

인공신경망과 가우시안 과정 회귀에 의한 규칙파의 조파기 입력파고 추정 Estimation of the Input Wave Height of the Wave Generator for Regular Waves by Using Artificial Neural Networks and Gaussian Process Regression

오정은* · 오상호**
Jung-Eun Oh* and Sang-Ho Oh**

요 **지** : 2차원 조파수조 내에서 취득된 규칙파 실험데이터를 머신러닝 기법으로 분석하여 천수 변형을 경험한 파랑으로부터 조파기의 입력파고를 예측하는 모델을 수립하고 그 성능을 검증하였다. 이를 위해 가장 대표적인 머신러닝 기법인 인공신경망(NN)과 비모수 회귀분석 방법 중 하나인 가우시안 과정 회귀(GPR) 모델을 각각 수립하고 두 모델의 예측 성능을 비교하였다. 전체 실험자료를 모두 한꺼번에 활용한 경우와 쇄파 발생 여부에 따라 자료를 구분한 경우에 대해 독립적으로 분석을 수행하였다. 데이터를 구분하지 않은 경우에는 NN 및 GPR 모델 모두 조파기 입력파고 값과 계측값 사이의 오차가 비교적 크게 나타났다. 반면에 데이터를 비쇄파 및 쇄파 조건으로 구분하면 조파기 입력파고의 예측 정확도가 크게 향상되었다. 두 모델 중에서는 NN 모델보다 GPR 모델의 성능이 전반적으로 더 우수한 것으로 나타났다.

핵심용어 : 머신러닝, 인공신경망, 가우시안 과정 회귀, 조파, 수리실험

Abstract : The experimental data obtained in a wave flume were analyzed using machine learning techniques to establish a model that predicts the input wave height of the wavemaker based on the waves that have experienced wave shoaling and to verify the performance of the established model. For this purpose, artificial neural network (NN), the most representative machine learning technique, and Gaussian process regression (GPR), one of the non-parametric regression analysis methods, were applied respectively. Then, the predictive performance of the two models was compared. The analysis was performed independently for the case of using all the data at once and for the case by classifying the data with a criterion related to the occurrence of wave breaking. When the data were not classified, the error between the input wave height at the wavemaker and the measured value was relatively large for both the NN and GPR models. On the other hand, if the data were divided into non-breaking and breaking conditions, the accuracy of predicting the input wave height was greatly improved. Among the two models, the overall performance of the GPR model was better than that of the NN model.

Keywords : machine learning, artificial neural network, gaussian process regression, wave generation, physical experiment

1. 서 론

1.1 데이터 기반 연구

데이터 과학 및 컴퓨터 과학 분야의 머신러닝(machine learning) 기법이 근래 획기적으로 발달함에 따라 각 분야에서 활용 범위가 넓어지고 있다. 특히, IT 및 센싱 기술의 발전으로 활용 가능한 데이터의 수량이 크게 증가하여 머신러닝 기법을 활용한 우수한 성과가 지속적으로 산출되고 있다. 이처럼 데이터에 의존하여 예측을 하거나 또는 경험적 관계

식을 찾는 연구방법을 데이터 기반 연구(data-driven research)라고 통칭할 수 있다. 이러한 데이터 기반 연구는 어떤 관심대상에 대한 이해를 쉽게 할 수 있을 뿐 아니라 시스템 형태로 표현할 수 있기 때문에 다른 사람과 정보를 쉽게 공유할 수 있다는 장점을 가지고 있다. 실제로 다양한 데이터를 활용하여 자연 속 물리 현상에 잠재된 미지의 관계성을 찾아내거나 귀납적 통찰력을 보편적으로 적용 가능한 이론으로 이끌어낼 수 있음이 여러 학문 분야에서 확인되고 있으며, 이는 해안공학 분야에서도 예외는 아니다.

*한국해양과학기술원 해양ICT융합연구센터 박사후연구원(Post-Doctoral Researcher, Maritime ICT R&D Center, Korea Institute of Ocean Science and Technology)

**창원대학교 스마트그린공학부 건설시스템공학전공 부교수(Corresponding author: Sang-Ho Oh, Associate Professor, Department of Civil Engineering, School of Smart and Green Technology, Changwon National University, Changwon 51140 Korea, Tel: +82-55-213-3779, coast.oh@gmail.com)

해안공학 분야에서도 데이터 기반 연구의 범위는 점차 여러 응용 분야로 확장되고 있다. 특히 머신러닝 방법이 많이 사용되고 있는데 몇 가지 사례로서는 수면변위를 추출하거나 (Kim et al., 2021) 해양에서의 파랑 조건을 추정하고(James et al., 2018), 쇄파 발생을 예측하며(Stringari et al., 2021) 해안구조물 주변에서의 파랑 전파(da Silva, 2018), 월파 (Formentin and Zanuttigh, 2019; den Bieman et al., 2021) 및 처오름(Beuzen et al., 2019)을 모델링하는 연구 등이 이루어졌다. 한편, 국내 해안공학 분야에서도 데이터 기반 연구 방법에 의한 연구 성과들이 최근 소개되고 있다(Lee and Suh, 2016; Lee et al., 2020). 그러나 아직까지는 그 수가 많지 않은 편이며, 향후 보다 다양한 데이터 기반 연구가 수행될 필요가 있다.

1.2 단면 조파수조 내 파고 추정

단면 조파수조를 이용한 수리실험에서는 계측 대상 위치에서의 파고 및 주기를 정확하게 재현하는 것이 우선적으로 중요하다. 그런데 상당수의 실험이 파도가 진행하면서 수심이 얕아지는 조건에서 이루어진다. 왜냐하면 외해에서 입사하는 파랑은 수심이 더 얕은 연안으로 진행하면서 천수 또는 쇄파 등 파랑변형이 발생하게 되는데 해안공학 분야의 대부분 실험에서 관심 영역은 이러한 천해역 또는 중간수심역에 해당하기 때문이다. 또 다른 한편으로는 조파기 설치 위치의 수심을 계측 대상 지역의 수심보다 더 깊게 설정하고 수조를 따라 바닥의 높이를 점차 높여서 수심을 작게 함으로써 계측 대상 위치에서의 파고를 폭넓은 범위에서 재현해 낼 수 있다.

그렇기 때문에 조파기가 설치된 수심에서의 파고와 계측이 이루어지는 위치에서의 파고는 대부분 다르다. 따라서 계측 위치에서 희망하는 파고 조건을 재현해 내기 위해서는 조파기 설치 위치의 수심과 계측 위치의 수심이 다른 경우가 대부분이어서 조파된 파랑이 계측 위치에 도달하기까지 천수, 쇄파 등 파랑 변형이 발생하기 때문이다. 불규칙파는 물론이고 상대적으로 조파시간이 짧은 규칙파 조건에서도 조파 시간 자체는 짧지만 수조 내 반사파가 소멸할 때까지 대기하는 시간이 필요하기 때문에 시행착오를 여러 번 할수록 계측을 위한 실험과 생성에 많은 시간이 소요된다.

이러한 어려움을 극복하기 위해서 사전에 수치모델링을 통해 적절한 입력파고를 찾는 방법을 활용하기도 하지만 이 또

한 모델링 시간과 그에 따른 노력이 듦다. 다른 방법으로 쇄파대 내에서의 파고 예측식을 활용하기도 하지만, 이러한 식은 사면 경사가 일정한 조건에서만 사용 가능하기 때문에 적용성에 한계가 있다. 따라서, 계측 대상 위치의 목표파고를 재현하기 위한 조파수조 입력파고를 예측하는 데이터 기반 모델을 개발하게 되면, 실험준비 과정에서 실험과 생성 시의 시행착오 과정을 없애거나 최소화함으로써 전체 실험 시간을 단축할 수 있다. 즉, 데이터 기반 모델에 근거하여 목표 천해파고를 재현하기 위한 조파기 입력파고 값은 사전에 추정하여 실험의 효율성을 높일 수 있다. 이러한 점을 고려하여 본 연구에서는 2차원 조파수조 내에서 계측된 규칙파 실험데이터를 데이터 기반 연구 기법으로 분석하여 관측된 파고로부터 조파 위치에서의 입력파고를 예측하는 모델을 수립하고 그 성능을 검증하였다.

2. 수리실험 자료

조파수조 내 입력파고 예측 모델 수립을 위한 규칙파 실험 자료는 길이 50 m, 폭 1.2 m인 2차원 조파수조에서 취득되었다(Oh and Lee, 2018). 이 조파수조는 수조의 바닥 높이가 일정하지 않고 Fig. 1에 보인 것처럼 수조 한쪽 끝으로부터 10 m 지점부터 20 m 지점 사이에서 0.35 m 높아지도록 제작되었다. 이처럼 조파수조의 바닥 높이를 달라지게 한 이유는 조파기 설치 위치의 수심을 더 깊게 함으로써 목표 파고를 더 잘 재현할 수 있기 때문이다.

조파판은 Fig. 1에 파란색으로 표시한 것처럼 수조 한 쪽 끝으로부터 3.4 m 위치에 위치하고 있다. 파랑은 조파판을 구동시키는 전용 소프트웨어에 목표 주기 및 파고의 값을 입력하여 생성된다. 그러나 통상적으로는 입력파고(H_{input})와 실제로 조파판 위치에서 생성되는 파고는 완전히 같지 않으며 약간의 차이가 발생하게 된다. 그 이유는 소프트웨어적으로 입력된 파고를 실제로 재현할 때 전기 신호를 기계적인 장치에 전달하여 구동시키는 과정에서 실제 조파판의 움직임이 목표로 하는 이상적(ideal)인 조파판의 움직임과 완전히 같을 수 없기 때문이다. 이러한 관점에서 입력 파랑 조건에 최대한 가까운 실제 파랑이 생성될수록 성능이 우수한 조파장치라고 할 수 있다.

총 217개의 규칙파 자료를 생성하였으며, Fig. 1에 붉은 색으로 표시한 것처럼 수조의 길이 방향으로 총 2개 지점($x = 9$ m, $x = 30$ m)에서 파랑을 계측하였다. $x = 9$ m에서의 파랑은 조파판 설치 위치($x = 3.4$ m)와 같은 수심에서 계측된 것이며,



Fig. 1. Locations of the wave board and the wave gauges placed along the wave flume.

Table 1. Ranges of the water depth, wave height, and wave period of the data

Location	d (m)	H (cm)	T (s)
$x = 3.4 \text{ m}$	0.7~1.0	5.0~47.9	1.00~5.15
$x = 9 \text{ m}$		4.5~43.9	1.00~4.95

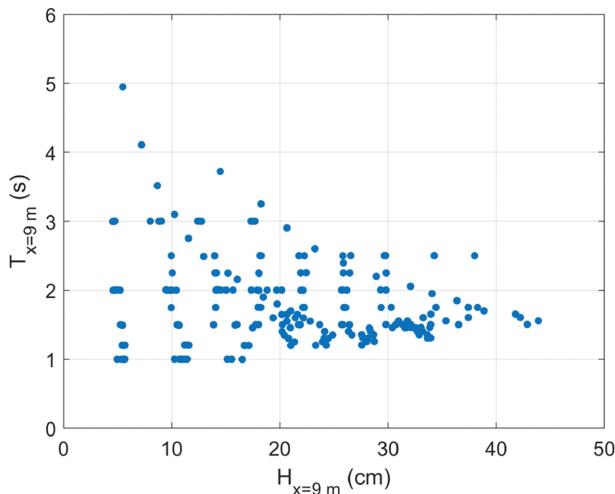


Fig. 2. Scatter plots of the wave height and period measured at $x = 9 \text{ m}$.

$x = 30 \text{ m}$ 에서의 파랑은 0.35 m 더 얕은 수심에서 계측된 것이다. Table 1에는 모든 데이터에 대해서 조파기 위치와 $x = 9 \text{ m}$ 에서의 수심, 파고 및 주기의 범위를 제시하였다. 또한 Fig. 2에는 $x = 9 \text{ m}$ 위치에서의 파고와 주기의 분포 경향을 보여주는 산포도를 제시하였다. 대부분의 수리실험이 실험실 환경에서의 주기 1~3 s 범위에서 주로 수행되는 점을 고려하여 이 범위에서 취득된 데이터가 대부분이며, 주기 3 s 이상의 데이터도 일부 존재한다.

Fig. 3에는 조파기에 입력한 파고와 $x = 9 \text{ m}$ 및 $x = 30 \text{ m}$ 에

서 계측된 파고를 각각 비교한 결과를 나타내었다. $x = 9 \text{ m}$ 에서의 파고($H_{x=9 \text{ m}}$)의 경우 입력파고(H_{input})와 같거나 극소하게 작음을 확인할 수 있다. $x = 9 \text{ m}$ 에서의 수심은 조파판이 설치된 위치($x = 3.4 \text{ m}$)의 수심과 동일하지만 앞에서 설명한 것처럼 조파장치의 전기/기계적 한계로 입력파고를 완벽하게 재현하지 못하기 때문에 차이가 발생하게 된다. 한편, $x = 30 \text{ m}$ 에서의 파고는 $x = 10\sim20 \text{ m}$ 구간의 경사면 및 $x = 20\sim30 \text{ m}$ 구간의 평탄면을 거치면서 일어나는 천수 또는 쇄파로 인해 파고가 증가하거나 감소하게 된다. 이처럼 $H_{x=30 \text{ m}}$ 와 H_{input} 사이의 관계가 비선형적인 관계를 나타내며 파랑의 주기 및 바닥 지형에 따라 영향을 받기 때문에 수리실험에서 $x = 30 \text{ m}$ 에서의 목표파고를 재현하기 위한 입력파고를 설정할 때 시행착오를 겪게 되는 것이다.

3. 자료 분석

머신러닝은 컴퓨터 알고리즘을 사용하여 데이터에 내재된 속성 및 경향을 찾는 인공지능 기법의 하나로서, 모형(model)을 구축하고 새로운 데이터에 대한 반복학습을 통해 지속적으로 성능을 향상시켜 나갈 수 있다는 장점을 가지고 있다 (Jordan and Mitchell, 2015). 머신러닝 기법 중에서도 가장 널리 알려지고 자주 쓰이는 기법 중 하나가 사람의 뇌 속 뉴런의 구조 및 작용을 모방한 인공신경망(artificial neural network)이다(Abiadun et al., 2018).

한편, 우리가 실험이나 관측 데이터로부터 어떠한 경향성을 찾아낼 때 흔히 사용하는 회귀(regression) 분석도 넓은 관점에서 보면 출력변수와 하나 이상의 입력변수 사이의 관계를 도출해내는 지도학습(supervised learning)의 일종이라고 해석할 수 있다. 지금까지 매우 다양한 회귀 분석 방법이 개발되어 왔는데 그 중에서도 가우시안 과정 회귀(Gaussian

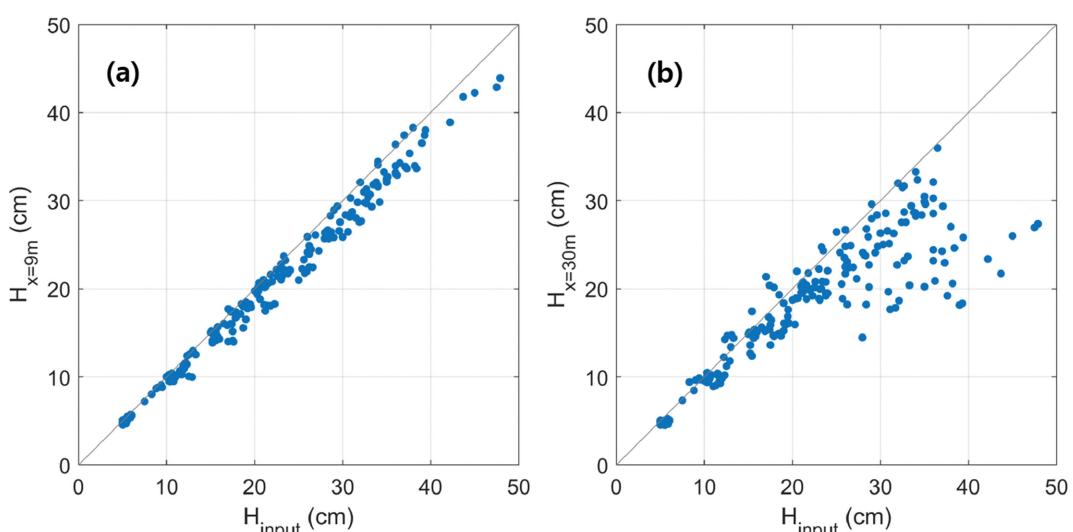


Fig. 3. Comparison of the wave height given as input to the wave generation system with the wave height measured at $x = 9 \text{ m}$ and $x = 30 \text{ m}$ where the water depths are different from each other.

process regression) 분석 기법은 변동성이 크고 유연한 회귀 결과를 제공하는 장점이 있다(Rasmussen and Williams, 2006).

본 연구에서는 조파수조 내 입력파고를 예측하는 모델 수립을 위해 인공신경망(NN)과 가우시안 과정 회귀(GPR) 모델 두 가지를 각각 적용해 보고 그 결과를 비교하였다.

3.1 인공신경망 모델

인공신경망 모델에서 신경망은 보통 여러 계층으로 구조화되며, 각 계층에는 사람의 뇌 속에 있는 생물학적 뉴런을 모방한 여러 개의 노드가 있다. 첫 계층인 입력층(input layer)에서는 외부에서 유입된 입력 신호가 수신된다. 이후의 계층은 각각 앞선 계층에서 보낸 출력을 받으며 마지막 계층인 출력층(output layer)에서는 최종 결과값이 도출된다. 입력층과 출력층 사이의 계층은 은닉층(hidden layer)이라고 한다. 다층 퍼셉트론(multi-layer perceptron)은 입력층과 출력층 사이에 하나 이상의 은닉층이 존재하는 인공신경망으로 Fig. 4와 같은 계층구조를 갖는다.

본 연구에서는 입력층과 출력층 사이에 10개의 노드를 가지는 1개의 은닉층으로 구성된 인공신경망을 구성하였다. 입력층의 자료로는 $x = 30\text{ m}$ 에서의 파고($H_{x=30\text{ m}}$) 및 수심($d_{x=30\text{ m}}$), 조파판 위치에서의 수심($d_{x=3.4\text{ m}}$) 및 입력파랑의 주기(T_{input})를, 출력층의 자료로는 입력파랑의 파고(H_{input})를 사용하였다. Fig. 4에는 이렇게 구성된 인공신경망의 개념도를 나타내었다. 활성함수로는 미분식이 간단하다는 장점으로 인공신경망에 자주 사용되는 시그모이드(sigmoid) 함수를 사용하였다. 확보된 217개의 전체 자료 중 85%에 해당하는 184개의 자료를 각각 인공신경망의 학습 및 검증을 통한 모델 구축(building)에 사용하였으며, 나머지 15%에 해당하는 33개 자료를 예측(prediction)에 사용하였다.

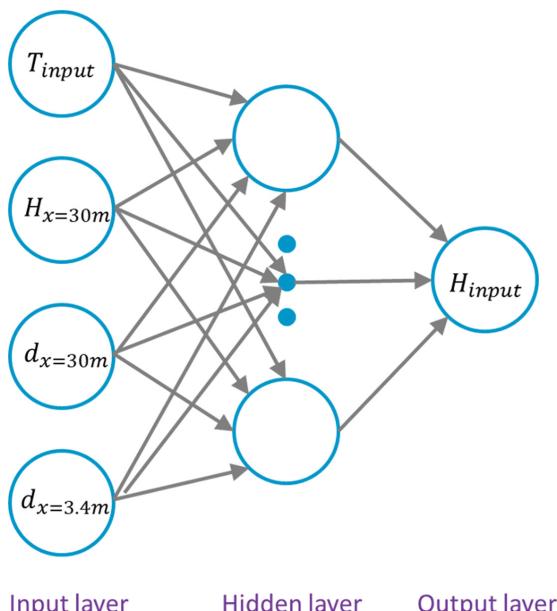


Fig. 4. Conceptual diagram of the artificial neural network.

3.2 가우시안 과정 회귀 모델

가우시안 과정 회귀(GPR) 모델은 비모수 커널 기반의 확률적 모델로서 다변량 정규분포를 무한한 차원으로 확장시키는 통계적 개념을 도입한 것으로 다음과 같이 표현될 수 있다(Rasmussen and Williams, 2006).

$$y = f(x) + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2) \quad (1)$$

여기에서 x 는 측정값, y 는 예측값이며, $f(x)$ 는 측정 잡음 ε 이 추가된 가능도(likelihood) 함수를 의미한다. 정규분포를 따르는 측정 잡음의 오차분산 σ^2 은 데이터로부터 추정된다. 가우시안 과정은 확률 변수들의 집합으로 취급되며 이 집합에 속한 임의의 유한 개의 확률 변수들은 결합 가우스 분포를 가지게 된다. 즉, n 개의 관측값 x_1, x_2, \dots, x_n 이 주어진 경우 확률변수 $f(x_1), f(x_2), \dots, f(x_n)$ 의 결합 분포는 가우스 분포이다. 자료의 분포를 가우시안 분포로 가정한다는 측면에서 가우시안 과정 회귀는 선형회귀와 유사하지만 무한차원의 특성함수를 이용할 수 있고 예측에 따른 불확실성을 다룰 수 있다는 측면에서 선형회귀와 차별성을 갖는다 (Kim et al., 2016).

가우시안 과정은 평균함수 $m(x)$ 와 공분산함수 $k(x, x')$ 에 의해서 정의되는데, 평균함수 $m(x) = 0$ 으로 가정하며 공분산함수 $k(x, x')$ 는 임의의 두 데이터 사이의 공분산을 구하는 커널(kernel)함수이다. 즉,

$$m(x) = E[f(x)] = 0 \quad (2)$$

$$k(x, x') = Cov[f(x), f(x')] \quad (3)$$

여기에서 x 와 x' 은 어떤 집단 내의 서로 다른 값을 나타낸다. 일반적으로 커널함수에는 방사상 기반함수(radial basis function)가 주로 사용되며 그 형태는 아래와 같이 표현된다.

$$k(x, x') = \sigma_f^2 \exp \left[-\frac{(x - x')^2}{2\sigma_m^2} \right] \quad (4)$$

여기에서 σ_m 은 입력변수 간 상대거리에 따른 상관성을 나타내는 커널모수이고 σ_f 는 출력변수의 크기를 조절하는 커널모수이다.

가우시안 과정 회귀 모델도 인공신경망 모델을 구축할 때와 동일한 184개의 자료를 이용하여 구축하였다. 마찬가지로 모델 수립 시 활용되지 않은 나머지 33개의 자료를 이용하여 모델의 성능을 예측하였다.

3.3 비쇄파와 쇄파 자료의 분리

본 연구에서는 모든 실험 자료를 한꺼번에 활용한 경우와 쇄파 및 비쇄파 조건에 해당하는 자료를 구분한 경우 각각 대해서 모델을 수립하였다. 이렇게 두 가지 경우로 나누어서 연구를 수행한 이유는 Fig. 3(b)에 보인 것처럼 쇄파와 비쇄파가 혼재된 실험 자료에 대해서도 두 가지 모델이 만족할 만한 수준의 성능을 보이는지를 우선 살펴보고, 쇄파와 비쇄파

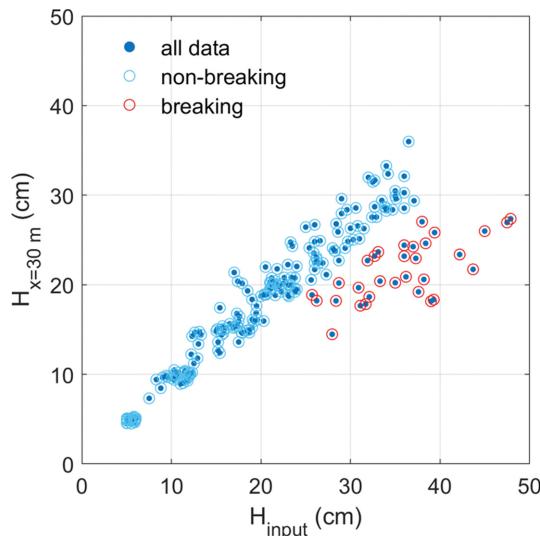


Fig. 5. Classification of the data depending on non-breaking or breaking conditions.

를 구분하여 모델을 수립했을 때 어느 정도의 성능 개선이 이루어지는지를 확인하기 위함이었다.

실험자료에서 쇄파와 비쇄파를 구분하기 위해서 쇄파 파고를 계산하는 다양한 경험식들(Goda, 2010)을 검토해 보았다. 그러나 이러한 쇄파 파고 예측식들은 해저면의 경사가 일정한 조건에서 적용 가능하며, Fig. 1에 보인 것처럼 일정한 경사면 이후 수평 구간이 이어지는 복합 단면에서의 쇄파 파고 예측에는 적합하지 않다. 실제로 이러한 식들을 사용하여 실험 자료로부터 비쇄파와 쇄파를 구분할 수 있는지 여부를 검토해 보았을 때 합리적인 결과를 제공하지 못함을 확인하였다.

이러한 점을 고려하여 본 연구에서는 아주 단순한 기준인 $H_b = 0.75H_{input}$ 을 적용하여 쇄파와 비쇄파를 임의적으로 나누었다. 즉, $x = 30\text{ m}$ 에서의 파고($H_{x=30\text{ m}}$)가 H_b 보다 크면 비쇄

파로, 그렇지 않으면 쇄파로 구분하였다. Fig. 5에는 이 기준에 따라서 실험 자료가 비쇄파와 쇄파로 구분된 결과를 나타내었다. 여기에서 주목할 점은 Fig. 5에 제시된 결과가 실제 수리실험에서의 쇄파 또는 비쇄파 조건에 잘 대응되는지 여부가 그다지 중요하지 않다는 점이다. 오히려, 상호 배타적인 두 집단의 자료를 이용하여 NN 모델 또는 GPR 모델을 독립적으로 구성할 때 각 집단에 속한 데이터의 경향성을 모델이 잘 인식할 수 있는 형태로 군집화가 이루어지는데 초점을 두는 것이 더 중요하다.

4. 분석 결과

4.1 비쇄파와 쇄파 자료를 구분하지 않은 경우

Fig. 6(a)는 전체 217개의 자료 중에서 85%에 해당하는 184개의 자료를 이용하여 구축된 인공신경망 모델에 의해 계산된 조파기 입력파고(H_{input})와 $x = 30\text{ m}$ 에서의 파고($H_{x=30\text{ m}}$)를 서로 비교한 결과를 제시하였다. 이 그림에서 알 수 있듯이 조파기에서의 입력파고가 작은 경우에는 인공신경망 모델이 어느 정도 실험 결과에 상응하는 결과를 나타내지만, 쇄파와 비쇄파가 발생하는 자료들이 혼재하는 상대적으로 입력파고가 비교적 큰 범위에서는 실험 결과와 큰 편차를 보인다. 이러한 경향은 모델 구축에 사용되지 않은 나머지 33개의 자료에 대한 인공신경망 모델 예측 결과인 Fig. 6(b)에서도 확인할 수 있다. 즉, 쇄파가 발생하는 경우와 그렇지 않은 경우의 데이터 경향이 확연히 다르기 때문에 인공신경망 모델이 그 특성을 제대로 파악해내지 못한다고 할 수 있다.

Fig. 7에는 인공신경망 모델에 의해 도출된 조파기 입력파고 값을 실제 수리실험에서의 관측값과 1:1로 비교한 결과를 나타내었다. 그림에서 파란색 점은 184개의 자료를 이용하여 구축된 모델의 결과에 해당하며 빨간색 점은 나머지 33개의

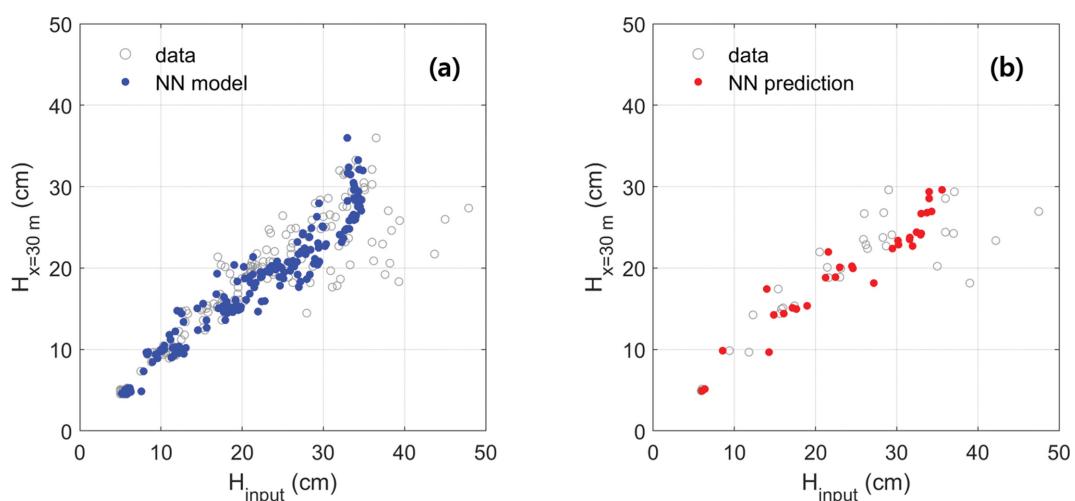


Fig. 6. Relationship between H_{input} and $H_{x=30\text{ m}}$ corresponding to the measured data (gray open dot), the builded NN model (blue solid dot), and the predicted result (red solid dot) by the NN model. The data were not classified according to non-breaking and breaking conditions.

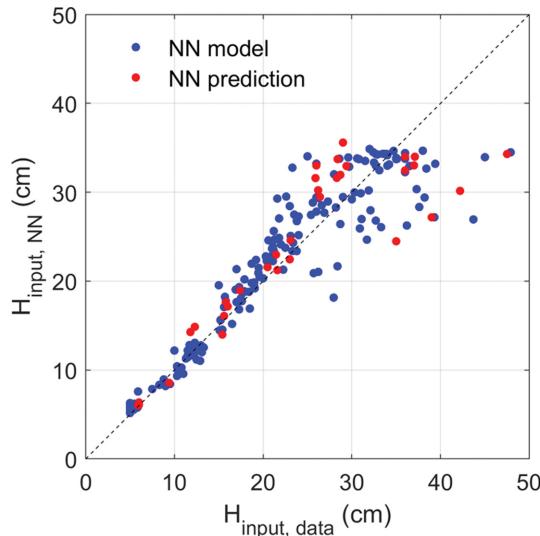


Fig. 7. Comparison of H_{input} from the measured data and the NN model. Blue and red dot correspond to the results of model building and prediction, respectively. The data were not classified according to non-breaking and breaking conditions.

자료를 이용하여 인공신경망 모델로 예측된 결과에 해당한다. 그림에서 알 수 있는 것처럼 입력파고가 대략 20 cm보다 작은 경우에는 모델 결과가 실험 자료와 어느 정도 유사하지만, 이보다 더 큰 범위에서는 편차가 꽤 크게 나타난다.

Fig. 8에는 GPR 모델에 의한 조파기 입력파고(H_{input})와 $x = 30$ m에서의 파고($H_{x=30\text{m}}$) 사이의 비교 결과를 제시하였다. GPR 모델도 NN 모델과 마찬가지로 조파기에서의 입력파고가 비교적 작은 범위에서는 어느 정도 실험 결과에 상응하는 결과를 나타내지만 입력파고가 비교적 큰 범위에서는 실험과 모델 결과가 큰 차이를 보인다. 그리고 이러한 경향은 실험과 모델에서의 조파기 입력파고(H_{input}) 값을 1:1로 비교한 결

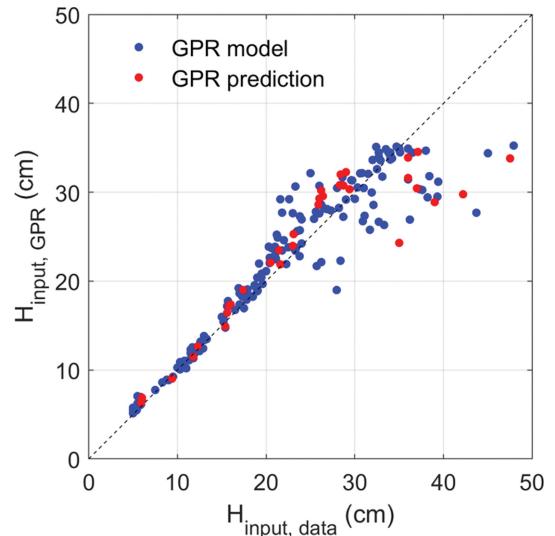


Fig. 9. Comparison of H_{input} from the measured data and the GPR model. Blue and red dot correspond to the results of model building and prediction, respectively. The data were not classified according to non-breaking and breaking conditions.

Table 2. RMSE values of the NN and GPR models

		All	Non-breaking	Breaking
NN	Model	3.83	1.10	2.58
	Prediction	5.16	1.39	3.02
GPR	Model	3.21	0.37	0.08
	Prediction	4.71	1.19	1.96

과를 나타낸 Fig. 9에서도 동일하게 확인된다. Table 2에는 NN 모델과 GPR 모델 수립 과정 및 각각 수립된 모델을 이용한 예측 과정에서의 평균제곱근오차(RMSE) 값을 나타내었다. 비록 두 모델 모두 RMSE 값이 큰 편이지만 상대적으

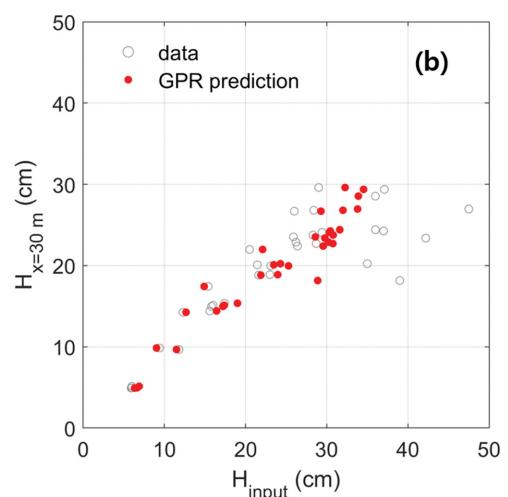
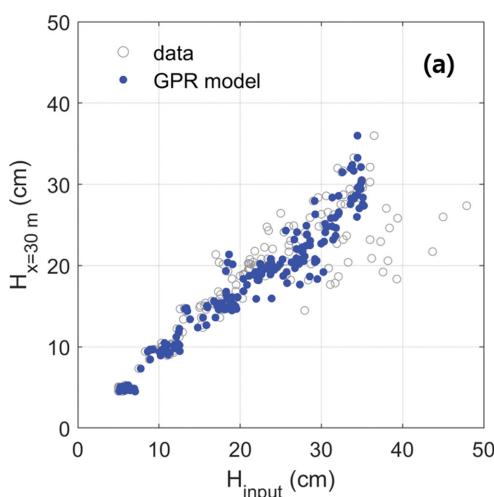


Fig. 8. Relationship between H_{input} and $H_{x=30\text{m}}$ corresponding to the measured data (gray open dot), the builded GPR model (blue solid dot), and the predicted result (red solid dot) by the GPR model. The data were not classified according to non-breaking and breaking conditions.

로는 GPR 모델이 조금 더 나은 결과를 제시함을 확인할 수 있다.

4.2 비쇄파와 쇄파 자료를 구분한 경우

Fig. 10에는 비쇄파와 쇄파 자료를 구분하여 NN 모델을 적용한 결과를 나타내었다. Fig. 6(a)와 비교해 보면 이렇게 자료를 구분함으로써 모델에 의한 H_{input} 과 $H_{x=30\text{m}}$ 의 상관관계가 크게 개선되었음을 확인할 수 있다. 쇄파 자료에 대한 모델의 경우 모델 구축에 사용된 자료의 개수가 상대적으로 적기 때문에 비쇄파 조건에 비해서 모델 적합도가 다소 저하되기는 하지만 Fig. 6(a)에서 쇄파 조건에 해당하는 자료를 제대로 모델링하지 못했던 점을 고려한다면 충분히 만족할 만한 수준의 결과를 얻었다고 할 수 있다. 실제로 Table 2에 보인 RMSE 값도 비쇄파와 쇄파를 구분하지 않은 경우에 비해서 더 작아졌음이 확인된다.

Fig. 11에는 이렇게 구축된 NN 모델을 이용하여 모델 수

립에 사용되지 않은 자료들에 대해 모델의 예측 성능을 평가한 결과를 비쇄파 및 쇄파 조건에 대하여 각각 제시하였다. 두 가지 조건에 대한 결과는 Fig. 10과 유사한 경향을 보이지만 자료와의 적합도는 다소 낮아지게 됨을 확인할 수 있다.

Fig. 12에는 NN 모델을 통해 도출된 조파기 입력파고(H_{input}) 값을 실험 자료와 1:1로 비교한 결과를 쇄파 및 비쇄파 조건에 대하여 각각 나타내었다. 쇄파와 비쇄파의 구분 없이 모든 자료를 한꺼번에 이용해서 구축된 NN 모델의 결과인 Fig. 7에 비해서 모델과 실험자료의 적합도가 크게 향상되었다. 그러나 여전히 비쇄파 조건에서는 일부 실험자료에서 모델과 실험자료의 편차가 존재한다.

한편, Fig. 13에는 비쇄파와 쇄파 자료를 구분하여 구축된 GPR 모델에 대한 결과를 나타내었다. 동일한 자료를 이용하여 구축된 NN 모델의 결과(Fig. 10)와 비교해 보면 GPR 모델이 실험자료와 더 잘 일치함을 알 수 있다. 즉, Fig. 10에 보인 것처럼 비쇄파와 쇄파 조건의 자료를 구분하여 NN 모

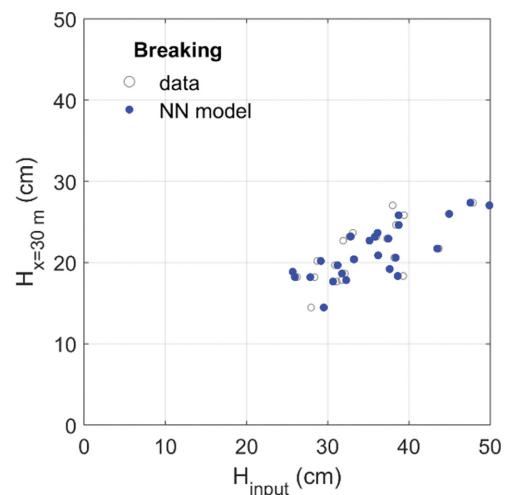
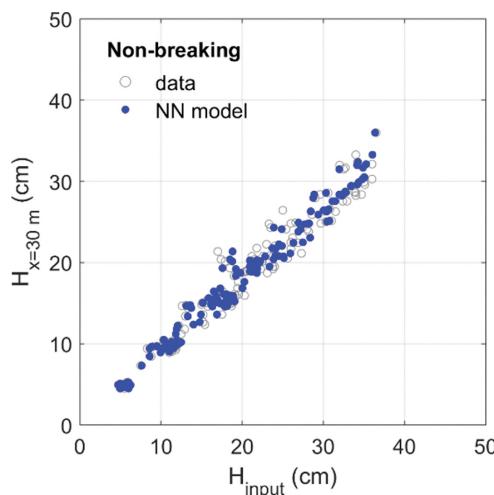


Fig. 10. Relationship between H_{input} and $H_{x=30\text{m}}$ corresponding to the measured data (gray open dot) and the builded NN model (blue solid dot). The model were built for the data corresponding to non-breaking and breaking conditions respectively.

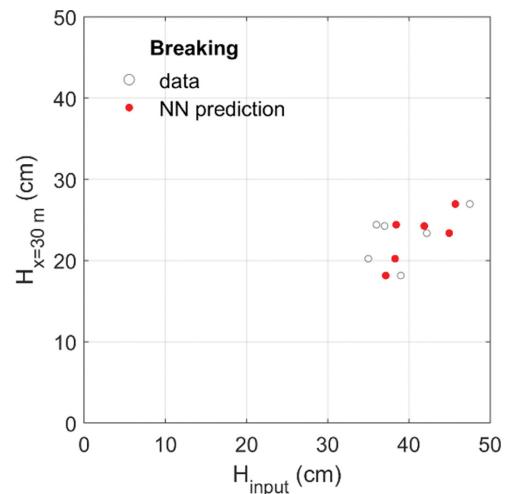
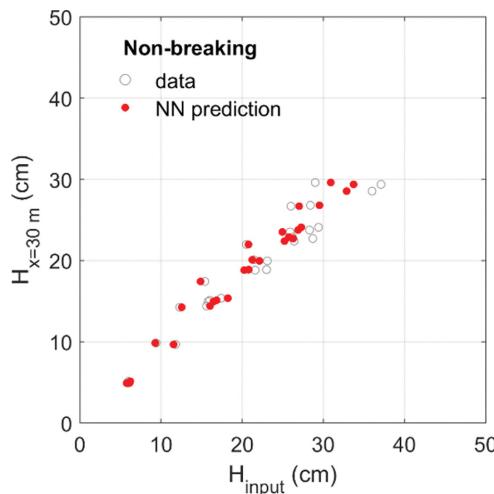


Fig. 11. Relationship between H_{input} and $H_{x=30\text{m}}$ corresponding to the measured data (gray open dot) and the predicted result by the NN model (red solid dot) for the non-breaking and breaking conditions.

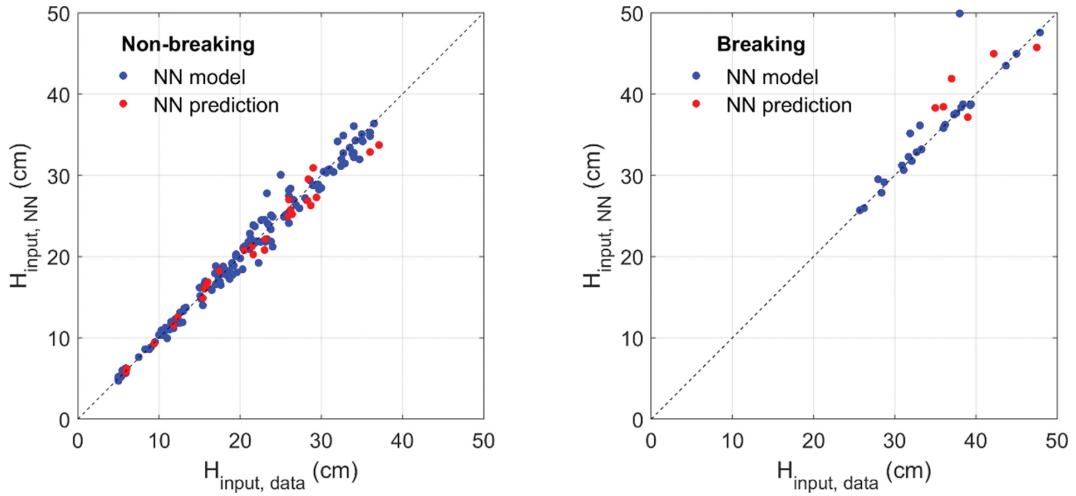


Fig. 12. Comparison of H_{input} from the measured data and the NN model for the non-breaking and breaking conditions. Blue and red dot correspond to the results of model building and prediction, respectively.

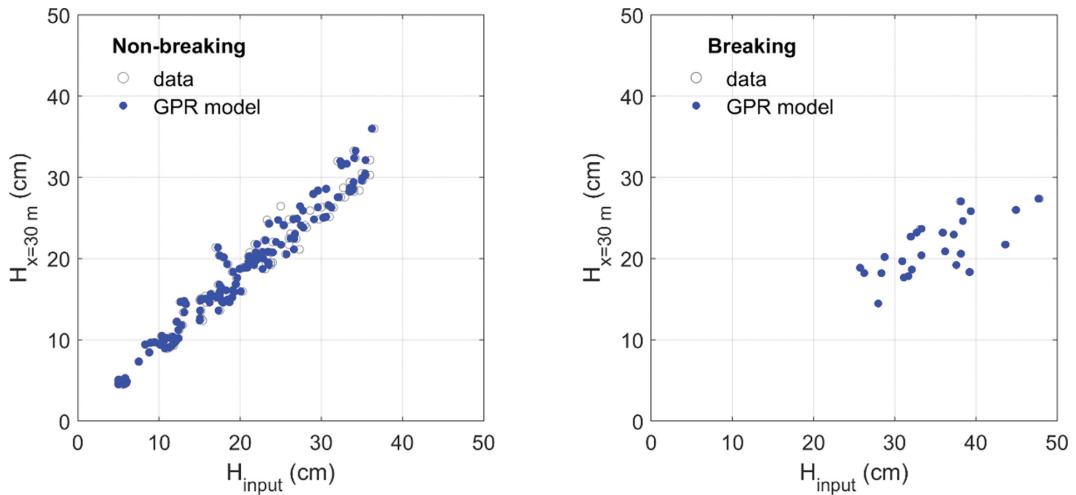


Fig. 13. Relationship between H_{input} and $H_{x=30\text{ m}}$ corresponding to the measured data (gray open dot) and the builded GPR model (blue solid dot). The model were built for the data corresponding to non-breaking and breaking conditions respectively.

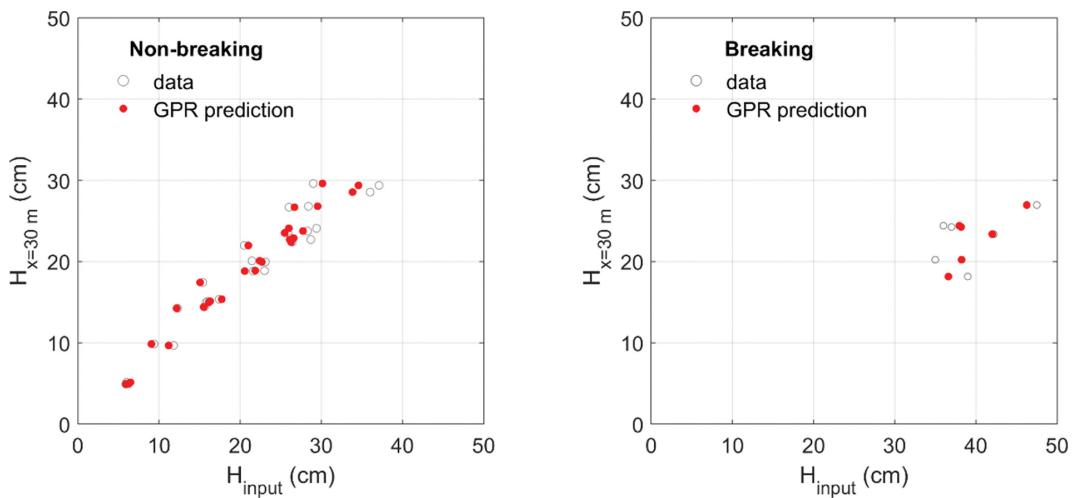


Fig. 14. Relationship between H_{input} and $H_{x=30\text{ m}}$ corresponding to the measured data (gray open dot) and the predicted result by the GPR model (red solid dot) for the non-breaking and breaking conditions.

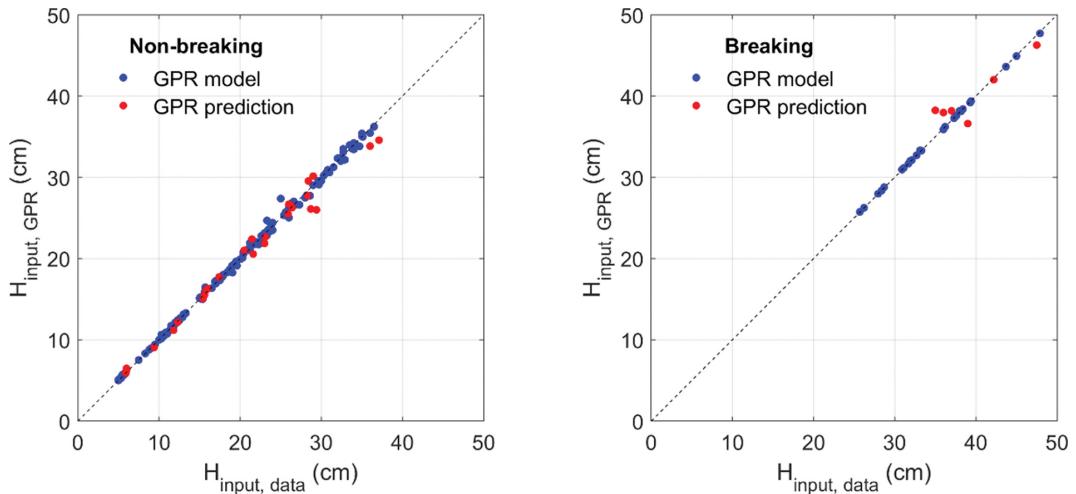


Fig. 15. Comparison of H_{input} from the measured data and the GPR model for the non-breaking and breaking conditions. Blue and red dot correspond to the results of model building and prediction, respectively.

델을 구축함으로써 H_{input} 과 $H_{x=30\text{m}}$ 의 상관관계가 만족할만한 수준으로 개선되었지만 GPR 모델은 NN모델보다 한층 더 좋은 결과를 보였다.

이러한 경향은 구축된 GPR 모델 수립에 사용되지 않은 자료로 모델의 예측 성능을 검증한 결과를 나타낸 Fig. 14에서도 확인된다. NN 모델과 마찬가지로 모델 구축 과정에 해당하는 Fig. 13의 결과에 비해서는 예측 과정에 해당하는 Fig. 14의 결과에서 실험자료와의 적합도는 다소 낮아지지만, Fig. 11에 보인 NN 모델의 예측 결과에 비해서는 H_{input} 과 $H_{x=30\text{m}}$ 의 상관관계가 더 나아졌음을 확인할 수 있다. 그리고 GPR 모델을 통해 도출된 조파기 입력파고(H_{input}) 값은 실험 자료와 1:1로 비교한 결과인 Fig. 15에서는 이러한 사실이 더 분명하게 나타난다. 즉, 비쇄파 및 쇄파 조건에서 모두 GPR 모델의 결과는 실험자료와의 편차가 NN모델에 비해서 훨씬 더 작게 나타났다. Table 2에 제시된 RMSE 값에서도 GPR 모델이 NN 모델보다 더 우수한 예측 성능을 보임을 알 수 있다.

4.3 토의

지금까지 살펴본 것처럼 조파 이후 천수 또는 쇄파 변형을 경험한 파랑으로부터 조파기 입력파고를 추정하는 데 있어 GPR 모델은 NN 모델에 비해서 일관되게 더 나은 예측 결과를 제공하였다. 연구를 수행하는 과정에서 NN 모델의 성능을 향상시키기 위해서 은닉층 개수를 2개로 늘리거나 신경망의 절점(node) 개수를 늘리거나 입력층 변수를 변화시키는 등의 시도를 해 보았지만 4.1절과 4.2절에 제시된 결과에 비해서 뚜렷하게 개선되는 결과를 얻지 못하였다.

인공신경망 기법은 다양한 분야에서 널리 활용되고 있는 반면 가우시안 과정 회귀 방법은 상대적으로 많이 알려진 편은 아니다. 이러한 이유로 NN 모델과 GPR 모델의 성능을 직접 비교한 연구가 거의 없지만 양자역학 분야에서 수행된 한 연구에서도 GPR 모델의 성능이 더 우수함을 보인 바 있다

(Kamath et al., 2018). 특히 데이터의 개수가 충분히 많지 않은 경우에도 GPR 방법이 꽤 만족할 수준의 예측 성능을 나타낸다는 점이 주요한 특징으로 설명되고 있으며(Knudde et al., 2020), 이러한 점은 본 연구에서도 확인되었다. Rasmussen and Williams (2006)에 의해 이론적 토대가 확립된 이후 여러 다양한 분야에서 GPR의 활용 범위가 증가하고 있는 추세이며, 해안공학 분야 연구에서도 비교적 적은 개수의 데이터를 이용한 회귀 분석 또는 시계열 분석이 자주 이루어지므로 GPR을 활용한 연구가 앞으로 많아질 수 있을 것으로 전망된다.

5. 결 론

본 연구에서는 단면 조파수조에서 규칙파를 이용한 수리실험을 수행할 때 수조 내 이격된 위치에서의 목표 파고를 재현하기 위한 조파기 입력파고를 예측하는 머신러닝 모델을 수립하고 그 성능을 검토해 보았다. 조파된 파랑이 수조 내 일정한 경사면과 그 이후의 평坦면을 진행하면서 파랑 변형을 거치기 때문에 파고와 주기 조건에 따라서 천수 과정을 거치게 되고 파고가 높은 경우에는 쇄파가 발생하게 된다. 이러한 파랑 변형 특성으로 인해 모든 실험자료를 이용하여 모델을 구축할 경우 입력파고 예측 정확도가 높지 않았다. 반면, 자료를 비쇄파와 쇄파 조건에 해당하는 두 집단으로 구분하여 독립적으로 모델을 구축하게 되면 입력파고의 예측 성능이 크게 개선됨을 확인하였다. 특히 인공신경망 모델에 비해서 가우시안 과정 모델의 성능이 더 우수한 것으로 나타났다.

이번에 구축된 가우시안 과정 회귀 모델은 본 실험자료를 취득한 단면 조파수조에서 향후 수리실험을 수행할 때 조파기 입력파고 설정에 활용될 수 있다. 그리고 이 모델을 적용할 경우 목표 실험파가 더 빠르게 취득되고 궁극적으로는 조파에 소요되는 시간이 실제로 단축되는지 여부를 확인하고자

한다. 그리고 모델의 효용성이 입증된다면 추가적인 계측자료를 취득하여 모델의 성능을 더 향상시키는 것도 가능할 것으로 기대된다. 특히 이번 연구에서는 규칙파 실험자료에 대해서만 연구를 수행하였지만 불규칙파 실험의 경우에도 가우시안 과정 회귀 모델 또는 인공신경망 모델로 조파기 입력파고를 적절하게 예측할 수 있다면 실험 수행 시의 효용성을 더욱 높일 수 있다.

비록 본 연구를 통해서 수립된 모델은 해당 조파수조에 대해서만 한정적으로 사용 가능하지만 근본적인 방법론은 다른 수리실험 또는 수치모델링 환경에서도 입력파고 설정 시 동일하게 적용될 수 있다. 따라서 조파수조 또는 수치파동수조를 빈번하게 활용하는 사용자라면 본 연구에서와 같이 입력파고를 예측하는 모델을 자체적으로 수립함으로써 실험 또는 모델링에 드는 시간 및 비용을 절감할 수 있을 것이다.

감사의 글

이 논문은 2021~2022년도 창원대학교 자율연구과제 연구비 지원으로 수행된 연구결과이며, 이에 감사드립니다.

References

- Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A. and Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), e00938.
- Beuzen, T., Goldstein, E.B. and Splinter, K.D. (2019). Ensemble models from machine learning: an example of wave runup and coastal dune erosion. *Natural Hazards and Earth System Sciences*, 19(10), 2295-2309.
- da Silva, S. (2018). Data-driven model identification of guided wave propagation in composite structures. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 40(11), 1-10.
- den Bieman, J.P., van Gent, M.R. and van den Boogaard, H.F. (2021). Wave overtopping predictions using an advanced machine learning technique. *Coastal Engineering*, 166, 103830.
- Formentin, S.M. and Zanuttigh, B. (2019). A genetic programming based formula for wave overtopping by crown walls and bulbous noses. *Coastal Engineering*, 152, 103529.
- Goda, Y. (2010). Reanalysis of regular and random breaking wave statistics. *Coastal Engineering Journal*, 52(1), 71-106.
- James, S.C., Zhang, Y. and O'Donncha, F. (2018). A machine learning framework to forecast wave conditions. *Coastal Engineering*, 137, 1-10.
- Jordan, M.I. and Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Kamath, A., Vargas-Hernández, R.A., Krems, R.V., Carrington Jr, T. and Manzhos, S. (2018). Neural networks vs Gaussian process regression for representing potential energy surfaces: A comparative study of fit quality and vibrational spectrum accuracy. *The Journal of Chemical Physics*, 148(24), 241702.
- Kim, H.G., Park, E., Jeong, J., Han, W.S. and Kim, K.-Y. (2016). Groundwater level trend analysis for long-term prediction based on Gaussian process regression. *Journal of Soil and Groundwater Environment*, 21(4), 30-41 (in Korean).
- Kim, J., Kim, T., Oh, S.-H., Do, K., Ryu, J.-G. and Kim, J. (2021). Deep visual domain adaptation and semi-supervised segmentation for understanding wave elevation using wave flume video images. *Scientific Reports*, 11, 21776.
- Knudde, N., Raes, W., De Bruycker, J., Dhaene, T. and Stevens, N. (2020). Data-efficient gaussian process regression for accurate visible light positioning. *IEEE Communications Letters*, 24(8), 1705-1709.
- Lee, J.S. and Suh, K.D. (2016). Calculation of stability number of tetrapods using weights and biases of ANN model. *Journal of Korean Society of Coastal and Ocean Engineers*, 28(5), 277-283 (in Korean).
- Lee, K.-H., Kim, T.-G. and Kim, D.-S. (2020). Prediction of wave breaking using machine learning open source platform. *Journal of Korean Society of Coastal and Ocean Engineers*, 32(4), 262-272 (in Korean).
- Oh, S.-H. and Lee, D.S. (2018). Two-dimensional wave flume with water circulating system for controlling water level. *Journal of Korean Society of Coastal and Ocean Engineers*, 30(6), 337-342 (in Korean).
- Rasmussen, C.E and Williams, C.K. (2006). Gaussian processes for machine learning, MIT press, Cambridge, Massachusetts.
- Stringari, C.E., Guimarães, P.V., Filipot, J.F., Leckler, F. and Duarte, R. (2021). Deep neural networks for active wave breaking classification. *Scientific Reports*, 11(1), 1-12.

Received 8 December, 2022

Revised 21 December, 2022

Accepted 21 December, 2022