있다. 전반적으로 통계적인 기준으로 신뢰구간의 크기가 매우 크게 나타나고 있지만, 월별 발생비율 변화양상은 뚜렷하게 나타나고 있으며, 절대적인 발생 빈도는 묵호가 가장 크게 나타나고 있으며, 죽변, 대진, 속초가 그 다음 그룹, 후포, 양포, 진하가 가장 적은 그룹을 형성하고 있음을 알 수 있다. 고파에 대한 돌연 고파의 발생 비율은 후포, 양포 지점에서는 50% 조금 못 미치는 수준이지만, 다른 지점에서는 60~66% 정도로 발생 비율에 있어서 지점별로 크게 차이가 나타나고 있음을 알 수 있다.

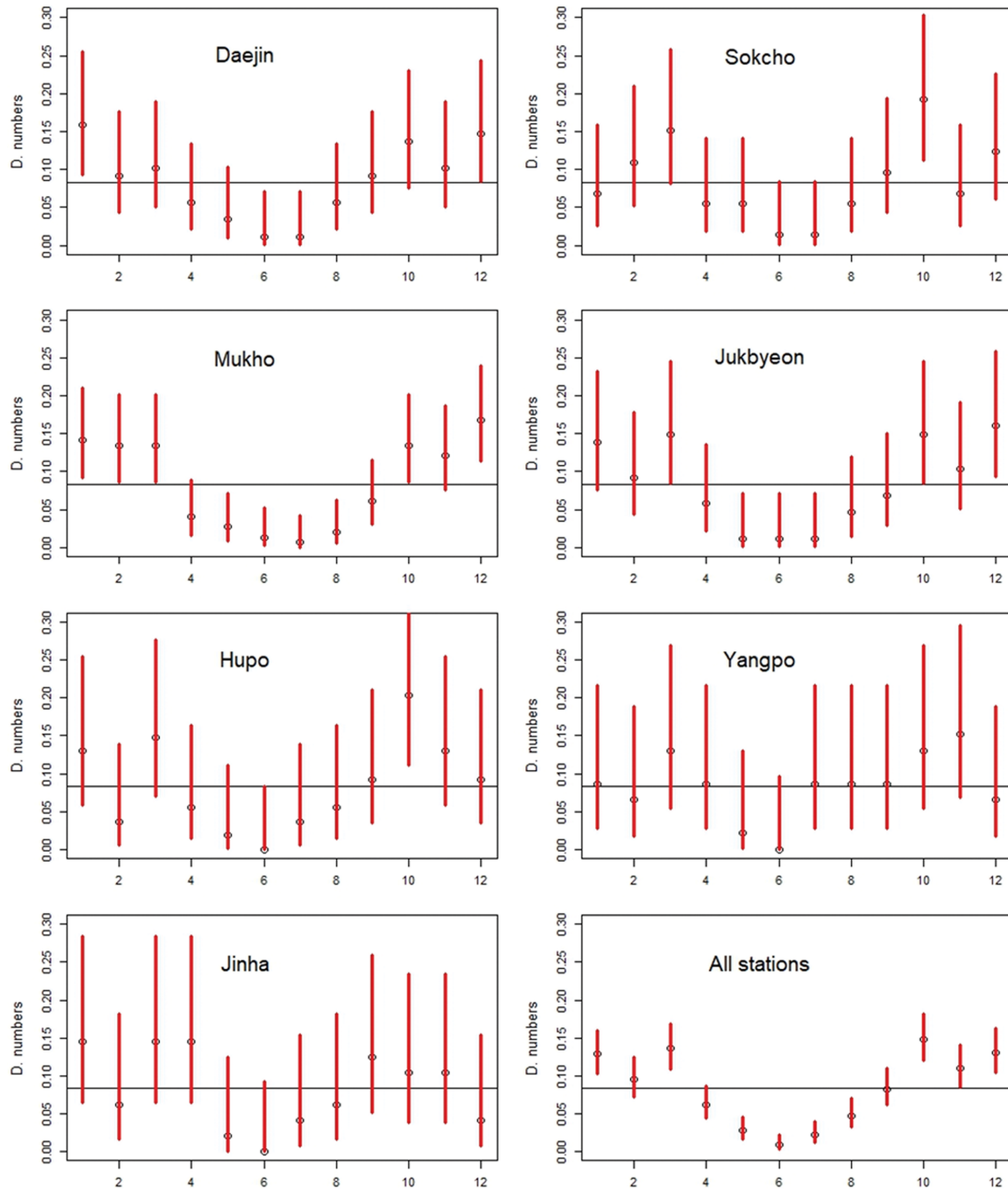


Fig. 11. Monthly variation and 95% confidence interval of the sudden high wave occurrence proportion (horizontal solid line = mean proportion).

3.4 공간 상관분석

동해연안 관측 지점의 공간적인 발생 빈도의 정량적인 차이 평가(Table 3. k_L , k_U 컬럼 범위 참조. 평균 발생 개수의 상한, 하한 - 95% 신뢰구간)와 더불어 고파, 돌연 고파의 발생 빈도의 변화 양상에 대한 공간 상관분석을 수행하였다. 상관분석은 연간 고파, 돌연 고파 발생 개수 자료를 이용하여 두 지점을 선택하여 수행하였다. 각각의 지점을 하나씩 선택하여 가능한 두 지점의 모든 쌍(pair)으로 조합하여 상관분석을 수행하였으며, 상관분석 결과는 Fig. 12에 그림으로 제시하였다. 그림에서 볼 수 있는 바와 같이 인접한 지점은 상대적으로 높은 상관 정도를 보이는 것으로 나타났다. 고파 조

건의 경우, 대진-속초, 후포-죽변의 경우 0.70, 0.76 정도의 높은 상관을 보이고 있음을 알 수 있는 반면, 양포-진하는 상관 정도가 매우 작게 나타나고 있는 것으로 파악되었다. 전체 7개 지점으로 구분되어 있으나, 고파 조건에서는 대진-속초, 묵호-죽변-후포, 양포, 진하 해역으로 구분되고 있는 것으로 파악된다. 반면, 돌연 고파 조건에서는 대진-속초에서 0.57 정도의 상관계수가 나타나고 있으며, 다른 모든 지점의 경우, 0.3 내외의 작은 상관 정도를 보이는 것으로 파악되었다. 따라서 전반적으로 약한 거리 상관과는 달리 각각의 관측 지점이 위치하는 해역의 독립적인 특성이 우세하게 나타나고 있는 것으로 판단된다.

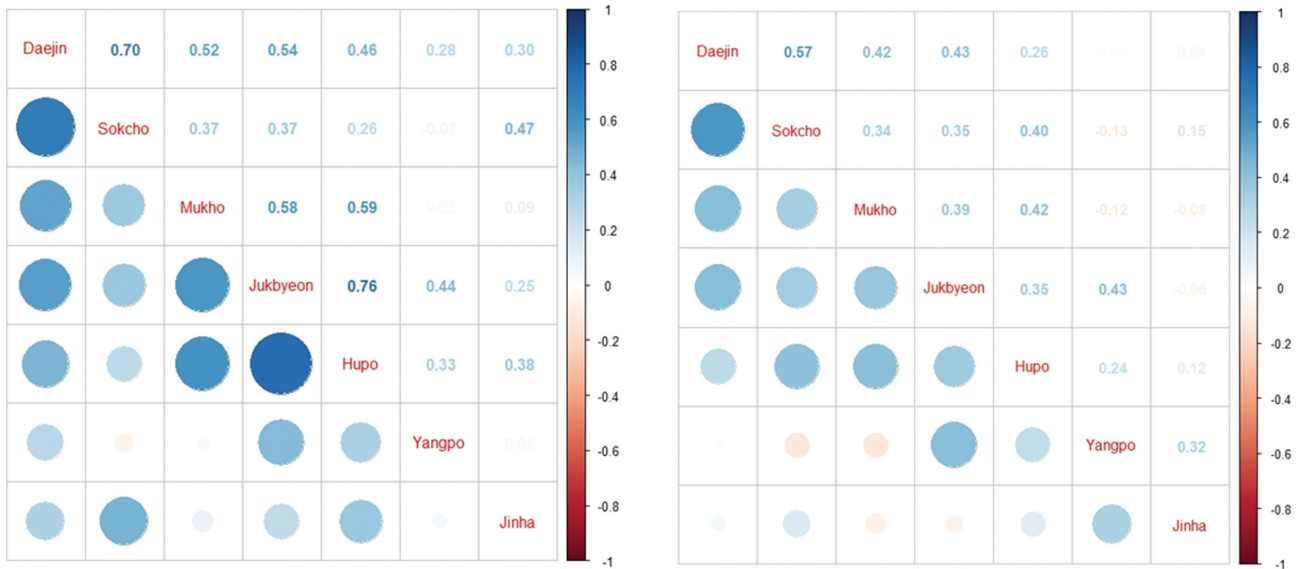


Fig. 12. Correlation plots (left panel = high waves, right panel = sudden high waves).

4. 결론 및 제언

본 연구로부터 얻은 주요 결론은 다음과 같다.

- 먼저 고파와 돌연 고파의 선정 방법을 제안하였다. 고파에 대해서는 유의과고가 3 m 이상인 연속된 고파 이벤트의 간격이 48시간 이내이고 그동안의 최소 유의과고가 1.5 m 이상이면 단일 고파 이벤트로 간주하였다. 돌연 고파에 대해서는 시점 기준 파고를 1.25 m, 고파 기준으로 3 m, 정온 유지 시간(t_c)으로 3시간, 돌연출현 시간(t_d)으로 12시간을 설정하고 24시간 이내에 부여된 조건을 만족시키는 고파 이벤트를 선정하는 방법을 제안하였다.
- 동해 연안 7개 지점의 장기 파랑자료에서 추출한 고파, 돌연 고파 발생 빈도 자료를 이용하여 Poisson 분포를 추정하고, 분포 적합검정을 수행한 결과 Poisson 분포가 통계적인 유의미한 수준에서 발생 빈도를 적절하게 적합하고 있는 것으로 파악되었다. 그리고 불연속 자료의 분포 적합 검정은 자료의 개수가 증가할 때 실질적인 의미를 가지는 효과 크기 이상의 차이보다 크게 더 작은 차이를 판정하는 Cramer-von Mises, Anderson-Darling 검정보다는 자료를 발생 개수 그룹으로 분류하여 분포적합을 판정하는 χ^2 검정이 보다 적절한 것으로 파악되었다.
- 한편, 등분산 조건에서 벗어나는 묵호 지점에 대하여 Zero-Inflated Poisson 분포 적합을 수행한 결과, 시각적으로 뚜렷하게 개선되는 적합 수준을 보이는 것으로 파악되었다. 매개변수 추정은 최우도법으로 수행하였으며, 간단하게 유도되는 모멘트법으로 추정되는 결과와 큰 차이를 보이지 않은 것으로 파악되었다.
- 한편, 연간 발생 빈도에 대한 월별 발생 비율은 뚜렷한 계절 변화를 보이고 있으나, 가용한 자료의 개수 제한으로 95% 신뢰구간이 0.2 정도의 범위에서 0.1(50%) 정도로 크

게 나타나고 있는 것으로 파악되었다. 모든 지점을 통합하는 경우, 발생 빈도가 평균 이하인 기간은 5~8월로 파악되었다. 4월, 9월은 평균 구간, 나머지 기간은 평균을 상회하는 구간으로 파악되었다. 그리고, 공간적인 상관은 대진-속초 지점을 제외하고는 작은 상관 정도를 보이는 것으로 파악되었다.

한편, 본 연구의 한계는 다음과 같다. 더불어 이 한계를 극복하기 위한 연구내용을 다음과 같이 제안한다.

- 고파, 돌연 고파의 빈도 추정 과정에서 다양한 지속시간 및 변화 양상을 보이는 파랑에서 독립적인 고파를 선정하는 객관적인 기준 수립이 요구된다. 본 연구에서 사용한 고파 발생 빈도에서 고파의 복잡한 변동 양상으로 인한 독립 조건을 만족하는 객관적인 고파 선택에 한계가 있다. 경험적인 판단과 더불어 폭풍 파랑의 독립성 검정을 통한 객관적인 기준을 통한 선택이 요구된다.
- 본 연구에서는 동해 연안 7개 지점, 17년 동안의 자료를 사용하였으나, 독립 조건을 만족하는 자료를 확보하기 위해서는 연간 발생 빈도 자료가 요구되는 상황이다. 따라서, 상대적으로 짧은 관측자료를 이용한 분포 추정은 높은 불확실성을 보일 수 있기 때문에, 40년 이상의 장기 후측 자료를 이용한 발생 빈도 추정 연구가 수행되어야 할 것으로 판단된다. 또한 장기 후측 자료의 경우, Jeong et al.(2016) 논문에서 제시하는 바와 같이 추세 검정을 포함하는 매개변수의 시간변화 양상을 반영하는 분포 적합 연구가 요구된다.
- 통계적인 검정 결과의 제 1종, 제 2종 오류를 모두 고려하는 정량적인 분석평가가 요구된다, 이를 위해서는 통계분석의 대상이 되는 다양한 통계량에 대한 실질적인 크기 기준, 효과 크기(effect size)에 대한 “Domain” 영역에서의 다양한 크기 제안-설정이 요구된다. 더불어 통계적으로 유의

미한 크기와 실질적으로 유의미한 크기에 대한 비교-검토가 요구된다.

References

- Arnold, T.B. and Emerson, J.W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *The R Journal*, 3(2), 34-39.
- Borchers, H. (2023). *pracma: Practical Numerical Math Functions*. R package version 2.4.4, <https://CRAN.R-project.org/package=pracma>.
- Cho, H., Jeong, W., Baek, W. and Kim, S. (2012). Analysis of the variation pattern of the wave climate in the Sokcho coastal zone. *J. of Korean Society of Coastal and Ocean Engineers*, 24(2), 120-127 (in Korean).
- Cho, H. and Lee, G. (2023). *Statistical Data Analysis in the Ocean Science with R Program*. KIOST.
- Chun, H.S., Kang, T.S., Ahn, K.M., Jeong, W.M., Kim, T.R. and Lee, D.H. (2014). A study on the statistical characteristics and numerical hindcasts of storm waves in East Sea. *Journal of Korean Society of Coastal and Ocean Engineers* 26(2), 81-95 (in Korean).
- Corless, R.M., Gonnet, G.H., Hare, D.E.G., Jeffery, D.J. and Knuth, D.E. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5, 329-359.
- Dencks, S., Piepenbrock, M. and Schmitz, G. (2020). Assessing vessel reconstruction in ultrasound localization microscopy by maximum-likelihood estimation of a zero-inflated Poisson model. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 67(8), 1603-1612.
- Faraway, J., Marsaglia, G., Marsaglia, J. and Baddeley, A. (2021). *gofest: Classical goodness-of-fit tests for univariate distributions*. R package version 1.2-3, <https://CRAN.R-project.org/package=gofest>.
- Heckert, N. and Filliben, J. (2003). *NIST/SEMATECH e-Handbook of Statistical Methods; Chapter 1. Exploratory Data Analysis*.
- Jeong, W.M., Ryu, K.H., Oh, S.H. and Baek, W.D. (2016). Trends of the storm wave appearance on the East Coast analyzed by using long-term wave observation data. *Journal of Korean Society of Coastal and Ocean Engineers*, 28(2), 109-115 (in Korean).
- Jeong, Y.C., Ryu, K.H., Ahn, S.J. and Jeong, W.M. (2026). Analysis of sudden high wave occurrence along the central-northern East Sea coast of Korea. *Journal of Korean Society of Coastal and Ocean Engineers* 38(2), 68-79 (in Korean).
- Kang, M.K. and Seol, D.I. (2022). Evaluation of the appropriateness of high wind wave alert by comparing the marine meteorological observation buoy data. *J. of Navigation and Port Research*, 46(1), 11-17 (in Korean).
- Kim, C.K., Lee, H.J., Kim, S.D., Oh, B.C. and Choi, J.E. (2023). Assessment of wave change considering the impact of climate change. *Journal of Korean Society of Disaster Security*, 16(4), 19-31 (in Korean).
- Kim, I.C., Oh, J.H. and Seo, K.D. (2016). A proposal for criterion of sudden high waves in the East Sea. *Journal of Korean Society of Coastal and Ocean Engineers*, 28(3), 117-123 (in Korean).
- Kim, K.H. and Jang, C.J. (2020). Linear trend in the wave climate of the East Sea estimated from a 40-year high-resolution wave hindcast. *Journal of Coastal Disaster Prevention*, 7(2), 73-83 (in Korean).
- Kim, S.U. and Kim, H.B. (2016). Comparison of probability distributions to analyze the number of occurrence of torrential rainfall events. *J. of Korea Water Resources Association*, 49(6), 481-493 (in Korean).
- Kim, T.J., Kwon, H.H. and Shin, Y.S. (2019). Frequency analysis of storm surge using Poisson-Generalized Pareto distribution. *J. of Korea Water Resources Association*, 52(3), 173-185 (in Korean).
- KMA (2026). *Monthly Marine Weather and Climate Information, Number of days of the severe wind wave alerts*, <https://marine.kma.go.kr/mmis/Notice?board=publication>.
- KOGAS (2026). *KOGAS Wind Wave and Swell Occurrence Data in East Sea. Public Data Dissemination*, <https://www.kogas.or.kr/site/koGas/1060302000000>.
- Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65, 297-303.
- NIST (2012). *NIST/SEMATECH e-Handbook of Statistical Methods*. National Institute of Standards and Technology, US Department of Commerce.
- Oh, S.H., Jeong, W.M., Lee, D.Y. and Kim, S.I. (2010). Analysis of the reason for occurrence of large-height swell-like waves in the East Coast of Korea. *Journal of Korean Society of Coastal and Ocean Engineers*, 22(2), 101-111 (in Korean).
- Seol, D.I. (2017). Distribution characteristics of high seas watch and warning in Korea. *J. of Navigation and Port Research*, 41(3), 103-108 (in Korean).
- Son, B.K. and Do, K.D. (2022). Numerical simulation of storm waves during typhoon MAYSAK and HAISHEN in the Korean Peninsula. *Journal of Coastal Disaster Prevention*, 9(1), 43-59 (in Korean).
- Trapletti, A. and Hornik, K. (2023). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-55, <https://CRAN.R-project.org/package=tseries>.
- Weisstein, E.W. (2026). *Poisson Distribution*. From MathWorld - A Wolfram Resource. <https://mathworld.wolfram.com/PoissonDistribution.html>.

Received 27 July, 2026

Revised 14 August, 2026

Accepted 17 August, 2026