

정확도 향상을 위한 CNN-LSTM 기반 풍력발전 예측 시스템

박래진¹⁾ · 강성우²⁾ · 이재형³⁾ · 정승민^{4)*}

CNN-LSTM based Wind Power Prediction System to Improve Accuracy

Rae-Jin Park¹⁾ · Sungwoo Kang²⁾ · Jaehyeong Lee³⁾ · Seungmin Jung^{4)*}

Received 5 January 2022 Revised 22 March 2022 Accepted 24 March 2022 Published online 23 May 2022

ABSTRACT In this study, we propose a wind power generation prediction system that applies machine learning and data mining to predict wind power generation. This system increases the utilization rate of new and renewable energy sources. For time-series data, the data set was established by measuring wind speed, wind generation, and environmental factors influencing the wind speed. The data set was pre-processed so that it could be applied appropriately to the model. The prediction system applied the CNN (Convolutional Neural Network) to the data mining process and then used the LSTM (Long Short-Term Memory) to learn and make predictions. The preciseness of the proposed system is verified by comparing the prediction data with the actual data, according to the presence or absence of data mining in the model of the prediction system.

Key words CNN, LSTM, Wind power prediction(풍력발전예측), Machine learning(기계학습)

Nomenclature

x : input data

x_{\min} : minimum value of input data

x_{\max} : maximum value of input data

f_t : forget gate function

i_t : input gate function

o_t : output gate function

σ : sigmoid

W : weight

b : bias

n : number of data

i : data order

t_i : actual data

y_i : predict data

1) Researcher, Department of Electrical Engineering, Hanbat National University

2) Researcher, Department of Electrical Engineering, Korea University

3) Senior researcher, Korea Electric Power Corporation

4) Professor, Department of Electrical Engineering, Hanbat National University

*Corresponding author: seungminj@hanbat.ac.kr

Tel: +82-42-821-1096

Fax: +82-42-821-1088

Subscript

IRENA : international renewable energy agency

NWP : numerical weather prediction

AR : autoregressive

- MA : moving average
 ARMA : autoregressive moving average
 ARIMA : Autoregressive intergrated moving average
 RNN : recurrent neural network
 LSTM : long short-term memory
 CNN : convolutional neural network
 MSE : mean squared error
 RMSE : root mean square error
 WAPE : weighted absolute percentage error

1. 서론

탄소를 발생시키는 발전원으로 인해 환경에 대한 우려가 확산되며 신재생에너지원의 이용을 위한 정책이 확정되었다. 이에 따라 국제적으로 신재생에너지 발전원이 확대되고 있으며 특히 태양광 발전과 비교해 면적당 출력이 높고, 높은 설비 이용률을 확보할 수 있다는 측면에서 풍력발전이 주목을 받고 있다. IRENA의 통계에 따르면 세계적으로 풍력발전의 설비용량은 2010년 약 180 GW에서 2020년 약 730 GW로 급격한 확대를 보였으며, 국내 또한 382 MW에서 1636 MW로 제주지역을 중심으로 풍력발전 설비가 확대되고 있다.^[1] 하지만 풍속의 경우 불규칙성과 변동성의 특징을 가지고 있어 간헐적 문제 출력은 풍력발전의 수용량 측면에서 숙제로 남아있다.^[2] 최근 제주지역은 계통의 안정화를 위해 풍력발전기의 출력제한 횟수와 용량이 증가하고 있으며 2020년에는 77회의 출력제한이 이루어졌다.^[3] 전력시스템에 대한 안정화와 신재생에너지 발전원의 이용률을 향상시키기 위해 풍력발전 예측 정확도를 높이기 위한 연구가 진행되고 있다.

풍력발전 예측 시스템은 크게 물리적 방법과 통계적 방법으로 나뉜다. 물리적 방법은 온도, 기압과 같은 NWP와 위상 정보 그리고 풍력 터빈의 허브 높이, 출력 곡선과 같은 기술적 특성을 고려한 수학적 모델링을 바탕으로 예측이 진행된다. 통계적 방법은 많은 양의 과거 데이터를 예측 모델에 적용하여 진행되며 일반적으로 기계 학습, 시계열 분석 접근 방식을 포함한다.^[4]

최근 기계 학습을 활용한 기술이 주목을 받으며 풍력발

전 예측 시스템 또한 통계적 방법 중 인공지능을 이용한 연구가 많은 진행 중이다. 풍력발전 예측의 통계적 방법을 활용한 기존의 시계열 모델은 AR, MA 모델이 있으며 두 가지 모델을 결합한 ARMA 모델 그리고 가장 많이 사용되는 ARIMA 모델이 있다. ARIMA 모델은 회귀를 이용하여 오차를 개선하는 모델로서 시계열 데이터에 주로 활용되지만, 주기성 및 계절성을 나타내는 데이터의 정확한 예측은 어려운 것으로 평가된다.^[5] 풍력발전은 풍속과 풍향에 주로 의존하여 발전량이 산정되며, 각 시계열 데이터는 비선형을 이루고 있고 계절성을 보인다. 기계 학습을 기반으로 한 RNN 알고리즘은 모든 변수에 대한 가중치를 학습시킬 수 있어 시계열 데이터를 다루기 적합하다고 평가되고 있다.^[6] 하지만 RNN은 데이터의 양이 방대하면 과거의 데이터 정보가 소실되어 학습을 통해 가중치를 결정하는 과정에서 영향력이 없어지는 장기 의존성 문제가 나타날 수 있다.

LSTM을 이용한 예측 알고리즘은 이러한 문제를 완화하기 위해 산업계에서 활용되어 있다.^[7] 기존에 선행된 연구는 데이터보다 모델의 구성에 집중하여 예측 시스템을 구축하는 데 중점을 두는 경향이 많았으며, 시계열 데이터는 풍속과 발전량의 데이터를 이용하여 진행되는 경우가 많았다. 이러한 방법은 불규칙한 풍속을 뒷받침하는 데이터가 부족하여 모델을 수정하여 예측 정확도를 높이기 어렵음이 존재할 수 있다. 따라서 데이터의 전처리 및 데이터마이닝을 하여 모델에 적용시키는 것이 중요하다고 평가되고 있다.

본 논문에서는 풍력발전량을 예측하여 신재생에너지원의 이용률을 높이기 위해 기계 학습 그리고 데이터마이닝을 적용한 풍력발전 예측 시스템을 제안하고자 한다. 시계열 데이터는 풍속과 발전량 그리고 풍속에 영향을 미치는 환경데이터를 추가하여 데이터셋을 설정하였다. 데이터셋은 모델에 적합하게 적용 가능하도록 전처리를 진행하였다. 예측 시스템은 CNN을 적용하여 데이터마이닝 과정을 진행한 후 LSTM을 활용하여 학습하고 예측을 진행하였다. 예측 시스템의 모델에 데이터마이닝의 유무에 따른 예측 데이터를 실제 데이터와 오차 비교를 통해 제안된 예측 시스템을 검증하고자 한다.

2. 모델 최적화

예측 시스템을 구성할 경우 모델의 구축도 중요하지만 데이터에 대한 전처리가 중요한 것으로 평가된다. 본 논문에서는 데이터의 전처리와 특징 추출에 중점을 두어 예측 모델을 구성하였으며 이는 CNN을 활용하였다. 시계열 데이터는 LSTM을 사용하여 예측을 진행하며 CNN-LSTM 예측 시스템을 구축하였다.

2.1 데이터 구축

2.1.1 데이터 전처리

시계열 데이터를 사용하여 예측 시스템을 구성할 때 데이터의 전처리를 통해 모델에 입력하여 출력에 부정적인 영향을 미치지 않도록 하는 것이 중요하다.^[8] 데이터 전처리를 함으로써 불완전한 데이터를 처리하는 과정을 거친 후 정규화를 적용, 예측 시스템에 입력하는 과정이 필요하다. 정규화는 입력으로 사용되는 변수 간 스케일 차이를 조정하여 개별 데이터의 크기를 모두 똑같은 단위로 변경하는 것을 의미한다. Standard Scaler, Robust Scaler, Min-Max Scaler, Normalizer와 같은 방법이 존재하며 본 논문에서는 가장 일반적으로 사용되는 Min-Max Scaler를 활용했다. 학습시키는 데이터 변수(Feature)가 여러 개가 존재할 경우, 최댓값의 크기가 다른 각 데이터 변수의 값을 0~1 또는 -1~1 사이의 값으로 변환하여 스케일을 조정한 후 학습

을 진행하는 것이 일반적이며 해당 식은 (1)과 같다.

$$Min - Max Scaler(x) = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

본 논문에서는 풍력발전기에서 측정된 실측데이터와 풍력발전기가 위치한 지역 인근의 기상청 환경 변수를 입력으로 하였고 전처리를 통해 정규화를 진행하며 CNN 과정을 거친다. LSTM을 활용해 학습시킨 뒤 예측을 진행하고 검증 및 시각화를 진행한다. 해당 과정을 예측 시스템 구성과정으로 표현하면 Fig. 1과 같다.

시계열 데이터 중 패턴에서 벗어난 값을 이상치(Outlier)라 정의한다. 해당 시계열 데이터는 두산중공업의 3 MW급 TC-II, TC-S 모델의 풍력발전기를 기준으로 하고 있어 3 MW를 초과하는 값과 0 W 미만의 값을 이상치로 정의하고 해당 데이터를 각각 3 MW와 0 MW로 고정하였다.^[9] 서남해 실증단지 풍력발전기를 통해 측정된 10월 1일부터 10월 21일까지의 원천 데이터를 Fig. 2에 나타내었으며, 도출된 이상치에 대한 부분을 확인할 수 있다. y축은 발전량을 의미하여 단위는 kW이고 x축은 시간을 의미한다. 또한 연속된 시계열 데이터 중 73개의 시간 데이터가 생략된 것을 확인하였다. 시계열 데이터는 연속된 데이터 형태로 기계 학습을 할 때 시간의 공백이 존재하면 학습률이 낮아질 수 있으며 해당 데이터는 각각의 앞뒤 측정값의 평균으로 정의하였다.

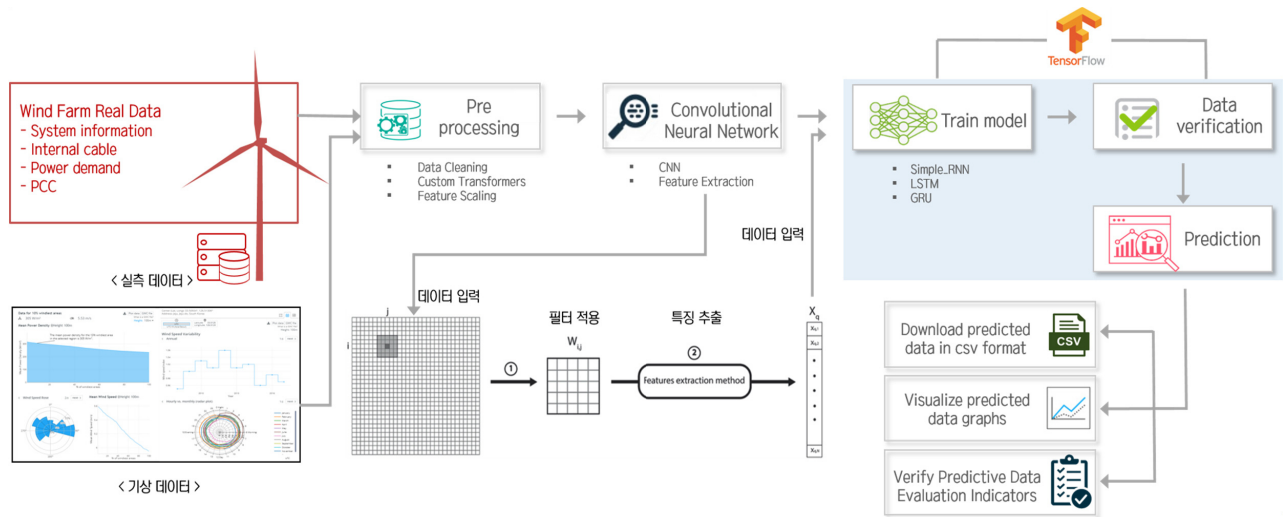


Fig. 1. Configuration of wind power prediction system

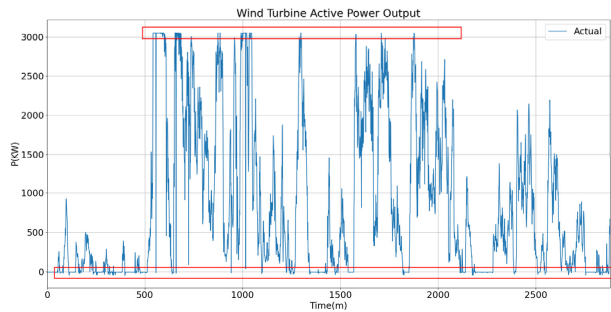


Fig. 2. Raw data of wind power extraction

2.1.2 데이터 히트맵

히트맵(Heatmap)은 데이터를 분석하는 데 있어서 중요한 과정으로 변수와 변수 사이의 선형적 관계가 있는지 보기 위해 시각화하는 방법이다. 해당 논문에서는 변수와 풍력발전기의 출력 값에 대한 상관관계 분석을 통해 변수를 설정하였다. 본 논문에서는 Seaborn 라이브러리를 활용한 데이터 히트맵을 통해 상관관계 분석을 진행하였으며, 풍력발전기를 통해 측정된 변수 5개와 환경 변수 3개를 더하여 출력변수(Output)를 제외한 총 8개의 변수가 출력값을 산정하는데 적합한 것으로 판단되었다. 상관계수가 높게 나왔을 경우 선형적 관계가 높다는 것을 의미하고 상관계수가 낮을 경우는 독립적인 데이터가 아닌 비선형적 관계가 있을 수 있어 각 변수들에 대한 고찰이 필요하다. Fig. 3은 각 8개의 변수에 대한 히트맵을 나타내고 있다. 히트맵으

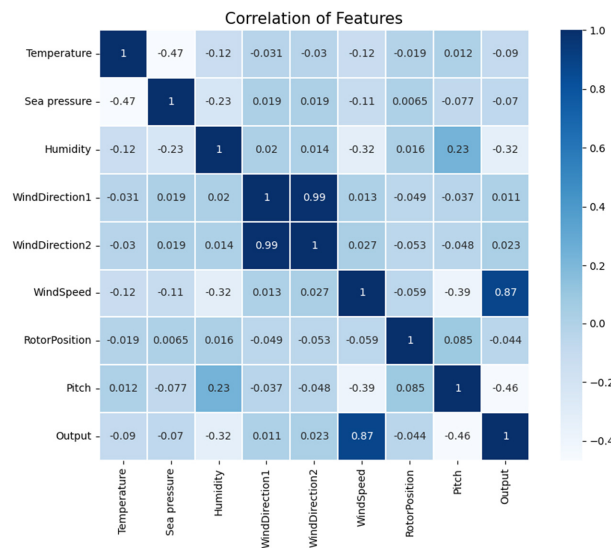


Fig. 3. Data heatmap

로 나타난 결과, 출력결과와 가장 선형적인 관계를 나타내는 것은 풍속이며, 블레이드의 피치 각도도 높은 선형적 관계를 이루고 있다. 피치 각도에 대한 상관계수가 음수를 이루고 있으며 이는 선형적인 관계지만 반비례를 의미한다. 따라서 피치 각도가 향상할 경우, 풍력발전기의 출력은 줄어든다는 것을 파악할 수 있다. 히트맵을 사용하여 변수들의 상관관계를 분석할 수 있으므로 추가적인 변수를 사용할 때 결과에 미치는 영향에 대한 분석이 가능하여 추후, 예측 정확도를 높이기 위한 데이터셋을 구성할 때 중요할 것으로 판단된다.

2.2 예측 알고리즘

2.2.1 Convolutional Neural Network

CNN은 합성곱 신경망으로서 이미지 분류에 주로 사용되지만 최근 데이터마이닝의 한 종류로서 사용되는 기계 학습 방법이다.^[10] CNN은 입력 데이터를 학습하고 패턴을 분석하며 데이터를 필터링하고 특징을 추출하는 Feature Extraction 역할을 수행한다. Convolution Layer와 Max Pooling 레이어를 반복적으로 Stack을 쌓는 방식으로 진행된다. CNN 레이어는 필터가 합성곱 연산을 시행하고 Max Pooling 레이어를 통해 최댓값을 추출함으로써 데이터의 공간적 특징을 확보한다.^[11] 따라서 복잡한 데이터셋을 효과적으로 구성할 수 있다. 이후 시계열 데이터에 적합한 모델인 LSTM에 적용하기 위해 CNN에서의 2차원 배열의 출력형태를 1차원으로 변환하는 Flatten 함수를 사용하였다. 차원을 변환시킨 데이터를 LSTM에 적용함으로써 데이터의 특징을 추출하고 시간적 추세를 반영한 CNN-LSTM 모델을 구축하였으며 Fig. 4와 같다.

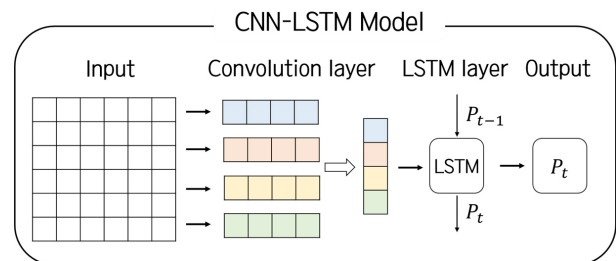


Fig. 4. CNN-LSTM model configuration

2.2.2 Long Short-Term Memory

최근에 많이 활용되는 LSTM은 RNN의 한 종류이며 긴 기간을 가지고 있는 시계열 데이터를 활용할 때 적합한 학습을 수행할 능력을 갖추고 있다. 본 논문에서 활용한 LSTM 구성은 Fig. 5와 같다. 기본적인 LSTM 구조는 같지만 입력하는 값에 환경변수를 추가하여 학습시키는 양을 증가시켰다. 추가한 환경 변수는 온도, 습도, 기압이며 기존의 풍력발전기에서 측정된 변수의 기간과 시간 스케일을 동일하게 구성했다.

LSTM은 Cell state와 Forget gate, Input gate, Output gate로 구성된다. Forget gate는 Cell state로부터 정보의 보존 여부를 Sigmoid layer에 의해 결정한다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

Input gate는 들어오는 데이터를 Cell state에 저장 여부를 결정한다. Sigmoid layer를 거쳐 값을 저장할지 정하고 tanh layer에서 새로운 vector를 만든다. 이후 Cell state에 업데이트를 진행한다.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (4)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (5)$$

Output gate는 sigmoid layer와 tanh layer의 과정을 거쳐 정보를 선택적으로 내보내게 된다.

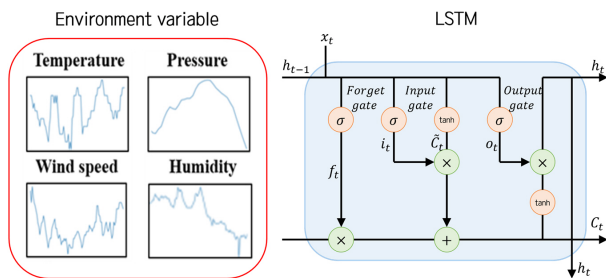


Fig. 5. LSTM configuration

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \times \tanh(C_t) \quad (7)$$

본 논문에서 제안하는 예측 시스템의 알고리즘은 Fig. 6과 같다. 합성곱 레이어와 Max Pooling을 이용하여 CNN을 구성하였고 CNN의 활성화 함수는 ReLU값으로 지정하였다. 또한 시계열 데이터의 특성을 고려하여 시간의 흐름에 따른 정확도를 향상시키기 위해 CNN-LSTM 모델에서 Moving Window 알고리즘을 적용하였다.

2.3 회귀지표평가

회귀지표평가는 학습데이터를 이용하여 훈련을 진행할

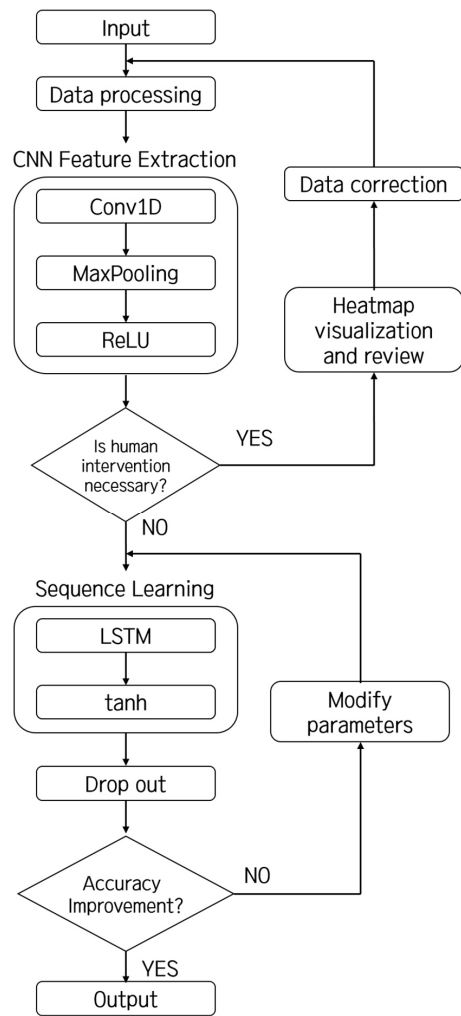


Fig. 6. Proposed flow chart for prediction

때 손실 함수를 활용해 훈련 과정에서의 오차를 산출하고 훈련 오차가 최소가 되는 방향으로 모델을 수정하기 위한 지표이다. 또한 결과를 도출하였을 때 실제 데이터와 예측 데이터 간의 오차를 비교하는 방법이다. 일반적인 회귀 지표 평가는 다음과 같은 방법이 존재한다. 예측 데이터와 실제 데이터 간의 오차를 보기 위한 척도는 평균 제곱근 오차인 RMSE와 가중 평균 백분율 오차인 WAPE를 사용하였으며 각각 식 (8), 식 (9)과 같이 나타낸다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2} \quad (8)$$

$$WAPE = \frac{\sum_{i=1}^n |y_i - t_i|}{\sum_{i=1}^n |t_i|} \times 100\% \quad (9)$$

2.4 시뮬레이션

전북 부안군 위도 및 고창군 해역의 서남해 실증단지에 있는 풍력발전기 20기 중 두산중공업 TC-S 모델 1기에 대한 시뮬레이션을 진행하였다. 풍력발전기에서 취득한 정보는 풍향, 풍속, 3개 블레이드 각각의 피치 각도 등이 있다. 또한 예측 정확도를 높이는 방안으로, 기상청에서 해당 풍력발전기 인근의 기상 데이터를 추가하여 데이터 보완을 수행하였다. 기상 데이터는 서남해 실증단지 인근 위도를 기준으로 하였으며 기온, 해면기압, 습도를 추가하였다.

데이터셋은 8개의 항목이 10분 데이터로 존재하게 된다. 적용한 LSTM 모델의 매개변수는 Table 1과 같다. 히든 레이어는 CNN, Max Pooling, LSTM을 포함하여 5개의 층으로 구성되어 있다. Epoch는 과적합을 고려하여 선정하였다. Batch Size는 데이터 학습의 제한으로 인해, 개수를 나눠주는 변수로써, 1일 10분 데이터 수가 144개로 존재하기 때문에 하루 예측 정확도를 높이기 위해 다음과 같이 선정했다. Learning Rate는 학습률을 의미하며 일반적으로 사용되는 학습률로 설정하였다.

Table 1. Parameters for LSTM model

Hidden layer	Epoch	Batch size	Learning rate
5	50	144	0.001

본 논문에서는 환경데이터를 추가한 데이터셋을 입력으로 하여 2020년 10월 1일부터 2021년 1월 29일까지의 데이터를 학습을 진행하였다. 이를 바탕으로, 2021년 1월 30일부터 2월 28일까지의 데이터를 테스트 데이터로 구성하여 일일 데이터 예측을 수행하였다. Fig. 7은 환경 변수를 포함한 데이터셋으로 CNN을 적용하지 않은 그래프이며 이를 사례 1로 정의한다. Fig. 8은 환경변수가 없는 데이터셋으로 CNN을 적용하여 예측한 발전량 그래프이며 이를 사례 2로 정의한다. Fig. 9는 환경 변수를 추가하고 CNN을 적용한 모델로 발전량을 예측한 그래프이며 사례 3으로 정

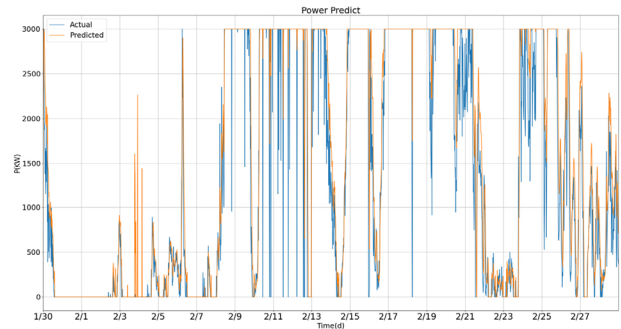


Fig. 7. Prediction curve without CNN

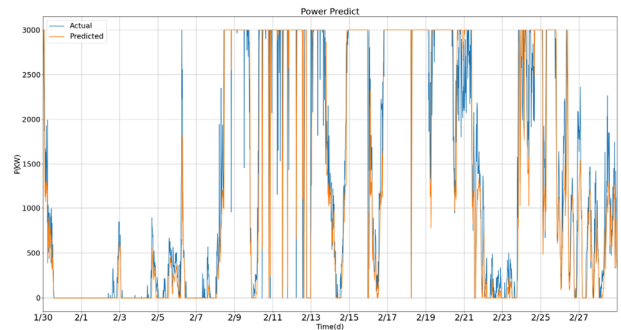


Fig. 8. Prediction curve with no environmental factors

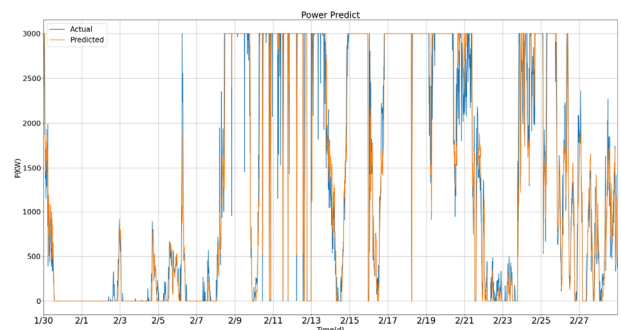


Fig. 9. Prediction curve with environmental factors

Table 2. Measured errors in case study

Case	RMSE	WAPE
1	389.9047	10.22961
2	340.2082	4.6354
3	275.9002	0.33066

의한다. 사례별 실제 값과 예측값에 대한 오차를 Table 2로 나타내었다.

Fig. 10은 환경 변수와 CNN의 유무에 따른 정확도 비교를 보이기 위해 2월 24일부터 2월 28일까지의 시간을 확대한 그래프이다. 환경 변수가 있음에 따라 실제 값에 가까워지는 것을 확인할 수 있다. Fig. 11은 사례별 연구 결과를 비교하기 위해 2월 13일에서의 시간을 확대한 그래프이다. 환경 변수와 CNN을 이용했을 경우 가장 실제 값에 근접하였고 사례 2, 사례 3의 순으로 예측값이 실제 값에 유사한 것을 확인할 수 있다. 그래프는 30일 동안 10분 단위의 예측을 진행하였으며 그래프의 y축은 발전량 x축은 순차적으로 예측되는 데이터를 의미한다.

시뮬레이션 결과 WAPE 기준으로 사례 1은 약 10.23%, 사례 2는 4.64%, 사례 3은 0.33%로 본 논문에서 제시한

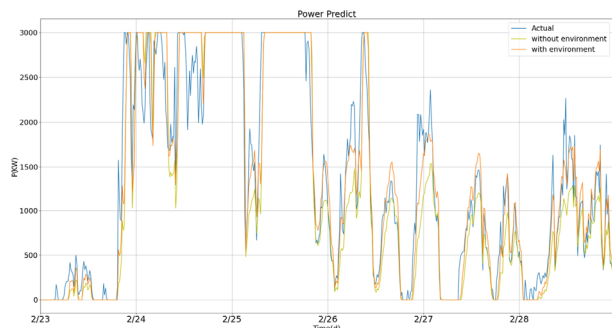


Fig. 10. Comparison curve for prediction accuracy

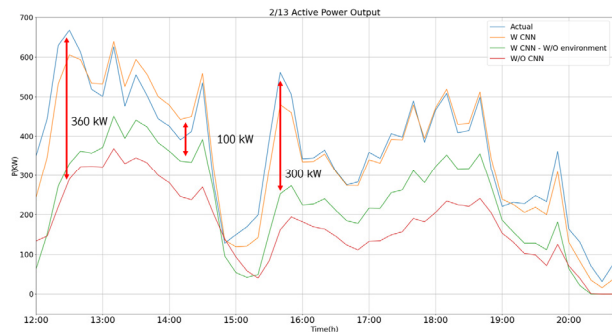


Fig. 11. Closed comparison curve for prediction accuracy

방법이 가장 오차율이 감소함을 확인하였다. 또한 Fig. 10의 그래프를 기준으로 가장 극단적인 오차를 확인하였을 때 사례 1은 약 360 kW, 사례 2는 약 300 kW, 사례 3은 약 100 kW의 발전량 오차가 발생한 것을 확인하였다. 이를 통해 본 논문에서 제안한 기법을 활용할 경우 상대적으로 적은 오차가 발생함을 검증하였다.

3. 결론

본 논문에서는 기계 학습 기반 풍력발전 예측 시스템의 입력값과 모델에 따른 정확도 비교를 위해 시뮬레이션을 진행하였다. 제안한 예측 모델의 신뢰성을 확인하고자 세 가지 연구 사례의 결과 값으로 비교하였으며 실제 발전량과의 정확도 비교를 하기 위해 평가지표를 작성하여 확인하였다. 시뮬레이션 결과 연구 사례 3인 환경변수를 고려하고 CNN을 통해 데이터마이닝을 진행하였을 때 오차가 저감됨을 확인하였다.

신재생에너지 보급률 확대로 불규칙한 에너지공급이 확대되고, 수요·공급제어에 예측 시스템 고도화가 요구되고 있다. 본 논문에서 제안한 예측기법은 전력계통에 투입되는 풍력발전을 안정적으로 예측하는 데 활용할 수 있으며, 풍력발전의 일반적인 제어관리시스템 구조에 적용할 수 있다. 대규모 풍력발전은 출력제한 과정에 예측 시스템 응용이 일반화되고 있으며, 제안된 방식의 예측 정확도 향상으로 비용의 경감이 가능하다. 향후, 예측의 신뢰성과 정확성을 향상하기 위해, 각 입력변수의 의미, 입력변수 간 상관관계, 다중공선성 등을 고려하여 최적화를 진행할 계획이다. 마지막으로, 시계열 예측에서 필수적인 3가지 요소인 추세, 계절성, 주기성의 분석의 중요성의 영향력이 커지는 추세이다. 특히, 풍력 터빈의 일시적 점검 및 통신오류로 인한 입력 데이터를 고려하고 이상치 데이터 수정 및 보완하여 예측의 정확도 향상을 위한 개선이 요구된다.

감사의 글

This work was supported by the Korea Electric Power Corporation grant (No. R21XO01-23) and the Korea

Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE) (No. 20183010025440).

References

- [1] IRENA, “Wind Energy”, <https://www.irena.org/wind>.
- [2] Byon, J.Y., Kim, Y.M., and Choi, B.C., 2014, “Variability of future wind and solar resource over the Korean peninsula based on climate change scenario”, *New. Renew. Energy*, 10(2), 29-39.
- [3] Nam, I.S., Lee, J.W., and Jung, S.M., 2021, “A study about the curtailment for wind turbines considering power sensitivity of distribution system”, *The Transactions of the Korean Institute of Electrical Engineers*, 71(3), 495-502.
- [4] Myung, H.S., Kim, H.C., Kang, N.H., Kim, Y.H., and Kim, S.H., 2018, “Analysis of the load contribution of wind power and photovoltaic power to power system in Jeju”, *J. Korean Sol. Energy Soc.*, 38(1), 13-24.
- [5] Jin, S.J., and Cho, H.J., 2020, “Seasonal time series forecasting using machine learning”, *JKDAS*, 22(5), 1779-1791.
- [6] Li, Z., Zou, H., and Qi, B., 2019, “Application of ARIMA and LSTM in relative humidity prediction”, *IEEE 19th International Conference on Communication Technology (ICCT)*, 1544-1549.
- [7] An, Y.J., Lee, T.K., and Kim, K.H., 2021, “Prediction of photovoltaic power generation based on LSTM considering daylight and solar radiation data”, *Trans. Korean. Inst. Elect. Eng.*, 70(8), 1096-1101.
- [8] Kim, D.J., Hwang, C.G., and Yoon, C.P., 2020, “Learning data preprocessing technique for improving indoor positioning performance based on machine learning.”, *JKIICE*, 24(11), 1528-1533.
- [9] Korea Offshore Wind Power, <http://www.kowp.co.kr/data/file.asp?menu=14>.
- [10] Zhao, B., Lu, H., Chen, S., Liu, J., and Wu, D., 2017, “Convolutional neural networks for time series classification”, *Journal of Systems Engineering and Electronics*, 28(1), 162-169.
- [11] Kim, T.Y., and Cho, S.B., 2019, “Predicting residential energy consumption using CNN-LSTM neural networks”, *Energy*, 182, 72-81.