

Original Article

딥러닝을 이용한 파랑·풍속 상호자료의 예측 기법

김태운[†]

한국환경정책·평가연구원 연구위원

A Study on the Prediction Technique for Wind and Wave Using Deep Learning

Taeyun Kim[†]

Senior Research Fellow, Korea Environment Institute, Sejong 30147, Korea

요 약

본 연구는 딥러닝 기법을 활용하여 해양에서의 파랑과 바람을 예측하였다. 2014년부터 2018년까지 해양에서 관측된 시간별 자료 41,693개를 딥러닝 모델의 학습자료로 활용하였다. 학습된 딥러닝 모델을 이용하여 2019년도의 바람과 파랑을 예측하고 관측된 자료로 비교·검증하였다. 유의파고 예측에서는 풍속, 풍향, 파향을 입력 자료로 사용하였으며, 모델 결과는 변동 패턴을 적절히 재현하고 상관관계 계수 0.83, 평균 절대 오차 0.283 m을 보였다. 풍속 예측에서는 입력 자료로 유의파고, 파향, 파주기, 풍향을 적용하였고 관측치와 예측치의 상관관계 계수는 0.836, 평균 절대 오차는 1.27 m/s를 보였다. 마지막으로 딥러닝 모델이 매우 효과적으로 최대파고의 예측을 수행하였다.

Abstract – This study aims to predict wave and wind using deep learning technique. The data for training the deep learning model were used hourly data (41,693) observed from 2014 to 2018, and the trained deep learning model was applied to the prediction of wind and wave in 2019. To predict the significant wave height, we used wind speed, wind direction, and wave direction as input data. The model results showed similar variation patterns with observed and predicted values. The correlation coefficient was 0.83 and the mean absolute error was 0.283 m. In the wind speed prediction, significant wave height, wave direction, wave period, and wind direction were used as input data. The deep learning model results showed error coefficients of correlation coefficient (0.836) and mean absolute error (1.27 m/s). Finally, the deep learning model reproduced the maximum wave height in good agreement with observations.

Keywords: Deep learning(딥러닝), Prediction technique(예측기법), Wave(파랑), Wind(바람)

1. 서 론

해양 및 연근해에 설치된 해상부이는 기상 및 해황을 관측하여 해상통항, 해양오염, 어업, 관광, 연안활동 등에 활용되고 있다. 측정된 자료는 사고시 대응, 어민에게 해황 정보 제공, 예측기술 개발, 연안재해 예방 등과 같은 다양한 분야에 중요하게 이용되고 있다(Kim and Kim[2014]). 한국에서는 1996년부터 해양기상부이가 운영되었으며 현재 21개소에서 기온, 바람, 기압, 습도, 파고 등을 관측하고 있다. 해양기상부이는 다양한 기상장비를 장착하고 있으며 관측된 정보는 위성 등과 같은 원격통신을 이용하여 자료를 전송하고 있다(KMA[2020]). 하지만 기상악화, 측정장비의 훼손, 자

료변환·처리의 오류 등으로 인하여 관측이 이루어지지 못하는 기간도 발생한다. 또한 넓은 해상에 해황정보 및 기상정보를 측정하기 위해 다수의 해상부이를 설치하기에는 경제적 부담이 있다.

이러한 미관측자료의 발생과 한정된 관측지점으로 인하여 연구에 필요한 정보를 추정 또는 예측하는 다양한 연구가 진행되어 왔다. 일반적으로 공간보간법(역거리가중법, 지역다항식보간법 등)이 널리 이용하고 있으나 이를 적용하기 위해서는 인근 지역의 공간적인 자료가 충분해야 한다(Lee[2010]). 관측된 자료들의 상관관계를 분석하거나 예측함수를 개발하여 필요한 자료를 얻을 수도 있다(Ko *et al.*[2017]; Choo *et al.*[2019]). 또한 수치모형을 활용하여 예측을 할 수 있으나 광범위한 지역의 경계자료 및 다양한 입력자료가 요구된다(Lee *et al.*[2014]).

최근에 발전하고 있는 추정 및 예측기술 중에 하나인 딥러닝

[†]Corresponding author: kimty@kei.re.kr

(Deep learning) 기법은 급격히 발전하고 있는 컴퓨팅 프로세스와 방대한 자료를 저장할 수 있는 저장 공간의 확보로 활발히 이용되고 있다(LeCun *et al.*[2015]). 딥러닝 기법의 가장 큰 장점 중에 하나는 의료, 이미지 분석, 기상예보, 무인 운전, 전력시스템, 금융 등과 같이 광범위한 분야에 적용할 수 있다는 것이다(Sengupta *et al.*[2019]; Shrestha and Mahmood[2019]). 최근에는 한국에서도 연안 방재 시스템, 어종 판별 시스템, 뜰개 관측등과 같은 해양 분야에 딥러닝 기술을 활용한 연구가 활발히 진행 중이다(Kim and Kim [2018]; Kang[2019]; Kim *et al.*[2019]).

따라서 본 연구에서도 딥러닝 기법을 적용하여 해상에서 관측된 자료들을 활용하여 연구에 필요한 정보를 추정 또는 예측하는 방법을 제시하고자 한다. 해상에서 발생하는 파고는 바람의 크기, 취송거리, 지속시간과 밀접한 관계를 이루고 있다. 이에 바람자료를 활용하여 유의파고를 추정하는 기법 및 결과를 제시하고 반대로 유의파고 자료를 활용하여 바람의 세기를 추정하고자 한다. 마지막으로 연안 범람 등에 직접적으로 영향을 미치는 최대 파고를 예측하기 위해 딥러닝과 유의파고자료를 활용하여 예측 정확도를 분석하고자 한다.

2. 방법론

2.1 해상부이자료

기상청에서 운영하고 있는 해상부이 자료(외연도 지점)를 딥러닝 기법에 적용하였다(Fig. 1). 외연도 지점의 해상부이는 충청남도

보령 인근 해안에서 약 80 km 이격된 서해 해상에 위치하고 있으며(위도 36.25, 경도 125.75) 풍속, 풍향, Gust 풍속, 기압, 습도, 기온, 수온, 최대파고, 유의파고, 평균파고, 파주기, 파향을 측하고 있다. 기상청 기상자료개방포털에서 시간별 관측 자료를 얻을 수 있으며 2017년 기준 품질검사결과 정상판정자료율이 90%이상을 보이고 있다(KMA[2020]). 해상부이에서는 매시간별 관측자료가 저장되고 있으며, 본 연구에서는 2014년부터 2019년까지 관측된 시간별 자료를 활용하였다. 매년 8,760개의 시간별 관측자료를 얻을 수 있으나 기상악화 및 시설물 수리 등으로 인하여 관측이 수행되지 못한 시점이 발생하였으며, 이들 자료를 제외한 자료의 갯수는 2014년 8,441개, 2015년 7,849개, 2016년 8,583개, 2017년 8,434개, 2018년 8,386개, 2019년 8,301개이다.

2.2 딥러닝 기법

인간의 뇌와 유사한 기능을 수행할 수 있는 인공지능을 지속적으로 개발해 오면서 기계학습의 효과성을 인지하게 되었으며, 기계학습의 여러 기법 중에 하나가 딥러닝 기법이다. 기존의 접근 방법은 데이터의 상호연관 관계를 인간이 지정하거나 수식으로 입력하였다면 딥러닝 기법은 대량의 자료를 학습시키면서 기계가 직접 데이터의 상호연관 관계를 찾아내도록 하는 것이다. 딥러닝 기법은 인간의 신경망(Neural Network) 형태를 기초로 하여 구성되어 있으며 입력층(Input layer), 중간층(Hidden layer), 출력층(Output layer)으로 구분된다(Fig. 2). 단순한 형태의 신경망 구조는 입력층, 중간층, 출력층이 각각 하나를 가지고 있으며, 다중의 중간층을 가진

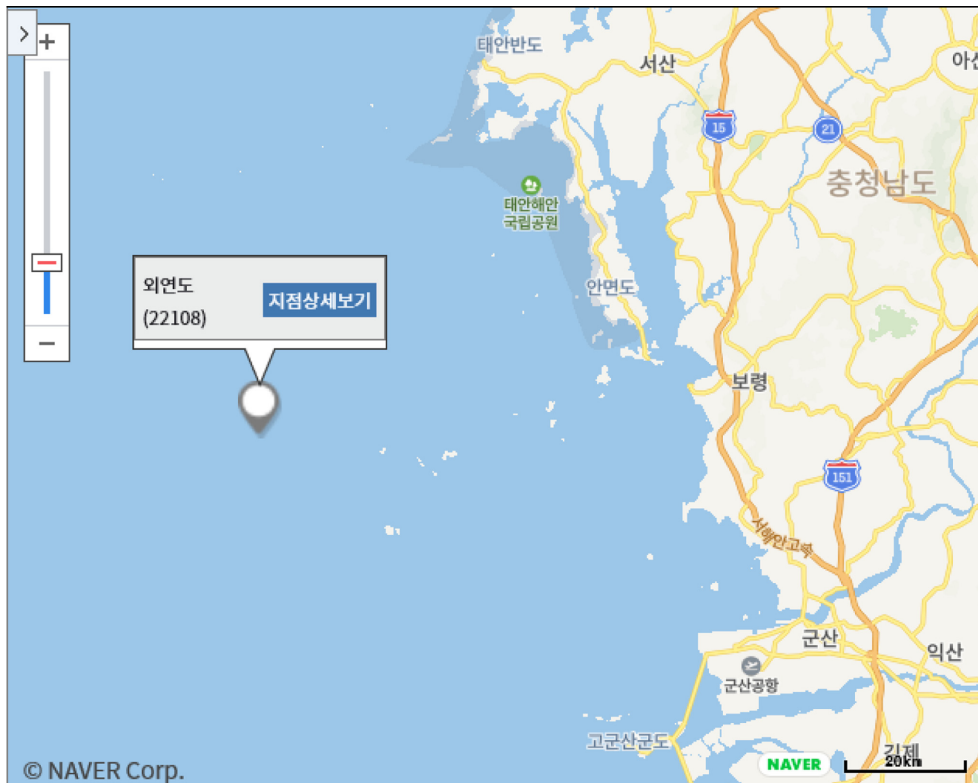


Fig. 1. Status of marine buoy site (KMA[2020]).

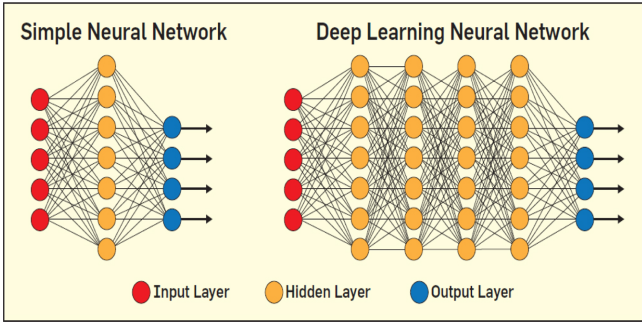


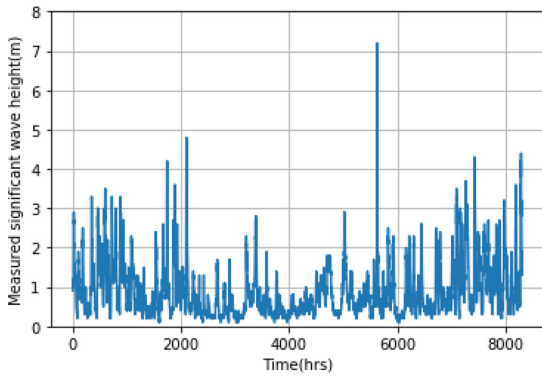
Fig. 2. Schematic of deep learning technique (Edwards[2018]).

신경망을 깊은 신경망(Deep Learning Neural Network, DNN)으로 정의하고 이를 활용한 기계학습을 딥러닝 기법이라 명명한다 (Kujira[2017]; Edwards[2018]; Cho[2020]).

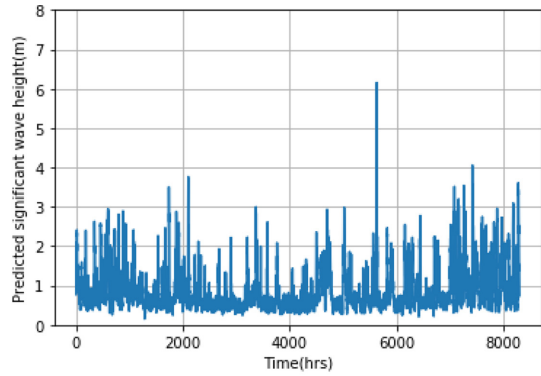
본 연구에서는 딥러닝 툴킷(toolkit)으로 텐서플로우(tensorflow) 1.12.3 버전을 포함한 keras 2.2.5을 사용하였으며 플랫폼은 파이썬(python) 3.6을 이용하였다. 딥러닝 기법은 가중치와 편향을 조절하여 오차를 최소화시키는 방법으로 가중치를 적절히 산정하는 것이 중요하다. 각각의 중간층을 거치면서 가중치가 조정되는 데, 중간층이 많을수록 기울기 소실이 발생하여 가중치의 수정이 어려워

진다. 이를 해결하기 위해서 활성화 함수가 사용되고 있으며, 시그모이드 함수, 하이퍼볼릭 탄젠트 함수, ReLU함수, 소프트플러스 함수 등이 주로 사용된다(Cho[2020]). 여러 개 중간층을 거치더라도 기울기 손실의 발생이 적은 ReLU함수를 동 연구에 적용하였다.

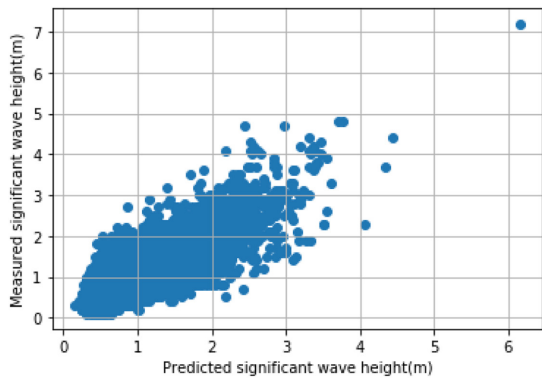
그리고 딥러닝 기법은 성능을 최적화하기 위해 예측값과 실측값의 오차를 비교하여 가중치를 조절한다. 이때 이용되는 오차함수는 mean squared error, mean absolute error, hinge, logcosh 등이 있으며, 유의파고 예측에서는 logcosh 오차함수를 적용하였다 (Keras[2020]). 딥러닝 기법에서는 가장 작은 오차를 찾아가기 위해서 미분 기울기를 이용하고 있으며 이를 경사 하강법이라고 명명한다. 미분 기울기를 토대로 가중치 수정이 이루어지는 경사 하강법은 정확성은 뛰어나지만 계산량이 매우 많아 계산속도가 느리다. 이를 보완하기 위해 고급 경사 하강법을 사용하고 있으며 확률적 경사 하강법, 모멘텀, 아다그리드, 알엠에스프롭, 아담 등의 다양한 방법이 이용된다. 확률적 경사 하강법은 전체 데이터를 사용하는 대신에 불규칙적으로 추출한 데이터를 사용하여 계산속도를 줄이는 것이며, 모멘텀은 이전 수정값을 참고하여 가중치를 조정하는 방법이다. 다양한 방법을 상호 보완하여 새로운 방법들이 개발되고 있으며 최근에 널리 사용되는 아담을 동 연구에 사용하였다(Cho [2020]).



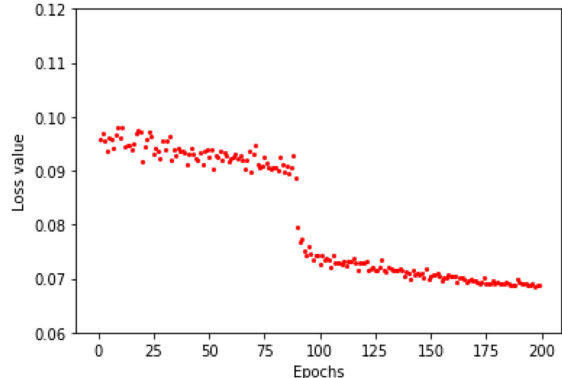
(a) Measured significant wave height



(b) Predicted significant wave height



(c) Measured significant wave height vs. Predicted significant wave height



(d) Loss value with increasing epoch

Fig. 3. Prediction results of significant wave height (a. measured significant wave height, b. predicted significant wave height, c. comparison of measured and predicted significant wave height, d. loss value with increasing epoch).

3. 예측기술결과

3.1 유의파고예측

외연도 지점의 유의파고를 예측하기 위하여 2014년부터 2018년까지의 데이터(41,693개)를 학습자료로 사용하였으며 학습된 딥러닝 모델을 활용하여 2019년도 유의파고를 예측하였다. 딥러닝 모델을 학습시키기 위해 풍속, 풍향, 파향을 입력층 자료로 사용하였으며 입력층 변수의 갯수가 많지 않아 2개의 중간층만 두었다. 첫 번째 중간층에서는 8개의 노드를 설정하고 활성화 함수로 ReLU를 사용하였다. 두 번째 중간층에서는 4개의 노드를 만들었으며 첫 번째 중간층과 같은 ReLU 활성화함수를 이용하였다. 오차함수로는 logcosh를 적용하고 계산속도와 정확성을 높이기 위해 고급 경사하강법으로 아담을 사용하였다.

학습자료를 반복할수록 오차가 줄어들지만 충분한 학습이 진행된 후에는 반복횟수를 증가시키더라도 오차가 일정한 값에 근접하여 수렴하게 된다. 동 연구에서는 학습자료의 중복 학습횟수가 100회에 도달하기 전에 오차가 급격히 줄어들었으며, 175회 이상에서는 오차의 크기가 일정한 값으로 수렴을 하고 있어 학습자료의 반복횟수를 200회로 설정하였다. 2019년도 유의파고 예측 결과에 따르

면 관측치와 예측치가 유사한 변동 패턴을 보이고 있으며, 상관관계 계수(R)는 0.83, 평균 절대 오차(mean absolute error)는 0.283 m을 나타내고 있다. 딥러닝 모델은 전반적인 피크 패턴을 잘 재현하고 있으나 2 m 이상의 파고에 대한 평균 절대 오차는 0.544 m을 나타내고 있어 다소 높은 오차값을 보였다. 이와 같은 결과는 딥러닝 모델이 고파랑에 대해서는 전반적으로 피크값을 과소평가하고 있는 양상을 보이고 있기 때문이다. 특히 최대파고 예측에서는 약 1 m 정도 낮은 파고를 예측하고 있다(Fig. 3). 딥러닝 모델의 구동시간은 Window를 사용하는 PC(intel core(TM) i9-9900K CPU 3.6GHz)에서 1분 45초로 짧은 시간 내에 8,301개의 유의파고 결과를 산출하였다.

3.2 풍속예측

유의파고 예측과 동일하게 풍속 예측도 2014년부터 2018년까지의 데이터(41,693개)를 학습 자료로 사용하였으며 학습된 딥러닝 모델을 적용하여 2019년도 풍속을 예측하였다. 풍속을 예측하기 위한 입력 자료로 유의파고, 파향, 파주기, 풍향을 사용하였으며 2개의 중간층을 설정하였다. 입력변수가 4개로 증가하여 노드를 12개(첫 번째 중간층, 6개(두 번째 중간층)로 정하고 활성화 함수는

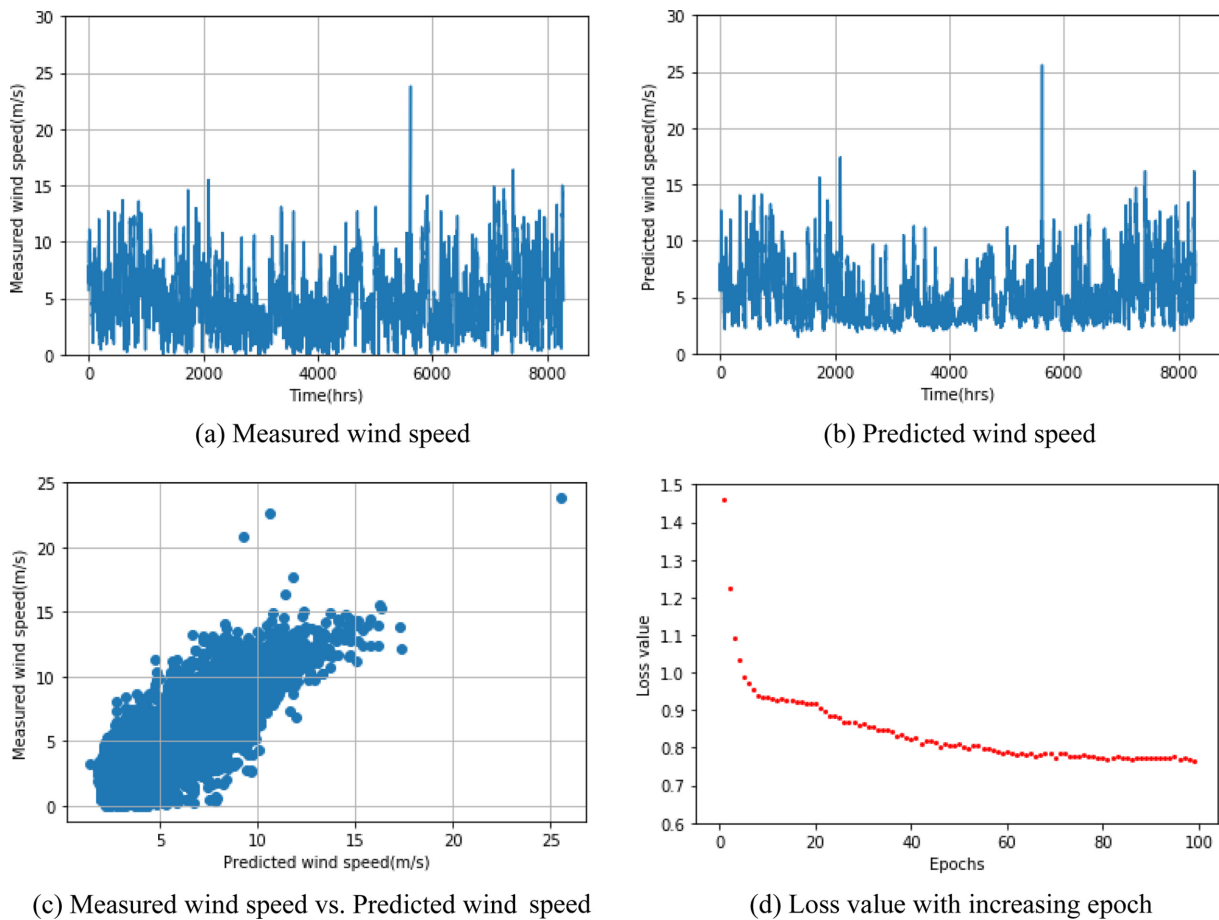


Fig. 4. Prediction results of wind speed (a. measured wind speed, b. predicted wind speed, c. comparison of measured and predicted wind speed d. loss value with increasing epoch).

ReLU를 사용하였다. 오차함수, 고급 경사 하강법은 유의파고 예측과 동일한 logcosh와 아담을 이용하였으며 학습 자료의 반복횟수(Epochs)를 100회로 설정하였다.

풍속 예측 결과에서도 유의 파고 예측 결과와 같이 관측치와 예측치가 유사한 변동 패턴을 보이고 있으며 피크시점을 정확히 재현하고 있다. 상관관계 계수도 0.836으로 유의파고 결과와 근접하며 평균 절대 오차는 1.27 m/s를 보였다. 풍속 예측을 위한 딥러닝 모델의 구동 시간은 약 54초로서 유의 파고 예측 시보다 절반 가량 감소하였으며, 이는 학습 자료의 반복횟수(Epochs)를 절반으로 줄였기 때문이다. 풍속 예측의 경우에는 학습 자료의 반복횟수가 50회를 초과하면 오차가 크게 줄어들지 않아 반복횟수를 100회로 설정하였다(Fig. 4).

3.3 최대파고예측

최대파고 예측에 적용된 딥러닝 모델은 유의파고 예측 시와 동일한 모델(오차함수, 활성화 함수, 고급 경사 하강법, 반복횟수)을 이용하였다. 상이한 부분은 입력 변수가 풍속, 풍향, 유의파고, 평균파고, 파향, 파주기로 늘어나 중간층의 노드 수를 32개, 16개로

증가시킨 점이다. 학습 자료와 예측 자료 또한 동일한 데이터를 사용하였으며 구동시간은 약 1분 40초가 소요되었다.

최대파고의 경우에는 입력 자료가 최대파고와 연관성이 매우 높은 유의파고, 평균파고, 풍속 등을 사용하였기 때문에 딥러닝 모델이 높은 정확도를 구현할 것으로 기대하였다. 모델 결과 역시 매우 우수한 예측 능력을 보여주었으며 관측치와 예측치의 상관관계 계수(R)가 거의 1에 가까운 0.986을 나타내었다(Fig. 5). 평균 절대 오차도 0.12m를 보였으며 특히 2019년에 한 번 기록된 10m 이상의 최대파고를 딥러닝 모델이 예측하였다. 관측된 최대파고는 10.2m였으며 예측된 값은 10.94m로 0.74m의 차이를 보였다.

4. 고찰

파랑은 풍속의 세기, 취송시간, 취송거리와 밀접한 관련이 있으며 딥러닝 기법이 이들 간의 상관관계를 적절히 재현할 수 있을지에 대한 의문이 연구 진행 과정에 있었다. 특히, 딥러닝 기법은 이론에 바탕으로 두고 추정 또는 예측을 진행하지 않고 단순히 방대한 자료를 토대로 상관관계를 유추하기 때문에 이런 의구심이 발

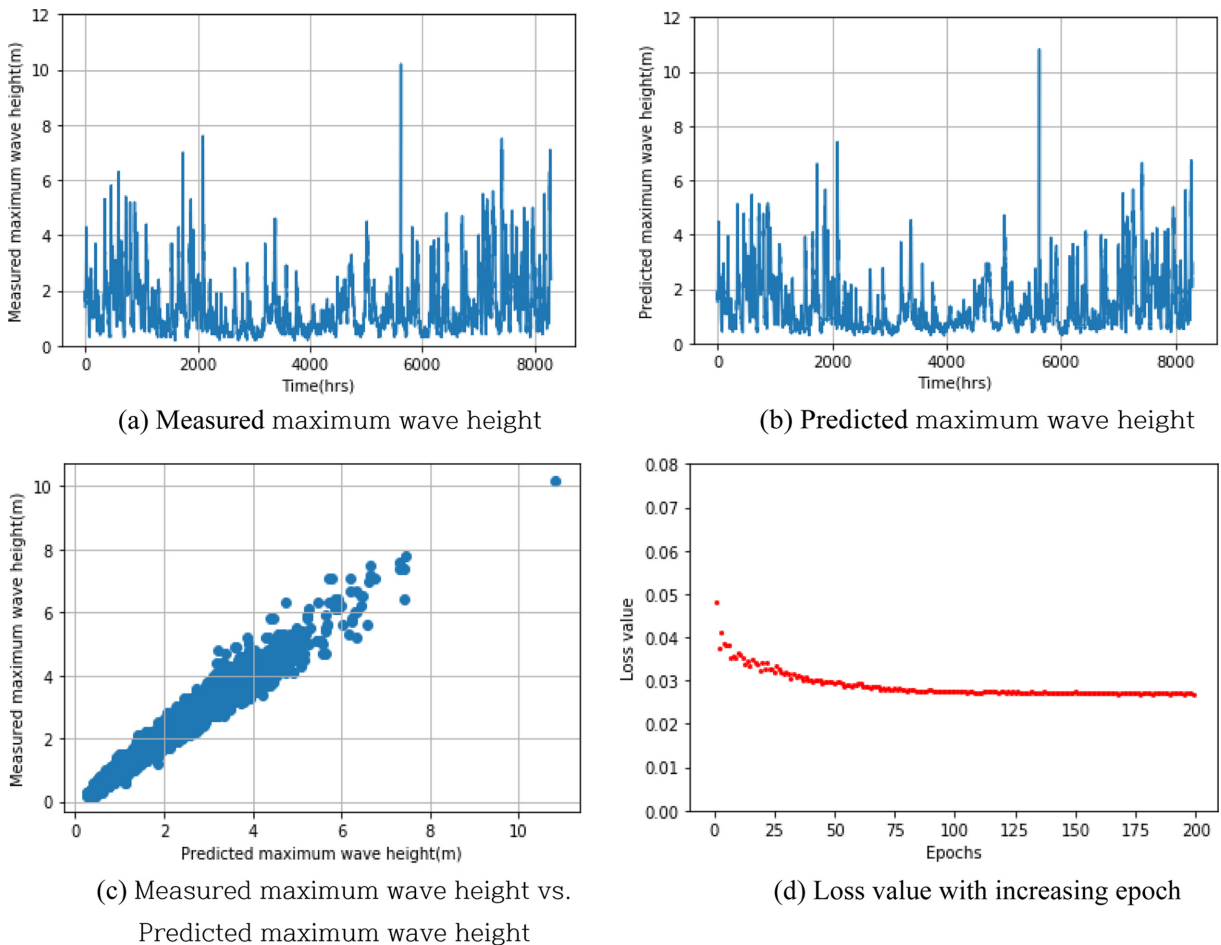


Fig. 5. Prediction results of maximum wave height (a. measured maximum wave height, b. predicted maximum wave height, c. comparison of measured and predicted maximum wave height d. loss value with increasing epoch).

생하였다. 하지만 다양한 연구를 통하여 지속적으로 단점을 보완하고 정확도를 향상시키고 있는 딥러닝 기법은 신뢰성 있는 연구 결과를 도출하였다. 그러나 딥러닝 기법에 적용되는 다양한 함수에 따라 딥러닝 모델의 정확도가 달라지며, 가용한 데이터의 양, 반복 횟수, 노드수에 따라 예측 결과치가 변화하는 특성을 보였다. 그럼에도 불구하고 충분한 데이터의 양, 다양한 함수를 테스트하여 최적의 함수의 적용, 적절한 노드수 및 반복횟수 등의 선정을 통하여 구축된 딥러닝 모델은 일관성 있는 예측 결과치를 제시해 주었다.

또한 연구에 적용된 딥러닝 모델은 모형의 구조를 전혀 바꾸지 않고 단순히 입력 자료만을 변경하더라도 다양한 목표 변수를 예측하였다. 단일 모형 구조를 통하여 다양한 변수를 예측할 수 있다는 것은 딥러닝 기법이 각각의 변수에 대한 상관관계를 충분히 파악하고 일관성 있게 예측하고 있음을 보여준다. 자료가 더 축적되고 딥러닝 기법이 진일보한다면 딥러닝 모델은 필요한 자료에 대한 추정뿐만 아니라 측정된 자료의 검증에도 이용될 수 있을 것으로 예상된다.

학습방법은 시계열적 학습이 아니라 데이터들간의 상호 연관관계를 찾는 방법으로 전체자료를 학습시켜 검증하는 것이 좋은 결과를 도출할 수 있다. 하지만 본 연구에서는 2014-2018년 데이터를 학습데이터로 사용하였으며, 학습데이터로 전혀 사용하지 않은 2019년 데이터를 검증자료로 활용하여 딥러닝 모형의 적용성을 보여주고자 하였다.

5. 결 론

관측된 자료를 토대로 자료들의 상관성을 찾아 예측을 수행하는 딥러닝 기법을 해상부이 자료를 활용하여 파랑-바람 관측 자료의 상호 예측 가능성을 검토하였다. 딥러닝 모델의 학습 자료는 2014년부터 2018년까지의 시간별 데이터(41,693개)를 이용하였으며 검증 자료는 2019년 시간별 데이터(8301개)을 사용하였다.

유의과고를 예측하기 위해서 입력 자료로 풍속, 풍향, 파향을 사용하였으며 관측치와 예측치가 유사한 변동 패턴을 보였고, 상관관계 계수는 0.83, 평균 절대 오차는 0.283 m 이었다. 풍속 예측에는 유의과고, 파향, 파주기, 풍향을 입력 자료로 활용하였으며 딥러닝 모델 결과는 상관관계 계수(0.836), 평균 절대 오차(1.27 m/s)의 오차를 보였다. 최대과고 예측에서는 높은 상관관계 계수(0.986)의 도출과 더불어 10 m 이상의 최대과고를 딥러닝 모델이 구현하였다.

후 기

본 연구는 한국환경정책·평가연구원(KEI)의 지원(딥러닝을 활용한 해양오염 예측도구 개발 및 적용 연구, 과제번호: RE2020-09)을 받아 수행한 연구과제입니다.

References

- [1] Cho, T., 2020, Deep learning for everyone (2nd edition). ISBN 979-11-6521-039-7, 368 p.
- [2] Edwards, C., 2018, Deep Learning Hunts for Signals among the Noise. *Communications of the ACM*, 61(6), 13-14.
- [3] Kang, H. J., 2019, Fish Species Recognition Using Deep Learning, M.S. thesis, Soongsil Univ., Seoul, South Korea.
- [4] Keras, 2020, The Python Deep Learning Library, <https://keras.io/> (accessed 2020.02.10).
- [5] KMA, Marine Meteorological Observations, <https://data.kma.go.kr/data/sea/selectBuoyRltmList.do?pgmNo=52> (accessed 2020.02.01)
- [6] Kim, K.D. and Kim, Y.H., 2018, Improved Correction of Drifter Data using Deep Neural Networks. *J. Korean Inst. Intelligent Sys.*, 28(5), 443-449.
- [7] Kim, Y.B. and Kim, S.M., 2014, Marine Meteorological Characteristics by Comparison of High Wind-Wave Alert and Moored Buoy Data off the Coast of the East Sea between 2006 and 2013. *J. Fisheries Marine Sci. Edu.*, 26(5), 1013-1025.
- [8] Kim, Y.J., Kim, T.W., Yoon, J.S. and Kim, M.K., 2019, Study of the Construction of a Coastal Disaster Prevention System using Deep Learning. *J. Ocean Eng. Technol.*, 33(6), 590-596.
- [9] 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. *J. Korean Soc. Coastal and Ocean Engineers*, 29(3), 129-138.
- [10] Kujira, H., 2017, Python for Machine Learning. ISBN 979-11-5839-067-9, 468 p.
- [11] LeCun, Y., Bengio, Y. and Hinton, G., 2015, Deep Learning. *Nature*, 521, 436-444.
- [12] Lee, H.S., 2010, Comparison and Evaluation of Root Mean Square for Parameter Settings of Spatial Interpolation Method. *J. Asso. Korean Geo.* 13(3), 29-41.
- [13] Lee, C.H., Ahn, S.J., Lee, B.W., Kim, S.W. and Kwon, S.J., 2014, Prediction of Swell-like High Wavers Using Observed Data on the East Coast of Korea. *J. Korean Soc. Coastal and Ocean Engineers*, 26(3), 149-159.
- [14] Choo, T.H., Sim, S.B., Jeon, H.S. and Moon, C.Y., 2019, Development of Tidal Prediction Functions Considering Ocean Meteorological Factors. *J. Korean Soc. Hazard Mitig.*, 19(6), 279-284.
- [15] Sengupta, S., Basak, S., Saikia, P., Paul, S., Tsalavoutis, V., Atiah, F., Ravi, V. and Peters, A., 2019, A Review of Deep Learning with Special Emphasis on Architectures, Applications and Recent Trends. *arXiv:1905.13294*.
- [16] Shrestha, A. and Mahmood, A., 2019, Review of Deep Learning Algorithms and Architectures. *IEEE Access*, 7, 53040-53065.

Received 11 May 2020

1st Revised 7 July 2020, 2nd Revised 28 July 2020

Accepted 29 July 2020