

LSTM 딥러닝 신경망 모델을 이용한 풍력발전단지 풍속 오차에 따른 출력 예측 민감도 분석

강민상* · 손은국** · 이진재* · 강승진*

Analysis of wind farm power prediction sensitivity for wind speed error using LSTM deep learning model

Minsang Kang*, Eunkuk Son**, Jinjae Lee* and Seungjin Kang*

Key Words : LSTM (장기 단기 기억), Recurrent neural network (순환 신경망), Power prediction (출력 예측), Wind farm (풍력발전단지), Wind turbine (풍력터빈), Wind speed error (풍속 오차)

ABSTRACT

This research is a comprehensive analysis of wind power prediction sensitivity using a Long Short-Term Memory (LSTM) deep learning neural network model, accounting for the inherent uncertainties in wind speed estimation. Utilizing a year's worth of operational data from an operational wind farm, the study forecasts the power output of both individual wind turbines and the farm collectively. Predictions were made daily at intervals of 10 minutes and 1 hour over a span of three months. The model's forecast accuracy was evaluated by comparing the root mean square error (RMSE), normalized RMSE (NRMSE), and correlation coefficients with actual power output data. Moreover, the research investigated how inaccuracies in wind speed inputs affect the power prediction sensitivity of the model. By simulating wind speed errors within a normal distribution range of 1% to 15%, the study analyzed their influence on the accuracy of power predictions. This investigation provided insights into the required wind speed prediction error rate to achieve an 8% power prediction error threshold, meeting the incentive standards for forecasting systems in renewable energy generation.

기호설명

$P_{p,F}$: 풍력발전단지 출력 예측값 [MW]
 $P_{p,t,N}$: N호기 풍력터빈의 출력 예측값 [MW]
 $P_{p,F,i}$: 풍력발전단지 i번째 시계열 출력 예측값 [MW]
 $P_{a,F,i}$: 풍력발전단지 i번째 시계열 출력 실제값 [MW]
 n : 시계열 데이터 수 [-]
 $\sigma_{p,F}$: 풍력발전단지 설비용량 [MW]

$\bar{\sigma}_{w,s,t,d}$: 개별 풍력터빈별 실제 일평균 풍속 [m/s]
 $WS_{e,t,i}$: 개별 풍력터빈 i번째 시계열 풍속 오차 [m/s]
 $WS_{a,t,i}$: 개별 풍력터빈 i번째 시계열 풍속값 [m/s]
 $\sigma_{p,t}$: 개별 풍력터빈 설비용량 [MW]
 $P_{p,t,i}$: 개별 풍력터빈 i번째 시계열 출력 예측값 [MW]
 $P_{a,t,i}$: 개별 풍력터빈 i번째 시계열 출력 실제값 [MW]
 Err_t : 개별 풍력터빈의 출력 예측 오차 [kW]
 Err_F : 풍력발전단지의 출력 예측 오차 [kW]

* 한국에너지기술연구원 풍력연구단

** 한국에너지기술연구원 풍력연구단(교신저자)

E-mail : Eunkuk.son@kier.re.kr

DOI : <https://www.doi.org/10.33519/kwea.2024.15.2.002>

Received : February 27, 2024, Revised : April 15, 2024,

Accepted : April 29, 2024

1. 서론

바람의 운동에너지를 이용하는 풍력발전의 원리상 풍력터빈의 전력생산량은 풍속에 따라 시시각각 변동하는 특성을 갖는다. 최근 국내에 설치되는 풍력터빈의 설비용량은 지난 5년간 연 평균 약 10 %의 지속적인 증가세를 보이며, 2022년 기준으로 1,803 MW가 설치되어 있다.[1] 또한 2023년 정부의 제10차 전력수급 기본계획에 따르면 2036년까지 풍력발전 설비용량을 34,089 MW 까지 늘리는 것을 목표로 하여 2022년 대비 약 18.9배 까지 증가할 것으로 전망된다. 이는 2036년 전체 설비용량 목표 144.1 GW의 약 23.7 %로 풍력발전의 변동성으로 인한 전력 계통의 불안정성은 증대될 것으로 예상된다.[2] 이러한 전력계통의 영향에 대응하기 위하여 전력계통을 운영하는 전력거래소에서는 2020년 ‘재생에너지 발전량 예측제도(이하 “예측제도”)’를 시행하여 사전에 전력생산량에 대한 정확한 발전계획을 제출한 사업자에게는 인센티브를 부여함으로써 예측 정확도를 향상하기 위한 노력을 하고 있다.[3] 또한 출력 예측 관련 연구로 기상수치 예측모델 WRF(Weather Research and Forecasting)을 이용하거나, 시계열 기반의 통계적 예측모델로 활용되는 ARIMA(Autoregressive Integrated Moving Average) 및 PER(Persistence) 모델 등을 이용한 연구도 수행된 바 있다.[4] 최근 머신러닝(Machine learning), 딥러닝(Deep learning)과 같은 기술들이 빠르게 발전하면서 최근 풍력 예측에 활용되고 있고, 대표적인 딥러닝 방법 중 RNN(Recurrent Neural Network) 기반의 LSTM(Long Short-Term Memory Network) 모델을 이용한 방법이 활용되어 많은 연구에서 단기 풍력발전 예측 성능 향상을 입증한 바 있다.[5]

풍력발전단지의 출력은 풍속에 종속적인 관계를 갖기 때문에 정확한 풍속예측 기술은 출력예측에 필수적이다. 풍력발전단지 출력 ‘예측제도’가 도입된 이후 하루 전 발전량 예측오차율 8 %가 출력예측의 기준점이 되었지만, 실제로 이를 달성하기 위해 필요한 풍속의 예측 오차율에 대해 다루어진 사례는 극히 드물다. 따라서 본 연구에서는 LSTM 모델을 이용한 출력 예측 모델을 생성하고, 출력 예측 모델에 입력되는 풍속 데이터 오차에 따라 풍력 발전량 예측에 미치는 민감도를 분석하였다. 이를 통하여 풍력 발전사업자가 ‘예측제도’에 참여하기에 필요한 풍력 발전량 예측오차율을 달성하기 위해 요구되는 풍속 예측 오차율을 추정 하였다.

2. 풍력발전단지 출력 예측 방법

2.1 LSTM 모델

LSTM 모델은 1997년 Hochreiter와 Schmidhuber에 의해 제안된 모델로 순환신경망인 RNN(Recurrent Neural Network) 모델을 개선한 방법이다. 일반적으로 RNN 모델은 순차적인 데이터 처리에 강점을 갖지만, Fig. 1.과 같이 입력된 모든 과거 데이터들이 예측에 일정한 영향을 주는 구조를 갖고 있어 장기 데이터를 이용한 데이터 학습 시에는 최근 데이터에 대한 비중이 상대적으로 낮아지는 데이터의 장기 의존성으로 인한 한계를 갖는다.[6]

이를 극복하기 위하여 LSTM 모델은 Fig. 2.와 같이 메모리 셀을 4개의 망각층(Forget gate), 입력층(Input gate), 출력층(Output gate) 및 셀 스테이트(Cell state) 레이어로 구성된다. 각 레이어는 상호작용하며, 과거 데이터에 대한 의존성을 줄이고 최근 데이터에 대하여 적절한 가중치를 반영함으로써 특히 시계열 데이터 예측에 강점을 갖도록 개발되었다.[7]

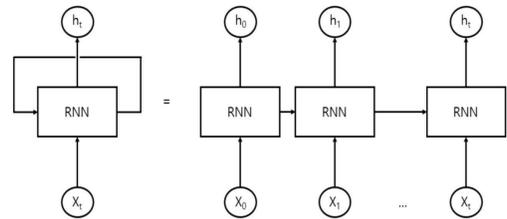


Fig. 1 The structure of RNN

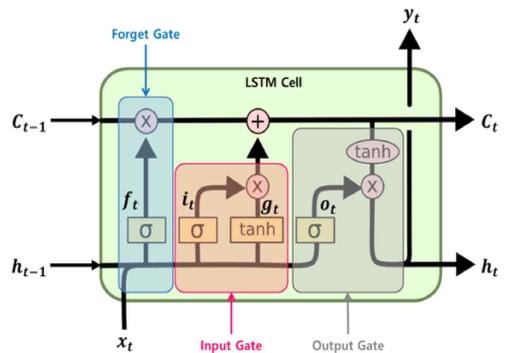


Fig. 2 The structure of LSTM

2.2 출력 예측 절차 및 분석 방법

딥러닝 기반 LSTM 출력 예측 모델을 이용한 풍력 발전단지 출력 예측 절차와 예측 결과에 대한 검증 및 분석방법은 Fig. 3과 같다. 먼저 풍력발전단지 내 모든 풍력터빈에 대하여 각각 LSTM 출력 예측 모델을 생성하고, 각 풍력터빈별 10분 및 1시간 간격의 시계열 출력 예측을 1일 단위로 수행하였다. 각 풍력터빈의 출력 예측 결과를 합산하여 풍력발전단지의 출력 예측 값을 도출하였다. 최종적으로 풍력발전단지 출력 예측 값은 실제 풍력발전단지의 SCADA 측정값과 비교하여 예측 결과에 대한 검증을 수행하였다. 풍력발전단지 출력 예측을 위하여 개별 풍력터빈 각각 출력을 예측한 이유는 풍력터빈별 출력곡선은 인접한 풍력터빈에 의한 후류영향으로 다른 특성을 보이며, 이를 반영하기 위해서는 각 풍력터빈의 풍속, 풍향에 대한 출력변화를 개별적으로 학습하여 모델을 생성하는 것이 필요하기 때문이다.

추가적으로 출력 예측 모델의 풍속 오차에 따른 민감도를 분석하기 위하여 출력 예측 모델에 입력되는 풍속을 1%, 3%, 5%, 10%, 15% 까지 총 5가지의

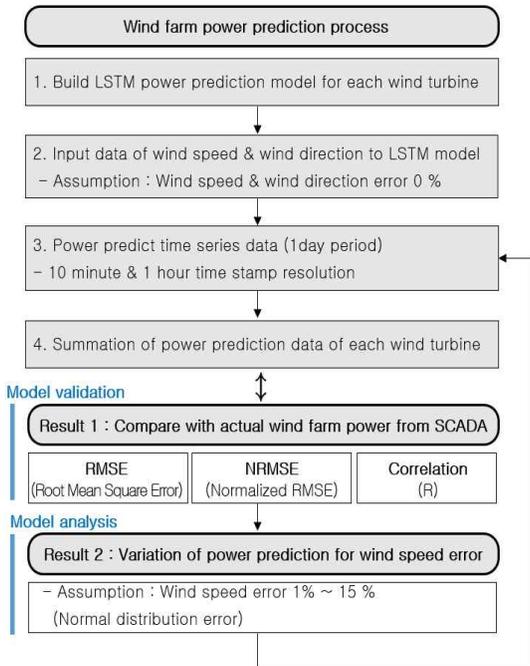


Fig. 3 The process of wind farm power prediction and model validation & model analysis method for the results

풍속 오차를 발생시켰을 때, 출력 예측 결과에 미치는 영향을 분석하였다. 이를 통해 출력 예측 오차 목표치에 따라 요구되는 풍속 예측의 오차율을 추정하고자 하였다.

2.3 출력 예측 모델

LSTM 출력 예측 모델은 풍력발전단지 내 풍력터빈 각각에 대하여 생성하며, 이들 예측 모델은 1일 단위로 출력을 예측하고 이를 합산하여 풍력발전단지 출력을 예측하고자 하였다. LSTM 출력 예측 모델의 학습 및 테스트, 예측 데이터의 구조는 Fig. 4와 같이 1일 단위로 이동하며 수행하였다. 출력 예측 모델 생성을 위한 데이터는 출력 예측을 수행한 기간의 하루 전까지의 SCADA 데이터를 포함하였다. 예를 들면 2020. 10. 01.일 출력데이터를 예측하고자 하는 경우, 학습에 사용한 데이터 기간은 2020. 09. 30.일까지의 데이터를 사용하였다. 마찬가지로 2020. 12. 31.일 출력데이터를 예측하고자 하는 경우, 학습에 사용한 데이터 기간은 2020. 12. 30.일까지의 데이터를 사용하였다. 1일 단위로 출력데이터 예측기간이 늦어질 때 마다 데이터 학습에 사용하는 데이터 기간이 1일씩 늘어나므로, 일별로 출력 예측 모델을 개별적으로 생성하였다. 이는 데이터 학습에 필요한 데이터를 최대한 확보하기 위한 방법으로 수행하였으나, 모델 학습에 필요한 시간 및 자원에 따라 장기간 데이터 확보가 가능한 경우에는 학습 데이터 기간을 최근 1년 정도로 한정하는 것도 필요할 것으로 보인다.

본 연구에서는 최초 예측 모델 생성에 사용하는 학습데이터 기간을 최근 9개월 동안의 데이터를 이용하

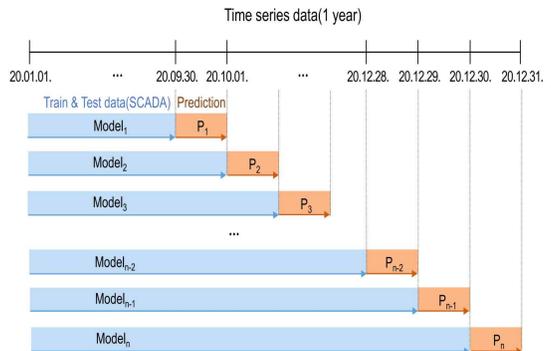


Fig. 4 The structure of modeling data

였는데, 이는 사전 연구를 통하여 최초 예측 모델 생성에 사용되는 학습 데이터 기간을 최근 1개월, 3개월 및 6개월로 달리하였을 때 가장 낮은 예측 오차를 보인 결과를 바탕으로 적용하였다. 출력 예측 모델 학습에 사용한 데이터 중 80 %는 딥러닝 학습을 위한 데이터로 활용하였고, 나머지 20 %의 데이터는 모델 평가를 위한 테스트 데이터로 활용하였다.

딥러닝 데이터 학습 전에 활용되는 데이터에 대한 전처리 과정에서 SCADA 데이터 중 결측치에 대한 데이터와 풍력터빈의 출력이 2 kW 이하에 해당하는 데이터는 삭제하였다. 풍력터빈이 2 kW 이하에서는 발전이 중지되고 가동이 빈번할 것으로 예상되어 정상적인 운전 데이터가 수집되지 못하기 때문이다. 추가적으로 데이터 학습에 활용되는 풍속, 풍향, 출력값은 데이터 범위가 서로 다르기 때문에 이를 스케일링 하기 위하여 최대값 1에서 최소값 0 범위로 설정하는 MinMaxScaler 클래스를 적용하였다.

데이터 전처리 후 LSTM 출력 예측 모델 생성을 위하여 학습 모델에 대한 셀의 수(은닉층 노드 수), 학습률(Learning rate), 활성화 함수(Activation function), 최적화 함수(Optimizer), 배치 크기(Batch size) 등의 하이퍼 파라미터(Hyper parameter) 값을 입력하여야 한다. 하이퍼 파라미터 값에 대하여 정해진 수치는 없으며, 학습 데이터의 특성에 따라 달리 설정하여야 한다. 따라서 본 연구에서는 하이퍼 파라미터 값을 조정 및 반복 설정하여 최적값을 구하였고, 설정한 하이퍼 파라미터 값은 Table 1과 같다.

Table 1 The values of hyper parameter

Contents	Values	Remarks
Number of unit	32	
Learning rate	0.001	
Activation function	Hyperbolic tangent	
Cost function	Mean Absolute Error	
Optimizer	Adam	
Epochs	100	
Batch size	200	

2.4 출력 예측 모델 평가 방법

풍력발전단지의 출력 예측은 각 풍력터빈별 예측 모델에 예측하고자 하는 1일 단위로 SCADA에서 실제 측정된 풍력터빈별 풍속 및 풍향데이터를 입력하여 개별 풍력터빈의 출력을 예측하였고, 이를 합산하여 (식 1)과 같이 풍력발전단지의 출력을 예측하였다.

$$P_{p,F} = \sum_{N=1}^6 P_{p,tN} \quad (\text{식 1})$$

출력 예측 모델의 검증은 (식 1)에 따라 계산된 풍력발전단지 출력 예측 데이터와 실제 측정된 풍력발전단지 출력 데이터를 비교하는 방법으로 검증하였다.

구체적인 출력 예측 모델 검증 방법으로 다음의 3가지의 항목을 이용한 실제 측정값과의 비교검증을 수행하였다. 첫 번째로 일반적으로 오차의 크기를 분석하기 위한 방법으로 활용되는 평균 제곱근 편차(RMSE, Root Mean Square Error)를 (식 2)에 따라 적용하였고, 두 번째로는 오차의 상대 크기를 파악하기 위하여 정규화된 평균 제곱근 편차(NRMSE, Normalized Root Mean Square Error)를 이용한 방법을 (식 3)에 따라 적용하였다. 여기에서, 데이터 정규화를 위한 $\sigma_{p,WF}$ 는 풍력발전단지 설비용량 12.3 MW를 적용하였다. 마지막으로 비교 데이터간의 데이터 상관성 분석을 위하여 상관계수(R, Correlation coefficient) 분석을 (식 4)에 따라 수행하였다.

$$RMSE = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n (P_{p,F,i} - P_{a,F,i})^2 \right)} \quad (\text{식 2})$$

$$NRMSE = \frac{1}{\sigma_{p,WF}} \sqrt{\frac{1}{n} \left(\sum_{i=1}^n (P_{p,F,i} - P_{a,F,i})^2 \right)} \times 100 \quad (\text{식 3})$$

$$R = \frac{\sum_{i=1}^n [(P_{p,F,i} - \bar{P}_{p,F}) \cdot (P_{a,F,i} - \bar{P}_{a,F})]}{\sqrt{\sum_{i=1}^n (P_{p,F,i} - \bar{P}_{p,F})^2 \cdot \sum_{i=1}^n (P_{a,F,i} - \bar{P}_{a,F})^2}} \quad (\text{식 4})$$

2.5 출력 예측 민감도 분석 방법

2.5.1 풍속 오차 추정

풍속 오차에 따라 출력 예측 오차에 미치는 영향을

분석하기 위하여 출력 예측 모델에 입력되는 풍속 데이터에 임의로 풍속 오차를 발생시켰다. 풍속 오차는 실제 측정된 풍속 대비 1 %, 3 %, 5 %, 10 %, 15 %로 설정하였고, 풍속 오차는 위의 오차 범위 내에서 정규분포의 확률로 오차가 발생하는 것으로 가정하였다. 이는 단순히 균등분포 확률로 오차를 추정하는 것보다 풍속 오차는 실제값을 기준으로 정규분포 확률로 발생한다고 가정하는 것이 실제 풍속 예측에서 발생하는 오차 형태를 추정할 수 있을 것이라 판단하였기 때문이다.

예를 들면 Fig. 5와 같이 실제 측정된 10분 평균 풍속이 10 m/s 이고 이에 대한 10 % 풍속 오차를 가정하는 경우, 10 m/s의 10 %의 오차인 9 m/s 또는 11 m/s를 랜덤하게 정규분포의 평균 풍속으로 가정하였다. 정규분포의 평균 풍속이 9 m/s 라고 하면, 정규분포의 표준편차는 2σ 값이 정규분포의 평균 풍속에 대한 10 % 가 되도록 0.45 m/s로 설정하였다. 즉, 해당 평균 풍속의 $\pm 10\%$ 범위 내에서 정규분포 내 95 %의 확률로 풍속 오차를 발생하였다.

정규분포를 이용하여 발생시킨 1 %, 3 %, 5 %, 10 %, 15 %의 풍속 오차가 출력 예측 기간 동안 제대로 추정되었는지 검증하기 위하여, 실제 각 풍력터빈에서 측정된 풍속과 비교는 (식 5)에 따라 각 풍력터빈별 실제 일평균 풍속 $\bar{\sigma}_{ws,t,d}$ 으로 정규화한 평균 제곱근 편차값에 대한 평균으로 확인하였다.

$$NRMSE = \frac{1}{\sigma_{ws,t,d}} \sqrt{\frac{1}{n} \sum_{i=1}^n (WS_{e,t,i} - WS_{a,t,i})^2} \times 100 \quad (\text{식 } 5)$$

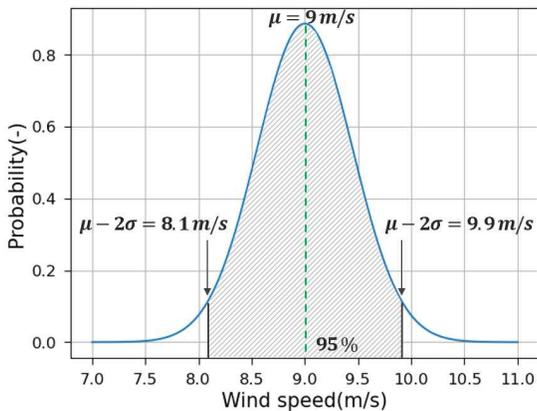


Fig. 5 The assumption example of 10 % wind speed error for 10 m/s by using normal distribution

2.5.2 출력 예측 민감도 분석 방법

각 풍력터빈에서의 실제 측정풍속과 2.5.1절에 따라 추정한 풍속 오차 1 % 부터 15 % 까지 결과를 2.3절에서 생성한 출력 예측 모델에 입력데이터로 활용하였을 때, 풍속 오차가 점차 커짐에 따라 출력 예측 오차에 미치는 영향을 분석하고자 하였다.

개별 풍력터빈에 대한 출력 예측 오차는 (식 6)에 따라 정규화한 평균 제곱근 편차값에 대한 평균으로 분석하였다.

$$NRMSE = \frac{1}{\sigma_{p,t}} \sqrt{\frac{1}{n} \sum_{i=1}^n (P_{p,t,i} - P_{a,t,i})^2} \times 100 \quad (\text{식 } 6)$$

여기에서, $\sigma_{p,t}$ 는 개별 풍력터빈의 설비용량 2.05 MW를 적용하였다.

풍력발전단지 전체에 대한 출력 예측 오차는 (식 3)에 따라 풍력발전단지 설비용량 12.3 MW로 정규화한 평균 제곱근 편차값에 대한 평균으로 분석하였다.

3. 풍력발전단지 출력 예측 및 민감도 분석 결과

3.1 대상 풍력발전단지 및 데이터

본 연구에 활용한 대상 풍력발전단지 및 데이터는 유럽 OpenAIRE 프로그램에 따라 개발되어 유럽 핵연구기구(CERN)에서 운영 중인 ‘제노도(Zenodo)’ 라는 범용 개방형 플랫폼에서 제공하는 켈마쉬(Kelmarsh) 풍력발전단지의 SCADA 자료 중 2020년도 1년간의 자료를 활용하였다.[8]



Fig. 6 The layout of kelmarsh wind farm

켈마쉬 풍력발전단지는 영국 내륙의 노샘프턴 해슬베르흐 인근에 위치한 풍력발전단지로서 Senvion사의 MM92 2.05 MW 풍력터빈 6기가 설치되어 2014년부터 상업 운영 중인 12.3 MW 규모의 풍력발전단지이다.(Fig. 6)[9]

본 연구에 활용한 켈마쉬 풍력발전단지의 1년치 SCADA 자료 내에는 각 풍력터빈의 풍속, 풍향, 나셀 위치, 출력, 로터 회전속도, 로터 베어링 온도, 기어 오일 압력 등 총 298개의 운전 데이터가 포함되어 있었지만 실제 딥러닝 분석에 사용한 데이터는 풍력터빈의 출력과 밀접한 관련을 갖는 풍속, 풍향데이터만을 학습 데이터(Train data)로 활용하고 출력데이터를 목표 데이터(Target data)로 활용하였다. 이는 다양한 종류의 데이터들을 이용하여 출력 예측 모델을 학습하기 보다는 최소한의 데이터를 이용하여 최적화된 출력 예측 모델을 생성하는 것이 향후 연구에서 효율적으로 활용 될 수 있을 것으로 판단하였기 때문이다.

출력 예측 기간은 데이터 분석에 활용한 2020년 1년치 SCADA 데이터 중 2020. 10. 01.부터 2020. 12. 31.일까지 3개월 동안 수행했으며, 1일 단위의 출력 예측은 10분 및 1시간 간격으로 예측하였다.

3.2 출력 예측 모델 검증

2.3절에서 생성한 LSTM 기반 출력 예측 모델의 반복 학습이 수행됨에 따른 손실함수 결과값은 Fig. 7과 같다. 반복 학습이 100회 까지 수행되는 동안 안정적으로 손실함수 결과값이 수렴하는 경향을 보이기 때문에 반복 학습이 제대로 수행된 것으로 판단하였다. 1

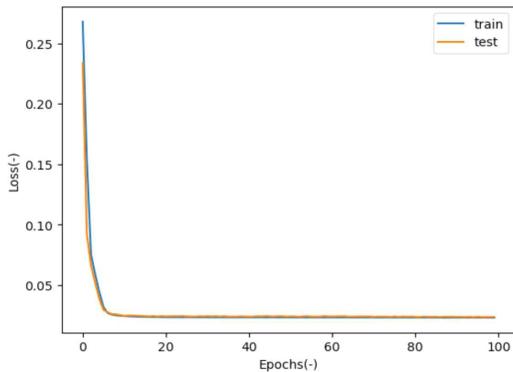


Fig. 7 The example of loss of train and test results(Training period : 2020. 01. 01.~2020. 09. 30. / Prediction period : 2020. 10. 01)

일 단위로 이동하며 예측을 수행하기 위하여 출력 예측을 수행한 2020. 10. 01.부터 2020. 12. 31.까지 92일 기간 중 비정상 데이터 수집기간을 제외한, 81일의 기간에 대해 각 풍력터빈별 81개의 모델을 생성하여 총 486개의 예측 모델을 생성하였다.

출력 예측 모델의 검증은 위에서 생성한 예측 모델에 측정된 풍속 및 풍향값을 입력하여 예측한 출력 예측값을 실제 측정된 출력값과 비교하는 방법으로 수행하였다. 이에 대한 구체적인 예측 모델의 검증 방법은

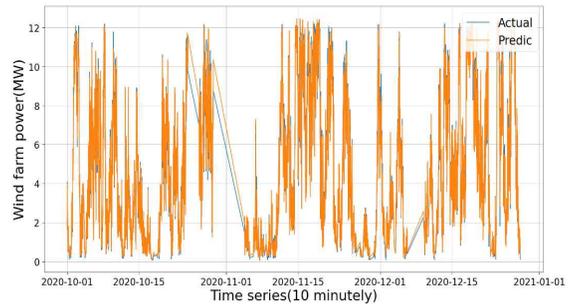


Fig. 8 The comparison of actual and prediction wind farm power for 10 minutely average data

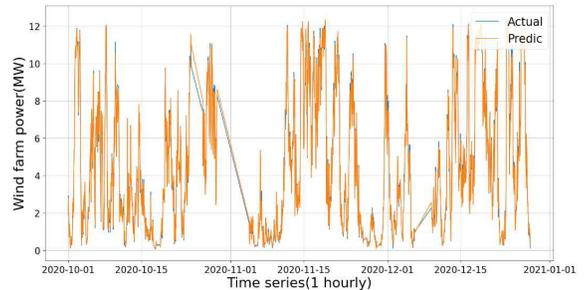


Fig. 9 Comparison between actual and predicted wind farm power using hourly average data

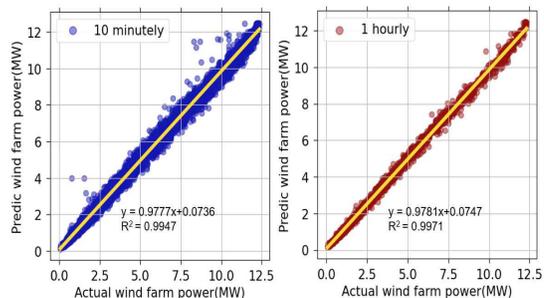


Fig. 10 Correlation between actual and predicted wind farm power using 10-minute and hourly average data

위의 2.4절에 따라 수행하였다.

실제 풍력발전단지의 출력과 Fig. 8 및 Fig. 9와 같이 출력 예측기간 전체 시계열로 비교하였을 때 전체적으로 유사한 트렌드를 확인하였고, 두 비교 데이터 간의 상관관계를 분석한 결과 Fig. 10과 같이 10분 단위 예측 데이터에서는 0.9947, 1시간 단위의 예측 데이터에서는 0.9971의 매우 높은 상관관계를 확인하였다.

Fig. 11은 출력 예측을 수행한 81일 기간 동안, 1일 단위로 10분 및 1시간 간격 출력 예측값을 (식 2)에 따라 분석한 평균 제공근 편차의 일별 평균을 나타낸 결과이다.

Fig. 12는 (식 3)에 따라 분석한 정규화된 평균 제공근 편차의 일별 평균을 나타낸다. 이를 전체 평균하면 전체 풍력발전단지 용량의 약 2.0% (10분 간격) 및 1.5% (1시간 간격) 예측 오차율을 갖는 것으로 분석하였다.

또한 Fig. 13의 상관계수 분석결과 1일 단위 10분 및 1시간 간격 상관계수 평균값이 약 0.987 및 0.992의 수치로 출력 예측값과 실제값이 매우 높은 상관관계를 갖고 있음을 확인하였다. 하지만, 2020.11.08.일 및 2020.11.26.일에는 낮은 평균 제공근 편차 및 정규화된 평균 제공근 편차를 보임에도 다른 예측기간 대비 낮은 상관계수 값이 확인된다. 이는 해당 예측일의 10분 단위 예측 데이터의 수가 73개 및 47개로 전체 예측기간 동안의 일별 평균 예측 데이터 수 119.5개에 비해 데이터 수가 충분치 않았고, Fig. 14와 같이 이 중 소수의 예측값과 실측값과의 차이가 큰 데이터로 인한 영향으로 낮은 상관계수 값을 보이는 것으로 분석된다.

정규화된 평균 제공근 편차 및 상관계수에 따른 예측 모델의 검증 결과는 Table 2 및 Table 3과 같이 도출되었고, 결과적으로 출력 예측 모델은 실제값과 유사한 출력이 예측되도록 성능을 발휘하는 것으로 검증을 완료하였다.

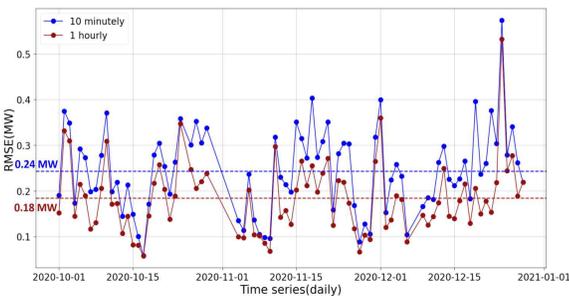


Fig. 11 Comparison of RMSE between actual and predicted wind farm power for 10-minute and hourly average data

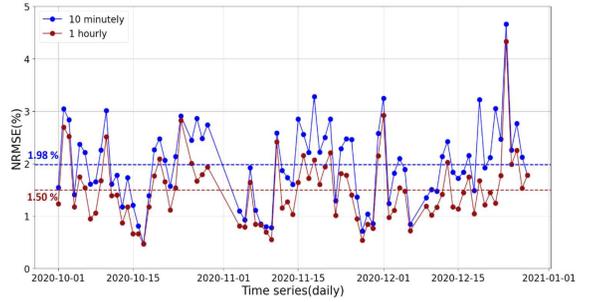


Fig. 12 Comparison of NRMSE between actual and predicted wind farm power for 10-minute and hourly average data

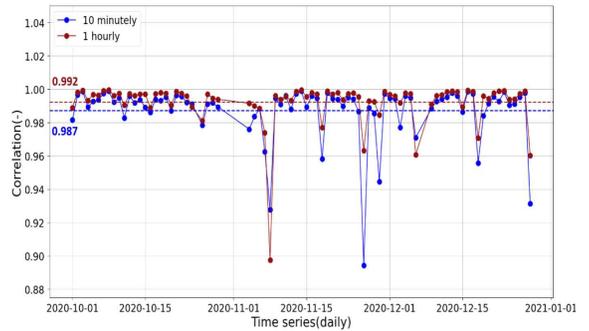


Fig. 13 Comparison of correlation coefficient between actual and predicted wind farm power for 10-minute and hourly average data

Table 2 The comparison of predicted and actual power for daily 10-minute averages during 2020.10.01.~2020.12.31.

Results	Daily averaged (10-minute)			Remarks
	RMSE [MW]	NRMSE [%]	Correlation [-]	
Average	0.24	1.98	98.73	
Minimum	0.06	0.48	89.43	
Maximum	0.57	4.66	99.89	

Table 3 The comparison of predicted and actual power for daily 1-hour averages during 2020.10.01.~2020.12.31.

Results	Daily averaged (1-hour)			Remarks
	RMSE [MW]	NRMSE [%]	Correlation [-]	
Average	0.18	1.50	99.22	
Minimum	0.06	0.47	89.76	
Maximum	0.53	4.33	99.96	

LSTM 딥러닝 신경망 모델을 이용한 풍력발전단지 풍속 오차에 따른 출력 예측 민감도 분석

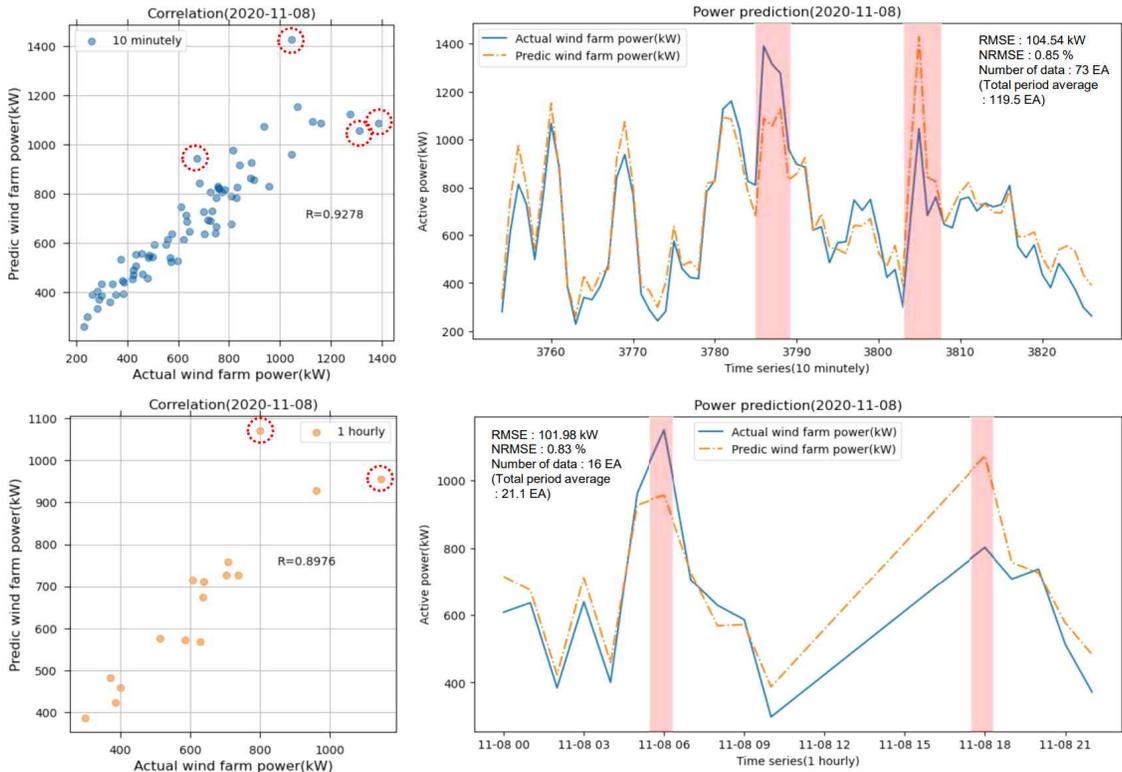


Fig. 14 Comparison of between actual and predicted wind farm power for 10-minute and hourly average data on 2020.11.08. (Red circle and shading: data with a large difference between actual and predicted values)

3.3 출력 예측 민감도 분석 결과

3.3.1 풍속 오차 추정 결과

2.5.1절의 풍속 오차 추정방법에 따라 1%, 3%, 5%, 10%, 15% 범위의 풍속 오차를 풍력발전단지 출력 예측기간 2020. 10. 01.부터 2020. 12. 31.까지 발생하였다. Fig. 15는 이 중 1일 동안의 풍력발전단지에서의 풍속 오차 발생 예시를 나타낸다. 가정된 풍속 오차가 커질수록 정규분포에 따른 풍속 오차 범위도 커지기 때문에 실제 풍속 보다 풍속의 변화폭이 점점 크게 변화하는 양상을 보여준다.

Table 4 및 Table 5는 각 풍력터빈 및 풍력발전단지에서 10분 간격 및 1시간 간격의 랜덤 풍속 오차를 1%부터 15% 까지 발생시켰을 때, 실제 풍속과의 일별 정규화된 평균 제곱근 편차를 평균 처리한 결과이다.

Table 4에서 볼 수 있듯이 10분 간격에서의 랜덤

풍속 오차 결과를 살펴보면 각 풍력터빈에서 가정된 랜덤 풍속 오차가 1%에서 15%까지 풍속 오차를 가정하였을 때, 최소 약 1.14%에서 최대 약 17.26%까지 실제 풍속 오차가 발생하는 것으로 분석되어 전반적으로 가정된 풍속 오차에 따라 풍력터빈의 풍속 오차가 발생된 것으로 확인하였다.

또한 Fig. 15와 Table 4에서의 10분 간격 풍속 오차 분석결과 보다 Fig. 16과 Table 5에서의 1시간 간격 풍속 오차 분석결과에서 실제 풍속 오차가 낮게 발생되고 있음을 확인할 수 있는데, 이는 10분 간격의 데이터를 1시간 간격으로 평균화하여 처리하는 과정에서 발생하는 결과로 판단된다.

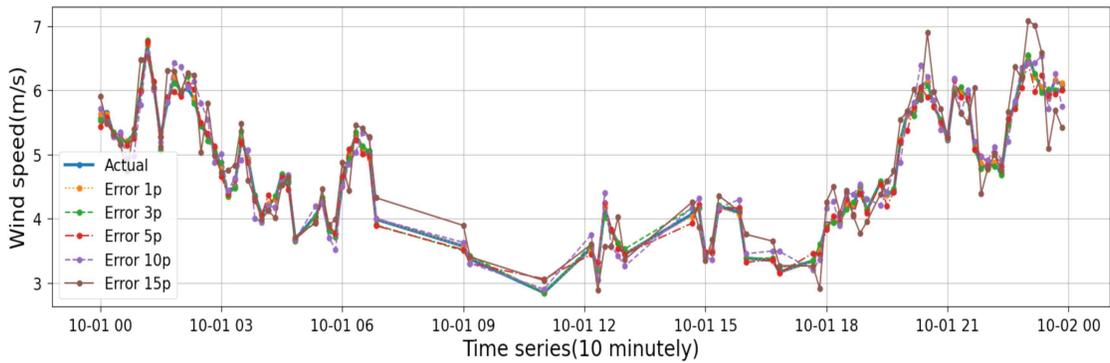


Fig. 15 The generation of 10 minutely wind speed error of wind farm for each 1 %, 3 %, 5 %, 10 %, 15 % of assumed wind speed error in 1 day (2020. 10. 01)

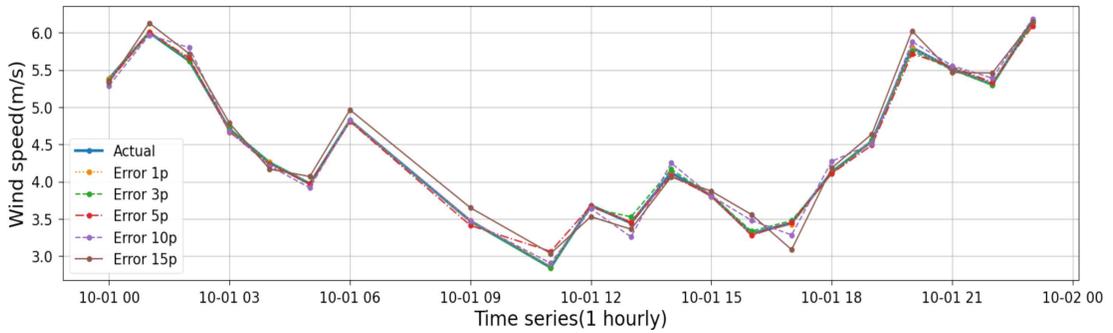


Fig. 16 The generation of 1 hourly wind speed error of wind farm for each 1 %, 3 %, 5 %, 10 %, 15 % of assumed wind speed error in 1 day (2020. 10. 01)

Table 4 The results of daily averaged NRMSE of 10 minutely averaged wind speed error for each wind turbines and wind farm during 2020.10.01.~2020.12.31.

Assumed wind speed error (%)	NRMSE of wind speed (10 minutely)					
	WT1	WT2	WT3	WT4	WT5	WT6
1	1.14	1.14	1.15	1.13	1.15	1.15
3	3.44	3.44	3.44	3.43	3.44	3.43
5	5.67	5.70	5.73	5.70	5.68	5.75
10	11.45	11.30	11.41	11.49	11.47	11.43
15	17.25	16.90	17.18	17.16	17.12	17.26

Table 5 The results of daily averaged NRMSE of 1 hourly averaged wind speed error for each wind turbines and wind farm during 2020.10.01.~2020.12.31.

Assumed wind speed error (%)	NRMSE of wind speed (1 hourly)					
	WT1	WT2	WT3	WT4	WT5	WT6
1	0.51	0.50	0.50	0.50	0.50	0.51
3	1.50	1.51	1.51	1.48	1.46	1.52
5	2.47	2.48	2.48	2.55	2.45	2.53
10	5.07	4.92	4.91	5.06	4.97	4.94
15	7.53	7.44	7.51	7.70	7.48	7.58

3.3.2 출력 예측 민감도 분석 결과

풍속 오차에 따른 출력 예측 민감도를 분석하기 위하여 3.3.1의 풍속 오차 1 % 부터 15 % 까지의 데이터를 이용하여 풍력터빈 및 풍력발전단지 출력 예측을 수행하였다.

Fig. 17 및 Fig. 18은 출력 예측 기간 중 1일에 대한 10분 및 1시간 간격 출력 예측 결과이다. 출력 예측 모델에 입력되는 풍속 오차가 커질수록 예측 출력의 결과값이 실제 출력값과 편차가 커지는 양상을 보여준다. 전체 출력 예측기간에 대한 출력 예측 오차는 2.5.2의 출력 예측 민감도 분석 방법에 따라 분석하여 Table 6 및 Table 7과 같이 도출하였다.

Table 6을 보면, 풍속 오차가 발생하지 않은 경우 풍력발전단지 정규화된 평균 제공근 편차가 1.98 %를 갖는 반면, 15 % 풍속 오차가 발생 시에는 10.96 %의 출력 예측 변동을 보인다. 각 풍력터빈 출력 예측 오

차에서는 풍속 오차가 발생하지 않은 경우 최소 3.13 %인 반면 15 % 풍속 오차 발생 시에는 최대 17.26 %의 출력 예측 변동을 보였다. 각 풍력터빈의 정규화된 평균 제공근 편차 보다 풍력발전단지의 정규화된 평균 제공근 편차가 작게 분석된다. 이는 풍력발전단지의 정규화된 평균 제공근 편차를 구할 때, 각 풍력터빈의 출력 예측 값을 모두 합산하여 풍력발전단지 예측 출력 결과를 계산하는데 각 풍력터빈별로 과대 예측 또는 과소 예측한 결과가 합산되며 상쇄되는 영향에서 기인한다.

예를 들면, Fig. 19는 풍속 오차가 없다고 가정한 10분 단위 출력 예측 결과 중 2020.10.01.일 21:10분 부터 2020.10.01.일 21:50분 까지 5개의 출력 예측 샘플 데이터를 추출하여 (식 7) 및 (식 8) 에 따라 계산한 각 풍력터빈별 출력 예측 오차(Err_t) 및 풍력발전단지의 출력 예측 오차(Err_F)를 나타낸다.

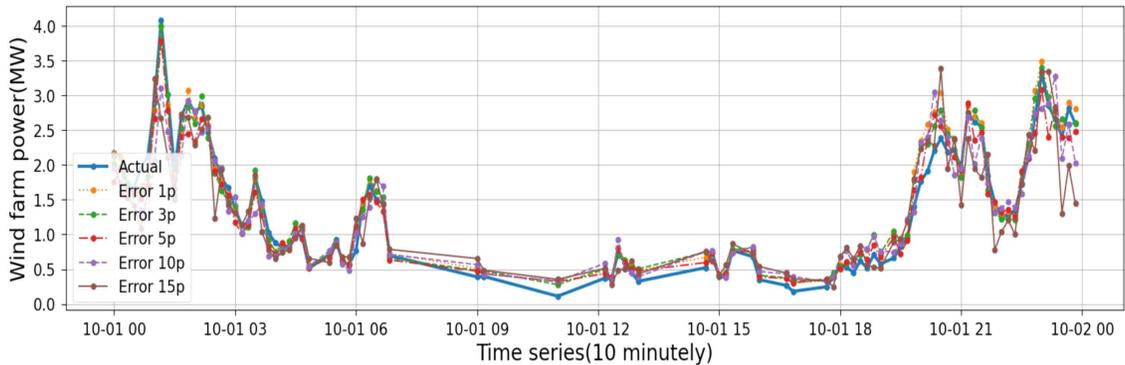


Fig. 17 The variation of wind farm power for 10 minutely wind speed error of wind farm for each 1 %, 3 %, 5 %, 10 %, 15 % of assumed wind speed error in 1 day (2020. 10. 01)

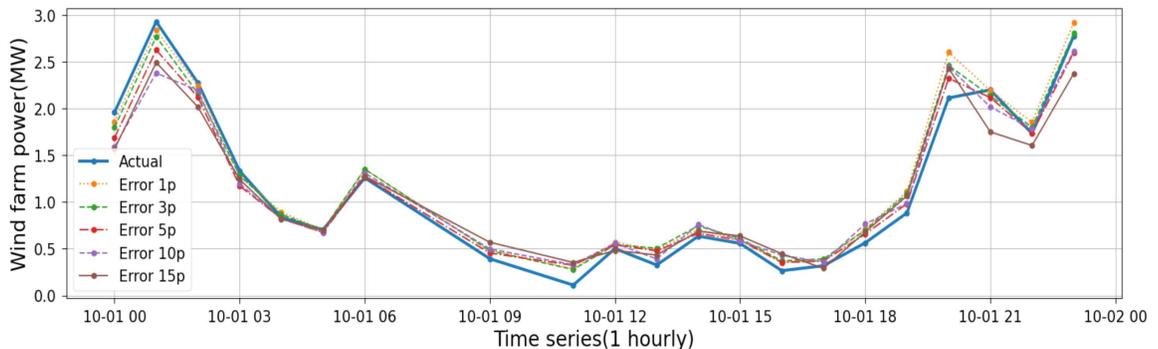


Fig. 18 The variation of wind farm power for 1 hourly wind speed error of wind farm for each 1 %, 3 %, 5 %, 10 %, 15 % of assumed wind speed error in 1 day (2020. 10. 01)

$$Err_t = P_{p,t} - P_{a,t} \quad (\text{식 7})$$

$$Err_F = P_{p,F} - P_{a,F} \quad (\text{식 8})$$

각 풍력터빈별 출력 예측 오차는 출력량의 과대 예측 또는 과소 예측에 따라 양의 값 또는 음의 값을 갖으며, 이는 풍력발전단지 예측 오차 합산 시에 상쇄됨으로써 개별 풍력터빈의 출력 예측 오차 보다 풍력발전단지의 출력 오차가 작게 계산된다.

Table 7은 10분 간격의 출력 예측 값을 1시간 간격으로 평균 처리하였을 때, 출력 예측 민감도 분석결과를 나타낸다. 풍속 오차가 발생하지 않은 경우 풍력발전단지 출력 예측 오차 1.50 %를 갖는 반면, 15 % 풍속 오차가 발생 시에는 9.65 %의 출력 예측 변동을 보였다. 각 풍력터빈 출력 예측 오차에서는 풍속 오차가 발생하지 않은 경우 최소 1.95 %인 반면 15 %의 풍속 오차

발생 시에는 최대 14.63 %의 출력 예측 변동을 보였다.

이는 Table 6의 10분 간격 출력 예측 민감도 분석 결과보다 낮은 출력 예측 변동 수치로, 10분 간격의 출력 예측 수치를 1시간 단위로 평균화 처리하는 과정에서 출력 예측 변동이 줄어들기 때문이다.

Table 6 및 Table 7의 출력 예측 민감도 분석결과를 이용하여 결과적으로 전력거래소에서 현재 운영 중인 ‘예측제도’의 SMP 정산금 인센티브를 받기 위한 풍력발전단지 출력 예측 오차 기준 8 %를 달성하기 위하여 요구되는 풍속 예측 오차율을 추정하였다.

Fig. 20을 보면 10분 간격 출력 예측 시에는 풍력발전단지 출력 예측 오차 8 % 달성을 위하여, 요구되는 풍속 예측 오차는 약 11.0 %로 추정하였다.

마찬가지로 Fig. 20의 1시간 간격 출력 예측 시, 풍력발전단지 출력 예측 오차 8 % 달성을 위하여 요구되는 풍속 예측 오차는 약 12.7 %로 추정하였다.

Table 6 The results of daily averaged NRMSE of 10 minutely averaged power prediction for each wind turbines and wind farm during 2020.10.01.~2020.12.31.

Assumed wind speed error (%)	NRMSE of power prediction (10 minutely)						
	Wind farm	WT1	WT2	WT3	WT4	WT5	WT6
0	1.98	3.40	3.48	3.84	3.23	3.56	3.13
1	2.05	3.87	3.86	3.99	3.59	3.80	3.35
3	2.61	4.58	4.92	5.11	5.14	4.72	4.38
5	4.18	7.33	7.51	6.79	6.50	7.90	5.91
10	7.23	11.11	12.46	10.24	11.69	11.75	10.70
15	10.96	15.10	17.26	15.30	15.90	15.99	17.25

Table 7 The results of daily averaged NRMSE of 1 hourly averaged power prediction for each wind turbines and wind farm during 2020.10.01.~2020.12.31.

Assumed wind speed error (%)	NRMSE of power prediction (1 hourly)						
	Wind farm	WT1	WT2	WT3	WT4	WT5	WT6
0	1.50	2.10	2.21	2.67	2.03	2.40	1.95
1	1.53	2.56	2.62	2.69	2.40	2.63	2.10
3	1.97	2.54	2.92	3.46	3.57	3.02	2.55
5	3.47	5.32	5.14	4.51	3.57	6.25	3.62
10	6.10	6.69	8.32	5.43	7.48	8.50	7.47
15	9.65	8.50	11.92	9.99	9.85	11.81	14.63

LSTM 딥러닝 신경망 모델을 이용한 풍력발전단지 풍속 오차에 따른 출력 예측 민감도 분석

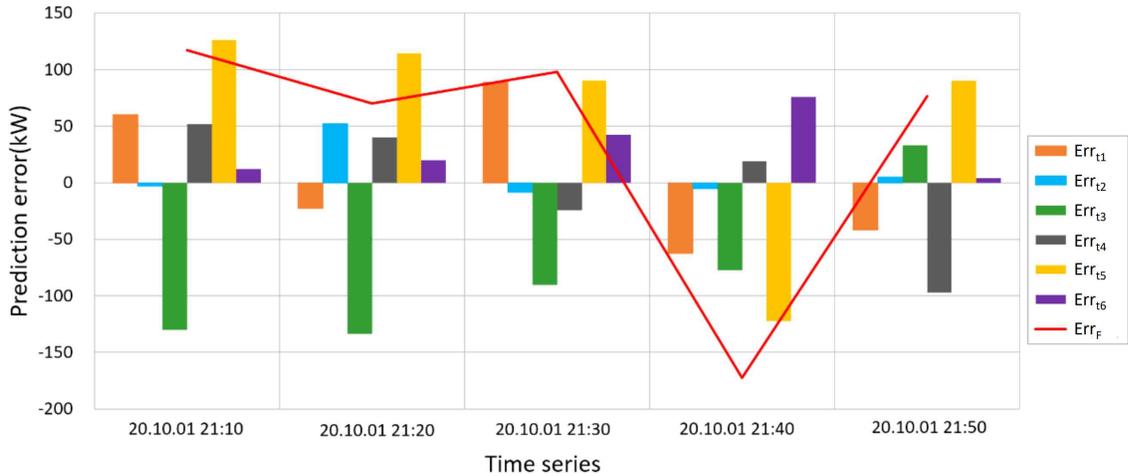


Fig. 19 Example of time series predicted power output error for each wind turbines and wind farm during 2020.10.01. 21:10~2020.10.01. 21:50 (Assumed wind speed error 0 %)

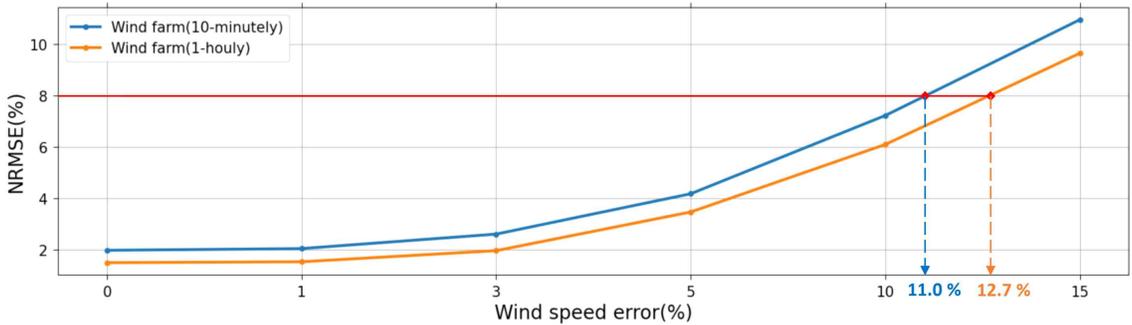


Fig. 20 The variation of wind farm power for wind speed error of wind farm for each 1 %, 3 %, 5 %, 10 %, 15 % of assumed wind speed error during 2020.10.01.~2020.12.31.

4. 결론

실제 풍력발전단지의 1년치 SCADA 데이터를 이용하여 LSTM 풍력터빈 출력 예측 모델을 생성하고, 이를 기반으로 총 81일 기간 동안 개별 풍력터빈 및 풍력발전단지 출력 예측을 10분 및 1시간 간격으로 수행하였다. 추가적으로 풍속 오차에 대한 출력 예측 민감도 분석을 수행하여 목표 출력 예측 오차 달성을 위해 요구되는 풍속 예측 오차율을 추정하였다.

본 연구에 활용한 LSTM 출력 예측 모델의 오차는 정규화된 평균 제곱근 편차를 이용하여 10분 간격 및 1시간 간격 실제 풍력발전단지 출력값과 비교분석 하였을 때, 평균 1.98 % 및 1.50 %로 확인하였으며, 상

관관계 분석결과 평균 98.73 및 99.22로 매우 높은 상관관계를 갖는 것으로 확인함으로써 LSTM 출력 예측 모델의 정확도를 검증하였다.

또한 추가적으로 풍속 오차에 대한 출력 예측 민감도 분석을 위하여 정규분포를 이용한 임의의 풍속 오차를 1 %, 3 %, 5 %, 10 %, 15 %로 발생시키고, 이를 생성한 LSTM 풍력터빈 출력 예측모델에 입력하여 결과를 분석하였다.

10분 간격의 풍력발전단지 출력 예측 오차 분석결과, 1 % 풍속 오차 발생 시 2.05 % 부터 15 % 풍속 오차에서는 10.96 % 출력 예측 오차를 도출하였고, 1시간 간격의 풍력발전단지 출력 예측 정확도 분석결과, 1 % 풍속 오차 발생 시 1.53 % 부터 15 % 풍속

오차에서는 9.65 % 출력 예측 오차를 도출하였다.

‘예측제도’의 출력 예측 오차 기준 8 % 달성을 위하여 요구되는 풍속 예측 오차는 10분 간격 풍속 예측 시 11.0 % 및 1시간 간격 풍속 예측에서는 12.7 %로 추정하였다. 따라서 이를 이용하여 향후 풍속 예측 모델의 정확도 목표를 설정하는 기준으로 활용할 수 있을 것으로 기대된다.

후기

본 연구는 산업통상자원부의 재원으로 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구결과(풍력발전제어시스템 국산화 기술개발, 과제번호 : 20213030020230 및 국제인증제도 대응 풍력 성능평가 요소 기술 개발 및 인증 체계 구축, 과제번호 : 2020020010)입니다.

참고문헌

- [1] Status of domestic wind turbine installation, Korea wind energy industry association, 2023
- [2] The 10th Basic Electricity Supply and Demand Plan, Ministry of trade, industry and energy, 2023
- [3] Electricity market operation rules, Korea power exchange, 2020
- [4] Kanna B.,S.N. Singh, 2012, “AWNN-Assisted Wind PowerForecasting Using Feed-Forward Neural Network,” IEEETrans. Sustain. Energy
- [5] Kim.J.J.,Baek.J.H.,Park.S.H.,Hur.J., 2021, “A Study on a Short-term Wind Power Output Forecasting using LSTM Method”, The korea institute of electrical engineers, 157-158
- [6] A Study on Wind Power Forecasting Using LSTM Method, Heungseok Lee · Kyuhan Kim · Heemyung Jeong · Hwaseok Lee · Hyungsu Kim · June Ho Park
- [7] A Study on Prediction of Wind Power Based on Deep-Learning Using Weather Data, Eun-Ji Kim · Taek-Kie Lee · Kyu-Ho Kim
- [8] <https://en.wikipedia.org/wiki/Zenodo>
- [9] <https://www.blue-energyco.com/our-projects/kelmarsh/>