Research Paper

Selection of Input Variables and Comparison of Artificial Neural Networks and One-Dimensional Convolutional Neural Networks for Prediction of Wind Power Generation in Yeongheung Wind Power Plant

Tae-Hui Park¹ · Da-Seul Jang¹ · Gyeong-Min Bae¹ · Kyung-Min Kim² · Johng-Hwa Ahn^{1,2,*}

¹Department of Environmental Engineering, Kangwon National University

²Department of Integrated Energy and Infra System, Kangwon National University

(Received January 12, 2021; Revised March 15, 2021; Accepted March 22, 2021)

Objectives : In this study, deep learning models of artificial neural network (ANN) and one-dimension convolutional neural networks (1D-CNN) were compared to predict nonlinear wind power generation at Yeongheung wind power plant.

Methods : The study site was Yeongheung-do, which has a 46 MW wind power plant. Hourly wind power and meteorological data from January to December 2018 were collected. After pre-processing with standard scaler, the training data were 64%, the validation data were 16%, and the test data were 20%. The optimum input variables of the model were selected using literature, and trial and error method. Rectified linear unit was used as the activation function. Hyperparameters were adjusted by trial and error method to optimized models. To compare the optimized models, the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) were used as the performance efficiency. Both ANN, and 1D-CNN were imported from the Keras library, and all of the performance efficiency was imported from the Scikit-learn library.

Results and Discussion : The optimized input variables in this study were wind speed, wind direction, temperature, and humidity. The optimized ANN performance was $R^2=0.848$, MAE=1.054, and RMSE=1.616, and the hyperparameters were 8 hidden layers with 100 nodes in each layer. The optimized 1D-CNN ($R^2=0.875$, MAE=0.982, and RMSE=1.583) had 4 convolutional layers and the number of filters were 64, 128, 64, and 32 in order from the first layer, and one hidden fully connected layer had 100 nodes. The 1D-CNN had higher R^2 , and lower MAE and RMSE than the ANN. Therefore, the 1D-CNN was selected as the optimized model to predict wind generation of the Yeongheung wind power plant.

Conclusions : The optimized 1D-CNN model in this study was more effective in predicting the Yeongheung wind power plant than the ANN. The optimal input variables were wind speed, wind direction, temperature, and humidity.

Keywords : Artificial Neural Networks, Artificial Intelligence, Convolutional Neural Networks, Machine Learning, Wind Power

The Korean text of this paper can be translated into multiple languages on the website of <http://jksee.or.kr> through Google Translator.

* Corresponding author E-mail: johngghwa@kangwon.ac.kr Tel: 033-250-6357 Fax: 033-259-5550

연구논문

영흥 풍력발전단지의 풍력발전량 예측을 위한 입력변수 선정 및 인공신경망과 1차원 합성곱 신경망 비교

박태희¹ · 장다슬¹ · 배경민¹ · 김경민² · 안종화^{1,2,*}

¹강원대학교 공과대학 환경공학과

²강원대학교 신산업개발 T-EMS 융합학과

목적: 본 연구에선 비선형적 풍력발전량 예측 모델 개발을 목적으로 총 설치 용량 46 MW의 풍력발전단지가 설치된 영흥 풍력발전 단지의 자료를 이용하여 인공신경망(artificial neural network, ANN)과 1차원 합성곱신경망(1-dimension convolutional neural network, 1D-CNN)의 성능을 비교하고자 하였다.

방법: 자료는 46 MW 발전능력을 가진 영흥 풍력발전단지의 2018년 1월부터 12월의 1시간 단위 풍력발전량 자료와 기상청에서 얻은 기상자료를 이용하였다. 최적 입력변수를 선정을 위하여 문헌연구를 바탕으로 시행착오를 거쳐 인자를 선정하였다. 전처리 과정을 거친 17,306개의 자료의 80%를 학습(training), 20%를 테스트(test)로 사용하였으며, 학습 자료의 20%를 검증(validation)자료로 구성하였다. 모델 내 활성화 함수로는 rectified linear unit를 사용하였으며, 시행착오법을 통해 하이퍼파라미터(hyperparameter)의 최적값을 도출하였다. 모든 모델은 Python의 Keras 라이브러리를 이용하여 설계하였으며, 성능지표인 결정계수(coefficient of determination, R^2), 평균제곱근오차(root mean square error, RMSE), 평균절대오차(mean absolute error, MAE) 등은 Scikit-learn 라이브러리에서 이용하였다.

결과 및 토의: 최적 입력변수는 풍속, 풍향, 온도, 습도 등이었다. ANN의 최적점으로는 은닉층 8층, 은닉층별 노드 수는 모두 100으로 나왔다. 최적 ANN 모델에서 성능지표는 $R^2=0.848$, MAE=1.054, RMSE=1.616이었다. 1D-CNN의 최적점으로는 합성곱층 4층, 층별 필터 수는 1층부터 64, 128, 64, 32개, 전결합층 1층에 노드 100개이다. 최적 1D-CNN 모델의 성능지표는 $R^2=0.875$, MAE=0.982, RMSE=1.583였다. 1D-CNN이 ANN보다 R^2 는 높고, MAE와 RMSE는 낮았다. ANN, 1D-CNN의 결정계수가 모두 0.8 이상으로 예측 성능이 우수하나, 1D-CNN이 ANN보다 모든 성능지표에서 높았다.

결론: 최적화된 모델의 성능지표 비교 결과 1D-CNN이 ANN보다 높은 성능을 보여 영흥 풍력 발전소 발전량 예측에 적합할 것으로 보인다. 최적 입력변수는 풍속, 풍향, 온도, 습도였다.

주제어: 인공신경망, 인공지능, 기계학습, 합성곱신경망, 풍력에너지

1. 서론

전 세계적으로 산업 및 농업 활동의 증가로 에너지 수요가 지속적으로 증가하고 있다.¹⁾ 이에 따라 환경오염도 심각해지고 있다. 2015년 12월 파리 기후 정상회의가 열린 이후 많은 나라가 온실가스 배출량을 줄이는 데 큰 노력을 기울이고 있다.²⁾ 이 문제를 해결하기 위해 고갈되는 에너지원보다 환경을 보호하고 지속 가능한 개발을 이룰 수 있는 에너지원인 신재생에너지가 주목을 받고 있다. 우리나라는 2017년 기준, 신재생에너지가 전체 발전량의 5.6%를 차지하여 주요국에 비해

낮은 편이다. 따라서 정부에서는 2030년까지 신재생에너지 발전량 비중을 20%까지 높이고, 신규설비의 95% 이상을 태양광, 풍력 등으로 공급하고자 하는 ‘재생에너지 3020 이행계획’을 발표했다.³⁾ 비슷한 사례로 2030년까지 독일은 전체 발전량의 65%, 스페인은 74%를 신재생에너지로 공급하겠다는 목표를 세웠다.^{4,5)}

풍력에너지는 근래 들어 가장 빠르게 성장하는 에너지원 중 하나이며, 화석연료 및 다른 신재생에너지에 비해 설치면적이 작고 발전단가가 저렴하여 여러 나라에서 많은 관심을 기울이고 있다.^{6,7)} 또한 풍력에너지는 온실가스 배출량을 줄일 수 있

는 잠재력을 가진 효율적인 신재생 에너지원이다.²⁾ 풍력에너지는 바람이 풍차의 날개를 돌리면 전기 발생 장치인 발전기에 연결되어 전기를 생산하고 생산된 전기는 바로 사용하거나 곳곳으로 보내 사용되도록 한다. 우리나라는 바람이 많이 부는 산, 긴 해안선, 섬 등에 풍력발전 설치 시 많은 전기를 생산할 수 있는 유리한 조건을 가지고 있다. 그러나 풍력에너지는 에너지 생성의 변동 등 예측하기 어려운 가변적인 성질을 띠어 전력 시스템에 많은 문제를 일으킨다.⁸⁾ 따라서 풍력발전량 예측은 전력 및 에너지 시스템의 계획, 관리, 운영 등에 필수적이다. 자원을 효율적으로 활용하기 위해서는 계획 수립 시 풍력발전 능력을 고려해야 한다.⁹⁾ 정확한 풍력발전 예측은 전력회사가 풍력발전 능력을 활용하고 비용을 일부 절감할 수 있다. 풍력발전 능력에 대한 신뢰성은 풍력발전량 예측 모델의 정확도에 달려 있는데, 풍력발전량 자료에 내재한 비선형 특성을 발견할 수 있는 인공지능 기술은 기계 학습이다.^{9,10)} 기계학습은 컴퓨터 알고리즘의 학습을 통해 자동으로 모델 성능을 개선하는 기술이다. 특히 비선형적 자료 예측에 있어 기존 수학적 모델에 비해 뛰어난 성능을 보여 환경 및 에너지 분야에 적용하기 위한 시도가 진행되고 있다. 이러한 사례로 인공신경망(artificial neural network, ANN), 서포트 벡터 머신(support vector machine, SVM) 등을 이용한 영산강, 주암호 조류예측¹¹⁾, ANN, SVM 등을 이용한 태양광 발전량 예측¹²⁾, 오차전파(back propagation) 신경망을 이용하여 지열발전 예측¹³⁾, SVM 등을 이용한 미생물 연료전지(microbial fuel cell) 성능 평가¹⁴⁾ 등이 있다. 이러한 인공지능을 이용한 다양한 인자의 예측 시도는 풍력 발전에서도 적용되어 널리 사용되고 있다.¹⁵⁾

기계학습을 이용하여 풍력발전량을 예측한 선행 연구에선 ANN, 합성곱 신경망(convolutional neural network, CNN), 서포트 벡터 머신(support vector machine, SVM), 자동회귀 누적 이동평균(autoregressive integrated moving average, ARIMA) 등의 모델이 사용되고 있다. 이 중 ANN을 사용한 빈도수가 가장 많았다.¹⁶⁻²¹⁾ CNN 모델 중 이미지 같은 2차원 자료를 다루는 2차원 합성곱 신경망(2-dimensional convolutional network)이 가장 널리 사용되고 있다. 이를 시계열 자료와 같은 1차원 자료에서 활용할 수 있게 변형한 것이 1차원 합성곱 신경망(1-dimensional convolutional neural network, 1D-CNN)이다.²²⁾ 1D-CNN을 사용하여 풍력발전량을 예측한 연구는 찾을 수 없었다. 또한, 인공신경망 기법에서 입력변수를 결정할 때 관련된 모든 입력변수를 선택하는 것보다 적절한 입력변수를 선택하는 것이 더 중요하다.²³⁾ 인천광역시 영흥도 지역에 풍력발전단지가 있고 발전량 자료를 공개하고 있으나 풍력발전량을 예측한 선행 연구는 찾아볼 수 없었다. 따라서 영흥 풍력발전단지의 풍력발전량 예측 성능을 향상하기 위해 최적 입력변수를 선정하고 ANN과 1D-CNN을 비교하고자 하였다.

Fig. 1. Location of Yeongheung wind-power plant.

2. 실험방법

2.1. 자료

연구지역은 대한민국 인천광역시 옹진군 영흥도에 위치한 영흥 풍력발전단지로 해당 단지의 풍력발전량 자료를 사용하였다. 영흥 풍력발전단지는 1단지(9대, 총 22 MW)와 2단지(8대, 총 24 MW)로 구분되어 있으며, 1단지는 서쪽에 9대, 2단지는 서쪽에 4대, 남쪽에 4대가 있다(Fig. 1). 본 연구는 1단지와 2단지 풍력발전량 자료를 통합하여 모델에 사용하였다. 대한민국 공공데이터포털(www.data.go.kr)에서 한국남동발전 영흥 풍력발전단지의 1년(2018.1.1.~2018.12.31.)간의 시간별 발전량 자료를 수집하였다. 기상청(weather.go.kr)에서 동일 기간의 영흥면 기상인자 자료를 수집하였다. 기상인자 중 풍력발전 예측에 풍속(wind speed, WS), 풍향(wind direction, WD), 습도(humidity, H), 온도(temperature, T), 압력(pressure, P) 등을 사용하였다. 자료 중 수치가 기재가 되어있지 않은 자료는 소거했으며 그 결과 약 17,306개의 자료를 시계열 순서대로 대입하여 사용하였다(Table 1). 이상치에 대해선 제외하지 않고 적용하였다. 자료 중 온도 및 풍향 등 상호간 편차가 큰 자료가 존재하여 전체 자료의 평균을 0으로, 표준편차를 1로 조절하는 표준화 작업을 진행하였다. 작업은 Scikit-learn 라이브러리의 `standardscaler` 함수를 활용하여 진행하였다.²⁴⁾ 최적 인자 선정은 기상인자 등을 모든 경우로 조합하여 총 15가지의 시나리오(scenario)를 ANN에 적용하여 선정하였다(Table 2).

2.2. 모델

2.2.1. 인공신경망

ANN은 인간의 뇌가 거치는 학습과정을 모방하여 정보를 처리하는 기법으로 뉴런을 통해 상호연결적인 예측을 이루어

Table 1. Statistics of power generation and meteorological data collected from January to December 2020 in the study site.

Variable	Average	Standard deviation	Minimum	Median	Maximum
Temperature (°C)	12.48	10.19	-14.10	13.10	35.40
Wind degree (°)	184.79	117.80	0	210.30	360
Wind speed (m/s)	3.64	2.28	0	3.20	18.90
Humidity (%)	79.48	19.72	20.50	87.10	98.80
Generation (MW)	2.31	3.55	0	0.49	19.22

Table 2. Result of artificial neural network model performance of various input scenarios in this study. All of the scenarios were used time-series data, and applied with data.

	Scenario	Train MAE (MW)	Test MAE (MW)	Train R ²	Test R ²	Train RMSE (MW)	Test RMSE (MW)
1	WS ^{a)} + WD ^{b)} + H ^{c)} + T ^{d)} + P ^{e)}	0.319	1.123	0.930	0.823	0.812	1.750
2	WS + WD + T + P	0.317	1.091	0.937	0.843	0.766	1.667
3	WS + WD + H + T	0.310	1.070	0.943	0.845	0.760	1.658
4	WS + WD + H + P	0.326	1.096	0.942	0.841	0.782	1.662
5	WS + H + T + P	0.309	1.100	0.936	0.843	0.793	1.686
6	WS + WD + H	0.328	1.070	0.940	0.837	0.788	1.671
7	WS + WD + P	0.314	1.083	0.943	0.830	0.766	1.718
8	WS + WD + T	0.319	1.120	0.933	0.834	0.808	1.668
9	WS + H + P	0.324	1.077	0.943	0.837	0.772	1.674
10	WS + H + T	0.333	1.077	0.939	0.844	0.783	1.671
11	WS + T + P	0.355	1.080	0.928	0.837	0.754	1.680
12	WS + WD	0.367	1.137	0.919	0.839	0.779	1.680
13	WS + H	0.352	1.090	0.925	0.836	0.791	1.680
14	WS + P	0.369	1.100	0.941	0.843	0.766	1.660
15	WS + T	0.344	1.107	0.944	0.842	0.769	1.667

^{a)} WS : Wind speed, ^{b)} WD : Wind Direction, ^{c)} H : Humidity, ^{d)} T : Temperature, ^{e)} P : Pressure

낸다.²⁵⁾ 반복적인 학습 과정을 거쳐 패턴을 찾아내고, 이를 일반화함으로써 예측하고자 하는 문제에 있어 유용하다. ANN은 모델 방정식을 명시적으로 정의할 필요 없이 변수 간의 상관관계를 추출하는 능력이 있다.⁹⁾ ANN은 다양한 입력 자료를 받는 입력층(input layer)과 이를 연산하기 위한 최소 1개 이상의 은닉층(hidden layer), 마지막으로 원하는 출력 자료의 예측값을 출력하는 출력층(output layer)으로 이루어져 있다(Fig. 2(a)). 연속적인 시간에 걸쳐 측정되는 시계열(time series) 자료로 훈련된 ANN은 임의의 선형 및 비선형 함수를 모델링 할 수 있는 능력을 갖추고 있어, 다양한 분야에서 폭넓게 활용되고 있다.²⁶⁾ 본 연구에서는 시계열로 구성된 비선형적인 풍력발전량 및 기상자료를 사용하여 풍력발전량 예측을 위해 다수의 은닉층과 노드로 구성되는 ANN 모델을 적용하였다. 입력 자료로는 기상청 자료를 활용하였으며, 예측 목표인 발전량을 출력 자료로 하였다. 모델의 과적합 방지를 위해 dropout과 학습 초기 종료(earlystop) 기법을 사용하였으며, 활성화함수(activation function)로는 rectified linear unit (ReLU)와 sigmoid를 적용하였다.

2.2.2. 1차원 합성곱신경망

최근 CNN은 전력 품질, 풍력발전단지의 공간적 특성 등 에너지 관련 문제를 해결하기 위해 사용되었다.²⁷⁻²⁹⁾ 또한, CNN을 통해 풍력 시계열 자료의 숨겨진 특징을 추출하는 것은 예측 성능을 향상하게 할 수 있는 방법이다.³⁰⁾ 따라서 본 연구에선 모델을 단순화하고 풍력 예측의 실시간 구현을 위해 1D 풍력 자료를 직접 예측할 수 있는 1D-CNN을 적용하였다.³¹⁾ CNN에서 풀링층(pooling layer)은 일반적으로 연속적인 합성곱층(convolution layer) 사이에 삽입된다(Fig. 2(b)).²⁷⁾ 본 연구에선 계산 복잡성을 줄이는 연산 기법인 max pooling을 사용하였다.³²⁾ 이후, 평탄화층(flatten layer)을 통해 자료를 배열 형태로 만들고, 전결합층(fully connected layer)에서 예측을 위한 연산 과정을 거친다. 모델의 과적합 방지를 위해 drop out과 early-stop 기법을 사용하고, 활성화 함수로는 ReLU와 sigmoid를 적용하였다. 입력, 출력 자료는 ANN과 동일하게 적용하였다.

2.3. 최적화 및 성능지표

수집한 자료를 총 15개의 조합으로 ANN을 이용하여 최적

Fig. 2. Architecture of artificial neural network and one-dimension convolutional neural network models.

자료를 선정한 다음 모델별 하이퍼파라미터를 시행착오법 (trial & error method)을 이용하여 최적화하였다(Fig. 3). 이후 최적화된 ANN과 CNN의 성능지표를 비교하여 영흥도 풍력발전량 예측에 적합한 모델을 선정하였다. 성능지표는 결정계수 (coefficient of determination, R^2), 평균절대오차(mean absolute error, MAE), 평균제곱근오차(root mean square error, RMSE)를 사용하였다. R^2 는 관측값으로부터 실험값의 분산을 보여준다(Eq. 1). 최상의 값은 1.0이며, 음수가 출력되는 경우 모델의 예측값과 실제 관측값이 어떠한 관계도 가지지 못하여 예측에 실패했음을 나타낸다.³³⁾

$$R^2 = 1 - \frac{\sum_{i=1}^n [P_i - O_i]^2}{\sum_{i=1}^n \left[O_i - \frac{1}{n} \sum_{i=1}^n O_i \right]^2} \quad (1)$$

여기서, n 은 자료의 수, O_i 는 관측값, P_i 는 실험값을 뜻한다.

MAE는 관측값과 실험값 간의 절대 오차의 평균을 보여준다(Eq. 2). 최상의 값은 0이며, 값이 상승할수록 성능이 좋지 않음을 의미한다.³⁴⁾

$$MAE = \frac{1}{n} \sum_{i=1}^n (|O_i - P_i|) \quad (2)$$

RMSE는 관측값과 실험값 간의 오차를 제곱한 뒤 평균을 내어 이를 제곱근한 값이다(Eq. 3). 최상의 값은 0이며, 값이 상승할수록 좋지 않은 성능이 나타난다.³³⁾

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (3)$$

Fig.3. Flowchart of modeling process to predict power generation of Yeongheung wind power plant.

모델링은 Python 환경에서 수행하였으며, ANN, CNN 모델은 Keras 라이브러리에서 이용하였다. R^2 , MAE, RMSE 등 성능지표는 Scikit-learn 라이브러리에서 이용하였다.

3. 결과 및 고찰

3.1. 입력변수

본 연구에서 풍력발전량 예측에 적절한 입력변수를 선정하기 위해 고려된 기상인자는 풍속, 풍향, 습도, 온도, 압력 등이었다. 풍력발전량은 이론적으로 동력전달장치 효율, 발전기 에너지변환 효율, 회전자 동력계수, 공기밀도, 블레이드 회전면적 등에 비례하고, 풍속의 세제곱에 비례한다.³⁵⁾ 즉, 풍속은 다른 변수에 비해 풍력발전에 더 큰 영향을 미치며 높은 상관관계를 갖는 것을 알 수 있다.⁹⁾ 따라서 풍속은 입력변수로 항상 사용하였다.

인자 최적화 결과 모든 인자를 조합한 경우(Scenario 1)에 비해 4개의 인자를 조합한 경우(Scenario 2-5)의 성능지표가 좋게 나왔는데, 이 중 Scenario 3이 상대적으로 우수하였다(Table 2). 즉 풍속, 풍향, 온도, 습도의 기상인자를 조합한 경우가 Train $R^2=0.943$, Test $R^2=0.845$, Train MAE=0.310, Test MAE=1.070, Train RMSE=0.760, Test RMSE=1.658로 다른 경우에 비해 좋은 결과를 보였다. 하지만, 기상인자가 3개 이하일 땐 Train과 Test의 MAE, RMSE가 높아지며 R^2 가 낮아지는 경향을 보였다(Scenario 6-15).

가장 우수한 지표를 보인 Scenario 3에서 압력 인자가 제외되었다. 압력을 포함하여 4개의 인자를 조합한 경우(Scenario

2-5)와 3개의 인자를 조합한 경우(Scenario 6-11)에서 RMSE가 높아진 경향을 보였다. 특히 Scenario 7에서 Test RMSE 1.718로 가장 높았다. 이와 다르게 2개의 인자를 조합한 경우(Scenario 12-15)에선 압력이 포함된 Scenario 14의 RMSE가 상대적으로 우수했다. 하지만 Scenario 3에 비해 낮은 성능지표를 보였다. 따라서 본 연구에서는 최적의 입력변수로, 압력을 제외한 풍속, 풍향, 온도, 습도의 4개 인자(Scenario 3)만 이용하였다. 선행 연구에서도 상대적으로 많이 사용된 입력변수는 풍속, 풍향, 온도, 습도이고, 압력은 그 빈도가 낮았다.³⁶⁾

3.2. 인공신경망

ANN에서 초기 훈련 중에 과적합 여부를 판단하기 위해 자료 사용 비율을 학습(training) 64%, 검증(validation) 16%, 확인(test) 20%로 하였다.³⁷⁾ 최적 결과를 도출하기 위해 은닉층 수, 은닉층별 노드(node) 수 등 하이퍼파라미터(hyperparameter)를 조정하였다. 활성화 함수는 sigmoid보다 ReLU를 사용했을 때 더 좋은 성능지표를 보였다. 은닉층의 경우 단일층보다 복수층에서 정확도가 높았다.³⁸⁾ 은닉층 수를 8층으로 하였을 때, 은닉층 수 7층과 9층으로 설정했을 때보다 Test MAE가 0.08, 0.12 만큼 감소했으며 Test RMSE가 0.069, 0.053 만큼 감소하였다. 은닉층 수를 8층으로 고정하고 은닉층별 노드를 90, 110으로 설정한 경우, 100으로 하였을 때 Test MAE가 0.23, 0.12 만큼 감소했으며 Test RMSE가 0.088, 0.113 만큼 감소하였다. 이에 은닉층 수를 8층, 은닉층별 노드수를 100개로 설정하였다. Drop out은 과적합을 줄여 신경망을 개선하는 기술로, 매우 다양한 영역에서 ANN의 성능을 향상하기 위해 사용된다.³⁹⁾ 본 연구에서 drop out을 적용하지 않은 경우의 성능지표가 우수하여 ANN에 drop out을 적용하지 않았다. Drop out을 사용했을 시, drop out이 없는 경우보다 훈련 길이가 2~3배 늘어나³⁷⁾ 성능지표가 좋지 않은 경향을 보였다. 학습 초기 종료를 적용한 경우 오차(loss)값이 10회 동안 개선되지 않을 때 모델 학습을 초기 종료하도록 적용하였다.⁴⁰⁾ 가장 좋은 성능을 나타낸 ANN 모델은 학습 초기 종료를 적용하여 epoch 30에서 최적화가 되었다. 학습 초기 종료를 적용한 경우, 적용하지 않았을 때보다 Test MAE가 0.007 감소했으며 Test RMSE는 0.02 감소했다. Batch size는 500으로 설정하였을 때가 성능지표가 최적 결과를 보였다. 따라서 최적 결과는 자료 사용 비율을 학습 64%, 검증 16%, 확인 20%이며, 활성화 함수 ReLU를 사용해 은닉층 수 8층, 은닉층별 노드 수 100개, drop out은 제외, batch size 500, epoch 36에서 최적점을 보였다. ANN의 최적 결과는 Train $R^2=0.947$, Test $R^2=0.848$, Train MAE=0.291, Test MAE=1.054, Train RMSE=0.750, Test RMSE=1.616이었다(Table 3).

ANN을 이용하여 터키의 한 풍력발전단지의 발전량을 예측한 연구에선 풍력, 풍속, 온도를 입력변수로 사용하였고, Test MAE=46.5120 KW, Test RMSE=61.3950 KW로 보고하여⁴¹⁾

Table 3. Performance of optimized model of artificial neural network and one-dimension convolutional neural network.

	Train MAE (MW)	Test MAE (MW)	Train R^2	Test R^2	Train RMSE (MW)	Test RMSE (MW)
ANN	0.291	1.054	0.947	0.848	0.750	1.616
CNN	0.215	0.982	0.961	0.875	0.537	1.583

Table 4. Literature review for wind power forecasting with artificial neural network and convolutional neural network.

Model	Literature	Region	Input	Outcome
ANN	Jiao et al. (2018) ¹⁸⁾	Homepage of EirGrid	-	Test RMSE 454.17 MW MAE 346.26 MW
ANN	Veena et al. (2020) ¹⁹⁾	-	NWP, SCADA	Test MAE 142.49 KW RMSE 216.52 KW
ANN	Nielson and Bhaganagar (2020) ²⁰⁾	Kirkwood, Iowa, USA	Wind speed, density, richardson number, turbulence intensity, wind shear	Average Test MAE 30.45 KW
ANN	Acikgoz et al. (2020) ²¹⁾	Turkey	Wind speed, wind direction	Test R^2 Spring 0.916, summer 0.929, autumn 0.899, winter 0.884
ANN	Hu et al. (2020) ³⁷⁾	Southeastern Australia	-	Test MAE 2.43 Test RMSE 3.75
ANN	Demirdelen (2019) ⁴¹⁾	Belen Wind Energy Power Plant in Hatay, Turkey	Atmospheric temperature and wind speed, wind power	Test MAE 46.5120 KW, RMSE 61.3950 KW
ANN	Carolin Mabel and Fernandez (2008) ⁴²⁾	Muppandal, Tamil Nadu, India	wind speed, relative humidity and generation hours	MAE Train 0.0588, Test 0.0586
ANN	Manobel et al. (2018) ⁴³⁾	Totoral wind farm	Wind speed, wind direction	Test RMSE 45.50 KW
ANN	Shi et al. (2012) ⁴⁴⁾	-	Wind speed, wind power	Test MAE 113.93
CNN	Mujeeb et al. (2019) ¹⁷⁾	New England's wind farm	Wind speed, wind direction, temperature, humidity, pressure	Test MAE Spring 3.34, summer 3.24, autumn 3.36, winter 2.89
CNN	Zhao and Jia (2020) ⁴⁷⁾	China	Wind speed, wind power, active power, theoretical power curve	Test R^2 0.961

(Table 4) 본 연구와 인자 및 성능에서의 차이를 보였다. 인도 무판달에 위치한 풍력 발전기의 발전량을 예측한 연구에선 풍속, 습도, 발전 시간을 입력변수로 사용하였으며, Train MAE=0.0588, Test MAE=0.0586이었다.⁴²⁾ 칠레 토토리알 풍력 단지의 풍력발전량을 예측한 논문에서는 풍속과 풍향을 입력 변수로 사용하였고, Test RMSE=45.50 KW였으며⁴³⁾, 풍속과 풍력을 입력변수로 사용하여 풍력발전량을 예측한 논문에서는 Test MAE=113.93로 나타났다.⁴⁴⁾ 본 연구를 포함한 대부분의 선행 연구에선 ANN 모델링 결과 풍속 인자가 적용되었으며, 나머지 인자는 지역에 따라 다르게 적용됨을 알 수 있었다. 선행 연구에서도 풍력발전은 지역, 기후마다 특성이 다를 수 있어¹⁰⁾, 효율적인 ANN의 예측모델 개발을 위해선 대상지역의 기후 등을 고려하여 적절한 변수의 선택이 중요하다는 것을 알 수 있었다.⁴⁵⁾

3.3. 1차원 합성곱신경망

CNN의 경우 변수인 합성곱층 수와 풀링층(pooling layer)이 결과에 영향을 주기 때문에 적절한 조정이 필요하다.^{27,46)} 이에 합성곱층과 필터 수를 여러 경우를 고려하여 조정했을 때, 합성곱층 4층, 필터(filter) 수 64, 128, 64, 32개에서 최적 결과를 얻었으며, 활성화 함수는 sigmoid보다 ReLU를 사용했을 때 최적 결과를 보였다. Test MAE 성능지표가 1.087로 제일 좋은 값을 보였다. 최적화 결과와 합성곱층 3층에 필터 수 48, 48, 16을 적용한 경우와 합성곱층 5층에 필터 수 64, 128, 64, 32, 16개인 경우를 비교했을 때, Test MAE가 0.02, 0.048 만큼 감소했으며 Test RMSE는 0.049, 0.074 만큼 감소하였다. 계산 복잡성뿐만 아니라 학습 가능한 변수의 수를 크게 줄일 수 있는 max pooling을 사용했다.³⁷⁾ Drop out은 적용하지 않았는데, 성능지표에 급격한 변화가 없고 R^2 의 성능지표가 적용하

지 않은 경우에서 좋은 경향을 보였기 때문이다. 학습 초기 종료를 적용한 경우가 적용하지 않았을 때보다 Test RMSE가 0.016 상승하였다. 학습 초기 종료를 적용하지 않고 batch size를 300으로 고정한 채로 epochs를 높여가며 실험을 진행했다. Epochs를 높일수록 Test R^2 가 상승하는 경향을 보였고, epochs 400으로 설정하였을 때 Test R^2 이 최고점을 보였다. Batch size는 300으로 설정하고, 완전결합층 수 1층, 노드 수 100개로 조정했을 때, 성능지표의 최적 결과를 얻었다. 따라서 최종 결과는 자료 사용 비율은 학습 64%, 검증 16%, 확인 20%로 ANN과 같게 하였고, 활성화 함수는 ReLU, 합성곱층 4층, 층별 필터 수는 순서대로 64, 128, 64, 32개, 전결합층 1개, 노드 수 100개, drop out은 제외, batch size 300, early-stop 제외한 경우로 epoch가 400에서 최적점을 보였다. 이때, Train $R^2=0.961$, Test $R^2=0.875$, Train MAE=0.215, Test MAE=0.982, Train=RMSE 0.537, Test RMSE=1.583로 나타났다(Fig. 3,

Table 3).

CNN을 이용하여 뉴잉글랜드에 위치한 풍력 발전소 자료로 풍력발전량을 예측한 논문에서는 합성곱층 3층, 전결합층 1층을 사용했다. Test MAE 봄=3.34, 여름=3.24, 가을=3.36, 겨울=2.89로 계절마다 기간을 나누어 결과를 비교했다(Table 4).¹⁷⁾ 중국에 위치한 풍력 발전소 자료를 사용하여 풍력발전량을 예측한 논문에서는 합성곱층 3층, 전결합층 1층을 사용했고, Test R^2 0.961로 나타났다.⁴⁷⁾ 본 연구에서는 합성곱층 4층, 전결합층 1층을 사용했다. CNN의 하이퍼파라미터를 조정된 결과, CNN의 효율적인 예측 모델을 위해서 지역과 상황에 맞는 하이퍼파라미터 조정이 중요한 것을 알 수 있었다.³⁷⁾

3.4. 모델 성능 비교

두 모델 모두 학습과 확인과정에서 관측값과 실험값의 상관관계가 0.8 이상으로 비교적 높았다(Fig. 4, Fig. 5). R^2 의 경우

Fig. 4. Prediction and observation data of wind power distribution with artificial neural network and one-dimension convolutional neural network.

Fig. 5. Correlation of prediction and observation data of wind power with artificial neural network and one-dimension convolutional neural network.

1D-CNN은 0.875, ANN은 0.848로, 1D-CNN이 ANN보다 0.027 높았다. ANN의 RMSE는 1.616, 1D-CNN의 RMSE는 1.583으로 1D-CNN이 0.033 낮은 결과를 보였다. RMSE는 관측값과 실험값의 오차를 나타내는 지표로, 1D-CNN의 풍력발전량 예측 정확도가 더 높음을 알 수 있다. ANN과 달리 1D-CNN은 가중치가 공유되기 때문에 학습 효율이 향상한 것으로 추정된다.¹⁷⁾ 따라서 ANN과 1D-CNN 모두 영흥도 지역 풍력발전량 예측에 적합하지만, 모델 성능을 비교해보았을 때 본 연구에서도 1D-CNN의 성능이 더 좋음을 알 수 있다.

4. 결론

본 연구에서는 효율적인 풍력 발전량 예측 모델 개발을 위하여 ANN과 1D-CNN을 이용하여 영흥도 풍력발전단지의 발전량 예측을 시도하였다. 풍속, 풍향, 온도, 습도를 입력변수로 사용한 모델에서 가장 좋은 성능지표를 보였다. ANN의 경우, 활성화 함수로 ReLU, 은닉층 8개, 은닉층별 노드 수는 모두 100개, early-stop은 10회, drop out은 적용하지 않았을 때 성능지표가 가장 우수하였다. 1D-CNN의 경우, 활성화 함수는 ReLU, 합성곱층 4층, 층별 필터 수는 64, 128, 64, 32개, 완전결합층 1개, 노드 수 100개, drop out 제외, early-stop 비

설정의 조건에서 가장 좋은 결과를 얻었다. 이러한 조건에서 1D-CNN은 ANN보다 모든 성능지표에서 우수하여 풍력발전량을 보다 효과적으로 예측할 수 있어 1D-CNN을 이용한 실제 풍력발전량 예측 모델 개발 가능성을 확인할 수 있었다. 또한, 본 연구의 입력 인자와 선행연구에서 사용인 인자들을 분석한 결과 풍력발전량은 지역마다 기후, 계절, 지형 등에 따라 사용하는 입력 인자가 달라지는 등 지역에 따른 특성을 보이므로, 이를 고려한 모델 개발이 필요함을 알 수 있었다. 따라서 향후 다양한 지역의 풍력발전량 예측을 통해 지역별 입력 인자의 차이의 원인을 분석할 필요가 있다.

Acknowledgement

본 논문은 4단계 BK21 사업(강원대학교, 다학제 융합 에너지지원 신산업 핵심인력 양성사업단)으로 지원된 연구입니다.

References

1. R. P. Shetty, A. Sathyabhama, P. S. Pai, Comparison of modeling methods for wind power prediction: a critical study, *Front. Energy*, 14, 347-358(2020).

2. Z. Wang, W. Wang, B. Wang, Regional wind power forecasting model with NWP grid data optimized, *Front. Energy*, 11, 175-183(2017).
3. Ministry of Trade, Industry and Energy, Implementation plan of renewable energy 3020, Ministry of Trade, Industry and Energy, Sejong, Korea, pp. 13(2017).
4. Korea Energy Economics Institute, World energy market insight, Korea Energy Economics Institute, 20(20), Ulsan, Korea, pp. 28(2020).
5. Korea Trade-Investment Promotion Agency, Analysis of energy policy of Spain and cooperation plan with Korea, Korea Trade-Investment Promotion Agency, Seoul, Korea, pp. 19(2019).
6. D. Fang, J. Wang, A novel application of artificial neural network for wind speed estimation, *Int. J. Sustain. Energy*, 36(5), 415-429(2017).
7. Korea Institute for Industrial Economics and Trade, Basic industries of the wind power plant industry, Korea Institute for Industrial Economics and Trade, Sejong, Korea, pp. 59 (2020).
8. L. Frias-Paredes, F. Mallor, M. Gastóon-Romeo, T. León, Assessing energy forecasting inaccuracy by simultaneously considering temporal and absolute errors, *Energy Conv. Manag.*, 142, 533-546(2017).
9. K. Methaprayoon, C. Yingvivanapong, W. J. Lee, J. R. Liao, An Integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty, in *Proceedings of IEEE Systems Technical Conference on Industrial and Commercial Power*, IEEE, NY, USA, pp. 116-124(2005).
10. H. Wang, Z. Lei, X. Zhang, B. Zhou, J. Peng, A review of deep learning for renewable energy forecasting, *Energy Conv. Manag.*, 198, 111799(2019).
11. Y. G. Park, G. H. Cho, J. H. Park, M. S. Cha, J. H. Kim, Development of early-warning protocol for predicting chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, *Korea, Sci. Total Environ.*, 502, 31-41(2015).
12. U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. V. Deventer, B. Horan, A. Stojcevski, Forecasting of photovoltaic power generation and model optimization: a review, *Renew. Sust. Energ. Rev.*, 81, 912-928(2018).
13. H. Esen, M. Inalli, ANN and ANFIS models for performance evaluation of a vertical ground source heat pump system, *Expert Syst. Appl.*, 37(12), 8134-8147(2010).
14. A. Garg, V. Vijayaraghavan, S. S. Mahapatra, K. Tai, C. H. Wong, Performance evaluation of microbial fuel cell by artificial intelligence methods, *Expert Syst. Appl.*, 41(4), 1389-1399(2014).
15. F. Shahid, A. Khan, A. Zameer, J. Arshad, K. Safdar, Wind power prediction using a three stage genetic ensemble and auxiliary predictor, *Appl. Soft Comput.*, 90, 106151(2020).
16. S. Kr. Jha, J. Bilalovic, A. Jha, N. Patel, H. Zhang, Renewable energy: present research and future scope of artificial intelligence, *Renew. Sust. Energ. Rev.*, 77, 297-317 (2017).
17. S. Mujeeb, T. A. Alghamdi, S. Ullah, A. Fatima, N. Javaid, T. Saba, Exploiting deep learning for wind power forecasting based on big data analytics, *Appl. Sci.*, 9(20), 4417-4435(2019).
18. R. Jiao, X. Huang, X. Ma, L. Han, W. Tian, A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting, *IEEE Access*, 6, 17851-17858(2018).
19. R. Veena, S. Mathew, M. I. Petra1, Artificially intelligent models for the site-specific performance of wind turbines, *Int. J. Energy Environ. Eng.*, 11(3), 289-297(2020).
20. J. Nielson, K. Bhaganagar, R. Meka, A. Alaeddini, Using atmospheric inputs for artificial neural networks to improve wind turbine power prediction, *Energy*, 190, 116273(2020).
21. H. Acikgoz, C. Yildiz, M. Sekkeli, An extreme learning machine based very shortterm wind power forecasting method for complex terrain, *Energ. Source Part A.*, 42(22), 2715-2730(2020).
22. X. Zhao, N. Jiang, J. Liu, D. Yu, J. Chang, Short-term average wind speed and turbulent standard deviation forecasts based on one-dimensional convolutional neural network and the integrate method for probabilistic framework, *Energy Conv. Manag.*, 203, 112239(2020).
23. Y. L. Tu, T. J. Chang, C. L. Chen, Y. J. Chang, Estimation of monthly wind power outputs of WECS with limited record period using artificial neural networks, *Energy Conv. Manag.*, 59, 114-121(2012).
24. R. G. Thippa, P. R. M. Swarna, M. Parimala, L. C. Chiranjii, K. Praveen, M. Reddy, H. Saqib, Z. K. Wazir, A deep neural networks based model for uninterrupted marine environment monitoring, *Comput. Commun.*, 157, 64-75(2020).
25. S. H. Hong, K. S. Shin, Using GA based input selection method for artificial neural network modeling: application to bankruptcy prediction, *Korea Intelli. Inf. Syst. Soc.*, 9(1), 227-249(2003).
26. G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, *Appl. Energy*, 87(7), 2313-2320 (2010).
27. Y. Hong, C. L. P. P. Rioflorido, A hybrid deep learning-based neural network for 24-h ahead wind power forecasting, *Appl. Energy*, 250, 530-539(2019).
28. S. Wang, H. Chen, A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network, *Appl. Energy*, 235, 1126-1140 (2019).
29. Y. Qin, K. Li, Z. Liang, B. Lee, F. Zhang, Y. Gu, L. Zhang, F. Wu, D. Rodriguez, Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal, *Appl. Energy*, 236, 262-272(2019).
30. S. Oehmcke, O. Zielinski, O. Kramer, Input quality aware convolutional LSTM networks for virtual marine sensors, *Neurocomputing*, 275, 2603-2615(2018).
31. H. Liu, X. Mi, Y. Li, Smart deep learning based wind speed

- prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network, *Energy Conv. Manag.*, 166, 120-131 (2018).
32. H. Zhiyi, S. Haidong, Z. Xiang, Z. Xianzhu, Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions, *Knowl. Based Syst.*, 207, 106396(2020).
 33. Y. H. Jung, J. H. Jung, B. G. Kim, S. U. Han, Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: case study of South Korea, *J. Clean Prod.*, 250, 119476(2020).
 34. M. Ibrahim, A. Alsheikh, Q. Al-Hindawi, S. Al-Dahidi, H. ElMoaqet, Short-time wind speed forecast using artificial learning-based algorithms, *Comput. Intell. Neurosci.*, 2020, 8439719(2020).
 35. M. A. Habib, S. A. M. Said, M. A. El-Hadidy, I. Al-Zaharna, Optimization procedure of a hybrid photovoltaic wind energy system, *Energy*, 24(11), 919-929(1999).
 36. Z. Lin, X. Liu, Wind power forecasting of an offshore wind turbine based on highfrequency SCADA data and deep learning neural network, *Energy*, 201, 117693(2020).
 37. T. Hu, W. Wu, Q. Guo, H. B. Sun, L. Shi, X. Shen, Very short-term spatial and temporal wind power forecasting: a deep learning approach, *CSEE J. Power Energy Syst.*, 6(2), 434-443(2020).
 38. Y. Cui, C. Huang, Y. Cui, A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm, *Environ. Sci. Pollut. Res.*, 27(7), 7353-7365(2019).
 39. N. Srivastava, G. Hinton, A. Krizhevsky, Ilya sutskever and ruslan salakhutdinov, dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.*, 15(56), 1929-1958(2014).
 40. B. J. Kim, M. A. Han, J. D. Baek, A convolutional neural network-based anthropomorphic model observer for signal detection in breast CT Images without human-labeled data, *IEEE Access*, 8, 162122-162131(2020).
 41. T. Demirdelen, P. Tekin, I. O. Aksu, F. Ekinci, The prediction model of characteristics for wind turbines based on meteorological properties using neural network swarm Intelligence, *Sustainability*, 11(17), 4803(2019).
 42. M. Carolin Mabel, E. Fernandez, Analysis of wind power generation and prediction using ANN: a case study, *Renew. Energy*, 33(5), 986-992(2008).
 43. B. Manobel, F. Sehnke, J. A. Lazzúus, I. Salfate, M. Felder, S. Montecinos, Wind turbine power curve modeling based on gaussian processes and artificial neural networks, *Renew. Energy*, 125, 1015-1020(2018).
 44. J. Shi, J. Guo, S. Zheng, Evaluation of hybrid forecasting approaches for wind speed and power generation time series, *Renew. Sust. Energ. Rev.*, 16(5), 3471-3480(2012).
 45. H. R. Fogno Fotso, C. V. A. Kazée, G. Djuidje Kenmoe, Optimal input variables disposition of artificial neural

networks models for enhancing time series forecasting accuracy, *Appl. Artif. Intell.*, 34(11), 792-815(2020).

46. Y. Ju, G. Y. Sun, Q. Chen, M. Zhang, H. Zhu, M. U. Rehman, A model combining convolutional neural network and lightGBM algorithm for ultra-short-term wind power forecasting, *IEEE Access*, 7, 28309-28318(2019).
47. Y. Zhao, L. Jia, A short-term hybrid wind power prediction model based on singular spectrum analysis and temporal convolutional networks, *J. Renew. Sustain. Energy*, 12, 053101(2020).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors and Contribution Statement

Tae-Hee Park

Department of Environmental Engineering, Kangwon National University, Bachelor, ORCID[®] 0000-0002-6016-3392: Data curation, Data analysis, Writing - original draft.

Da-Seul Jang

Department of Environmental Engineering, Kangwon National University, Bachelor, ORCID[®] 0000-0002-3971-8066: Methodology, Visualization.

Gyeong-Min Bae

Department of Environmental Engineering, Kangwon National University, Bachelor, ORCID[®] 0000-0002-0919-4871: Conceptualization, Validation.

Kyung-Min Kim

Department of Integrated Energy and Infra System, Kangwon National University, Master, ORCID[®] 0000-0002-7348-9367: Conceptualization, Methodology, Validation, Writing - Original Draft.

Johng-Hwa Ahn

Department of Environmental Engineering, Kangwon National University, Department of Integrated Energy and Infra System, Kangwon National University, Professor, ORCID[®] 0000-0001-5745-5338: Supervision, Project administration, Funding acquisition, Resources, Writing - review and editing.