IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 November 2023, accepted 18 November 2023, date of publication 27 November 2023,
date of current version 6 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3337210

RESEARCH ARTICLE

A Probabilistic Estimation of Transmission Congestion for Mitigating Wind Power Curtailments

MINJU LEE, (Student Member, IEEE), AND JIN HUR, (Senior Member, IEEE)
Department of Climate and Energy Systems Engineering, Ewha Womans University, Seoul 03760, Republic of Korea

Corresponding author: Jin Hur (jhur@ewha.ac.kr)

This work was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2018R1A6A1A08025520, and in part by Korea Electric Power Corporation under Grant R21XO01-1.

ABSTRACT The high penetration of renewable energy sources presents new challenges to power systems owing to the variability and uncertainty of renewable energy. When transmission congestion occurs, wind curtailment is required to ensure stability. Short-term wind power forecasting and transmission expansion are methods for managing the curtailment caused by transmission congestion. In addition, for transmission expansion, it is necessary to identify transmission congestion by assuming wind power scenarios. In this study, a short-term wind power forecasting model based on rolling long short-term memory (R-LSTM) is proposed for a 24-hour seasonal transmission congestion estimation. R-LSTM uses a recursive strategy to improve forecasting accuracy. Moreover, a method for probabilistic estimation of transmission congestion using seasonal hourly scenarios is proposed to account for variability and uncertainty. Wind power and electricity demand scenarios were generated using kernel density estimation (KDE) and the Metropolis-Hastings algorithm, a type of Markov chain Monte Carlo (MCMC) sampling. The R-LSTM model and transmission congestion estimation use empirical data from wind farm A and electricity demand on Jeju Island, where renewable energy is rapidly increasing. The results show that R-LSTM has a higher accuracy than long short-term memory (LSTM). The probability estimation results of transmission congestion revealed the season and time period with the highest probability of transmission congestion occurrence for each congestion components, as well as the lowest seasons and time periods.

INDEX TERMS Identifying transmission congestion, kernel density estimation, long short-term memory, Metropolis-Hastings algorithm, probabilistic estimation.

I. INTRODUCTION

Various policies have been implemented worldwide with the aim of achieving carbon neutrality to address climate change. Accordingly, the share of renewable energy in electricity generation is increasing. South Korea aims to increase the penetration of renewable energy resources by 30.6% by 2036 [1]. Jeju Island in South Korea aims to become carbon-free island JEJU by 2030 (CFI 2030). Owing to the rapid expansion of renewable energy to implement CFI 2030, Jeju Island has the highest renewable energy resource penetration in South Korea [1], [2].

However, as the penetration level of renewable energy sources increases, the characteristics of renewable energy, such as variability and uncertainty, pose new challenges to power systems. Furthermore, the probability of transmission congestion is increasing because of the differences in construction periods between renewable energy generators and transmission networks. This increased transmission congestion leads to higher levels of renewable energy curtailment.

Curtailment is particularly prevalent in wind power generation, leading to economic and energy-efficiency losses.

The associate editor coordinating the review of this manuscript and approving it for publication was Dinesh Kumar.

Curtailment is defined as "A reduction in the output of a generator from what it could otherwise produce given available resources, typically on an involuntary basis" [3]. Transmission congestion is a common cause of curtailment that occurs when the transmission capacity is insufficient to accommodate the available wind power output [3], [4]. In South Korea, delays in the construction of transmission lines, such as the East Jeju – Wando line, have led to curtailment [1].

Accurate wind power forecasting is one way to manage curtailments. Improved wind power forecasting reduces the curtailment of power systems with high wind penetration [3]. Additionally, the higher the accuracy, the better the reliability and stability of the power system. Wind power forecasts can be divided into short-, mid-, and long-term categories based on their timeframes. The short-term wind power forecasts are 30 min to 6 h ahead of each other. Short-term wind power output forecasting plays a crucial role in power system scheduling and congestion management. Therefore, several physical, statistical, and machine learning based methods have been studied to improve the accuracy of renewable energy output forecasting. Among these, deep learning methods can perform complex nonlinear modeling on large amounts of data and are well suited for time-series forecasting.

Many studies [5], [6], [7], [8], [9] have conducted short-term wind power forecasting using deep learning methods based on recurrent neural networks (RNN) and convolutional neural networks (CNN). In [5], [6], [7], and [8], diverse long short-term memory (LSTM) models were used to forecast short-term wind-power outputs. In [5], an error analysis was performed by adjusting parameters such as the number of epochs, batch size, and number of hidden layers. In [6], an improved LSTM was proposed that used the variational mode decomposition method to decompose the wind signal into its long-term components. In [7], a stacked RNN with a parametric sine activation function was proposed. In [8], a CNN-LSTM-LightGBM model was proposed using the MAPE-RM algorithm. In [9], a temporal convolutional network model was used to forecast the wind power output, which showed high accuracy. As reviewed above, previous studies have improved the performance using various LSTM model structures, data preprocessing through decomposition, and ensemble methods. However, these studies did not consider data update method, nor did they apply wind power forecasting results to power system analysis. In this paper, we compare the forecasting performance based on the application of a data update method called the recursive strategy and utilize seasonal short-term wind power forecasting results for a 24-hour transmission congestion estimation.

In addition, transmission expansion is a strategy for mitigating curtailment owing to transmission congestion. It is necessary to identify transmission congestion through a power system analysis. Existing power system analysis methods do not consider the variability and uncertainty of renewable energy. Therefore, methods that consider variability and uncertainty, such as probabilistic methods and the utilization

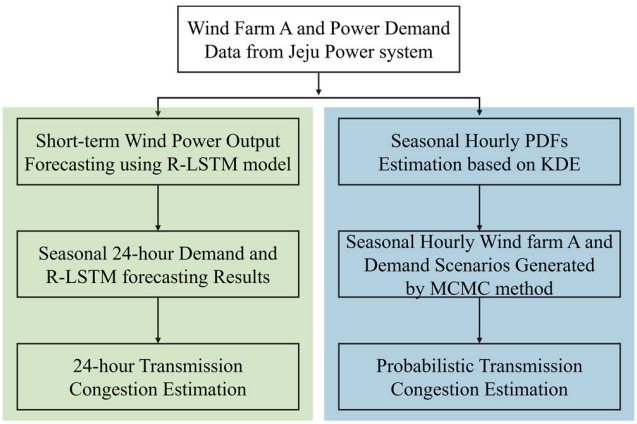


FIGURE 1. Overall flowchart of the seasonal hourly transmission congestion estimation.

of scenarios, have been studied [10], [11], [12], [13], [14]. Reference [11] presents multiple scenarios for analyzing the impact of wind siting and curtailment owing to transmission congestion based on these scenarios. In [12], short-term renewable energy curtailment prediction was conducted by detecting congestion in distribution grids. In this study, short-term forecasting results are used to estimate 24-hour transmission congestion, aiming to better account for the variability and uncertainty of wind power generation. Additionally, probabilistic transmission congestion estimation is performed by generating seasonal hourly scenarios.

Based on a review and comparison of previous studies, we propose an algorithm for a short-term wind power forecasting model based on rolling long short-term memory (R-LSTM) for 24-hour seasonal transmission congestion estimation and a method for probabilistic estimation of transmission congestion using seasonal hourly scenarios to account for variability and uncertainty. An overall flowchart of the deterministic and probabilistic transmission congestion estimations is shown in Fig. 1. As shown in Fig. 1, we examined the R-LSTM model proposed by Devi et al. [15]. In [15], historical wind power variables were used; however, in the present study, we used other variables such as historical wind speeds and hours. Moreover, wind power forecasting values were used to conduct 24-hour seasonal transmission congestion estimation. Additionally, seasonal hourly scenarios were generated to estimate the probability of transmission congestion. The probability density functions (PDFs) of the wind farm and power demand were obtained using Kernel Density Estimation (KDE), and Markov chain Monte Carlo (MCMC) sampling was used to generate seasonal hourly scenarios. The main contributions of this study are summarized as follows.

- The R-LSTM model was utilized for short-term wind power output forecasting, and its performance was compared with LSTM and analyzed based on its model structure.
- The KDE-MCMC method was applied to generate seasonal hourly scenarios for electricity demand and wind farm.

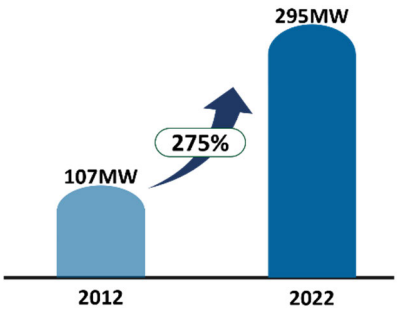


FIGURE 2. Installed wind power capacity in Jeju power system.

- Short-term wind power forecasting data were utilized to perform a 24-hour transmission congestion estimation.
- Probabilistic transmission congestion estimation was conducted using KDE-MCMC scenario data for wind power integration.
- The proposed methodology was applied to empirical data from wind farm A and the electricity demand on Jeju Island, which has a high penetration of renewable energy resources.

The remainder of this paper is organized as follows. In Section II, we discuss wind power generation and curtailment in the Jeju power system. Section III presents a short-term wind power forecasting model based on the R-LSTM model. In Section IV, we present a method for the probabilistic estimation of transmission congestion using KDE-MCMC sampling. Section V includes three components: short-term wind power forecasting based on R-LSTM, 24-hour transmission congestion estimation using R-LSTM, and probabilistic transmission congestion estimation using wind farm and electricity demand scenarios based on KDE-MCMC sampling. Finally, Section VI concludes the paper.

II. POWER SYSTEM OF JEJU ISLAND

In this study, wind farm A and power demand data from Jeju Island in South Korea were used. Jeju Island has a CFI 2030 target to meet its entire power demand with renewable energy by 2030 [2]. Owing to the rapid expansion of renewable energy to implement CFI 2030, Jeju Island has recently experienced an increase in the number of curtailments; therefore, improved wind power forecasting and transmission congestion estimations are necessary for management. The installed capacity of wind turbines on Jeju Island has increased 2.7 times, from 107 MW in 2012 to 295 MW in 2022, as shown in Fig. 2 [16]. The penetration of installed wind capacity increased from 5.3% in 2008 to 19% in 2022. In addition, wind power accounted for 19% of the generation mix in 2022.

Fig. 3 shows the seasonal average electricity demand in the Jeju power system from 2021 [17]. Peak demand mainly occurred between 16:00 and 22:00 in all seasons. The summer loads were higher than the others because the hot and humid weather increased the electricity demand for cooling.

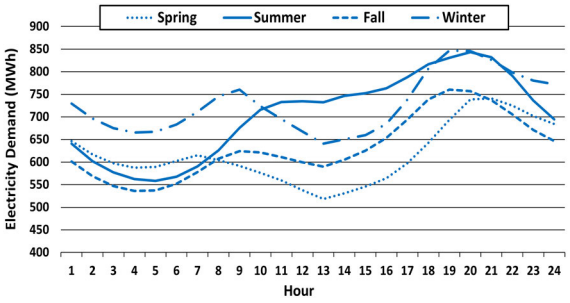


FIGURE 3. Seasonal average electricity demand in Jeju power system from 2021.

TABLE 1. Wind power curtailment in the Jeju power system.

Year	No. of Curtailment Event	Curtailment (MWh)	Curtailment rate (%)
2015	3	152	0.04
2016	6	252	0.05
2017	14	1300	0.24
2018	15	1366	0.25
2019	46	9223	1.66
2020	77	19449	3.24
2021	64	12016	-

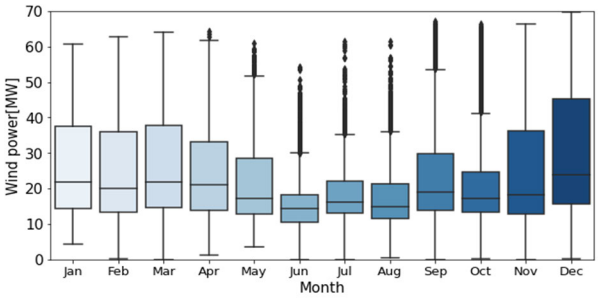


FIGURE 4. Monthly wind power boxplot in Jeju power system during 2021.

With a rapid increase in renewable energy in the Jeju power system, curtailment is necessary to protect the system from oversupply and transmission congestion. As listed in Table 1, the number of wind power curtailments in the Jeju power system has increased since 2015 [18]. Curtailments occur at dawn and are characterized by high supply and low demand. However, curtailment has occurred even during the daytime since 2017 because of the significant annual increase in the installed PV capacity to 39.6% in 2018 [17], [18], [19], [20]. In 2021, the curtailment was less than that in 2020 because it was supplied to the mainland using high-voltage direct current (HVDC); however, it still occurred 64 times with 12016 MWh of curtailment.

In addition, curtailment is expected to increase further to manage wind power in spring, fall, and winter, when there is a larger amount of wind power output, as shown in Fig. 4. Therefore, for efficient operation of the Jeju power system

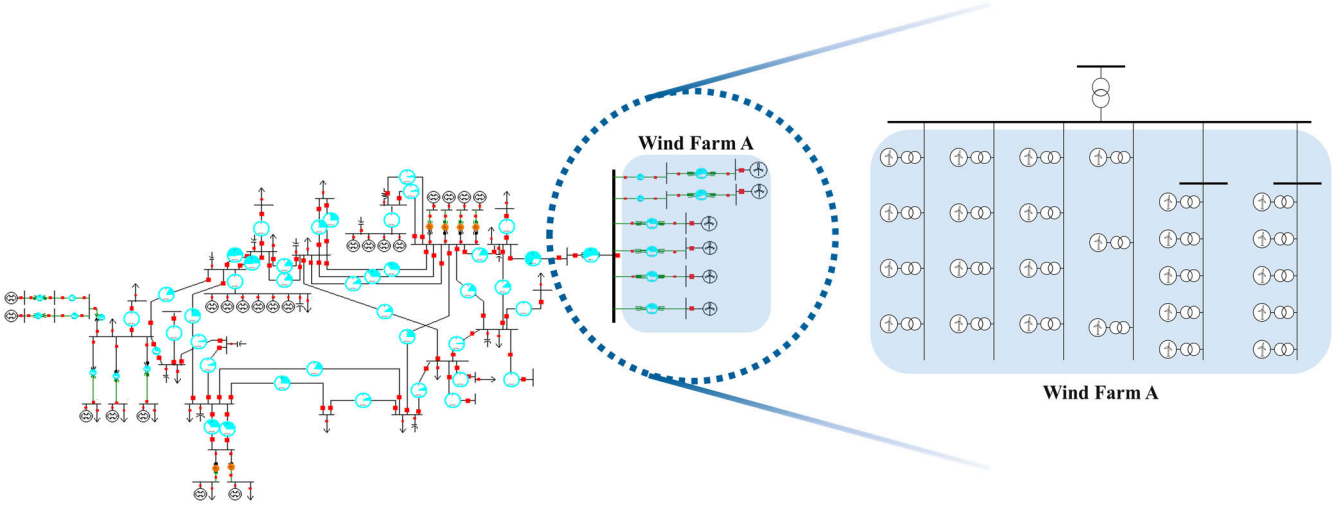


FIGURE 5. Diagram of the Jeju power system.

with high wind power penetration, it is necessary to improve the accuracy of wind power forecasts and estimate transmission congestion for effective curtailment management and estimation.

Fig. 5 shows a diagram of the Jeju power system. A short-term forecasting model and transmission congestion estimation were applied to wind farm A of the Jeju power system. The Jeju power system data used in this study comprise 74 buses, 46 generators, and 26 loads, with a total active load of 850 MW. Wind farm A consists of 25 wind turbine generators; in the diagram, three to five generators are represented on a single generator. In other words, a total of 60MW wind farm, consisting of 3MW 10 units and 2MW 15 units, is configured as an aggregated wind farm model.

III. PROPOSED A SHORT-TERM WIND POWER FORECASTING MODEL

A. LONG SHORT-TERM MEMORY

Machine learning can perform complex nonlinear modeling on large amounts of data and is well suited for time-series forecasting. Machine learning can use variables, such as wind speed and hours, to forecast wind power. Among deep learning models, the LSTM model was proposed by Hochreiter and Schmidhuber. The LSTM model improves the performance by preventing the vanishing gradient problem of the RNN [21]. The equations for the LSTM cell are as follows [21]:

$$i_t = \sigma\left(w_{xi}x_t + w_{hi}h_{t-1} + b_i\right) \quad (1)$$
$$o_t = \sigma\left(w_{xo}x_t + w_{ho}h_{t-1} + b_o\right) \quad (2)$$
$$f_t = \sigma\left(w_{xf}x_t + w_{hf}h_{t-1} + b_f\right) \quad (3)$$
$$\tilde{c}_t = \tanh\left(w_{xc}x_t + w_{hc}h_{t-1} + b_c\right) \quad (4)$$
$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (5)$$
$$h_t = o_t \cdot \tanh\left(c_t\right), \quad (6)$$

TABLE 2. Period for recursive strategy.

Period 1	Period 2	Period 3	Period 4
00:00 – 05:45	06:00 – 11:45	12:00 – 17:45	18:00 – 23:45

where i, o, and f denote the input, output, and forget gates, respectively, x_t denotes the input data, σ signifies the sigmoid function, w denotes the gate weight, b indicates the bias at each gate, c indicates the cell state, $\tilde{c}_t$ indicates the cell input activation vector, and h indicates the hidden state, which is also known as the output vector of the LSTM unit.

The LSTM uses gates based on a sigmoid function to allow cells to selectively transfer data. Input, forget, and output gates can be used to add or remove information from a cell state [22], [23]. Additionally, the memory cell remembers its state over time. A memory cell stores long-term dependencies, unlike an RNN that does not reflect long-term temporal dependencies [21].

B. ROLLING LONG SHORT-TERM MEMORY

The R-LSTM, proposed by Devi et al. [15], improves the accuracy of short-term wind power forecasting using a recursive strategy in the LSTM model. The existing R-LSTM model uses wind power variables; however, in this study, wind power, wind speed, and hour variables were used for forecasting. The main difference between the R-LSTM and LSTM is the use of a recursive strategy. The LSTM does not update the training set, preventing it from learning new data patterns. In contrast, R-LSTM conducts retraining in each period, which makes it computationally more time-consuming. However, this recursive strategy enables R-LSTM to learn new data patterns, thereby improving the forecasting accuracy.

The recursive strategy sets four periods by dividing 24 h of test data by 6 h, as shown in Table 2 [15]. The realized data

Algorithm 1 R-LSTM Model For Short-Term Wind Power Forecasting

Input: Wind speed, wind power output and hour data
Min-max normalization
Split data to training and test set
Convert data into pairs of feature variables and target variables
Training R-LSTM model with training data
loop:
  if *last forecasted period* $<$ *6 hours* **then**
    Short-term Wind power forecasting using R-LSTM model
  end
  if *Last forecasted period* $>$ *6hours* **then**
    Add realized data to training data
    Training R-LSTM model with updated training data
  end
  go to loop
Compare the forecasted values with measured values using the NMAE
end

for each period are included in the training data after the completion of each forecast period. Thus, R-LSTM was retrained to learn the latest wind power patterns, and short-term wind power forecasting was performed using the updated R-LSTM model [15].

The R-LSTM model conducts short-term wind power output forecasting based on Algorithm 1. Before entering the data into the model, normalization was required because the units and magnitudes of the variables were different. In this study, Min-Max normalization was applied to the wind power output, wind speed, and hour variables. The Min-Max normalization is defined as follows:

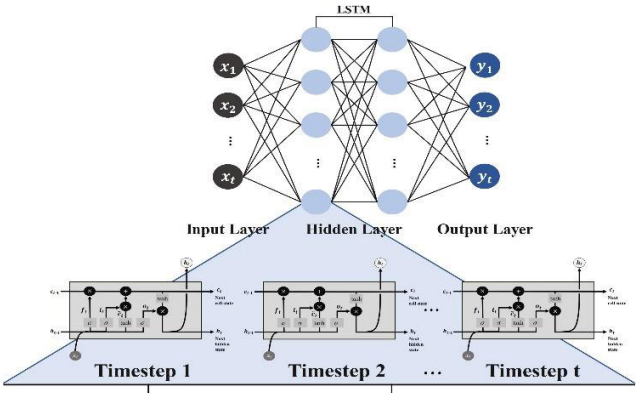
$$f(x) = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{7}$$

where x is a value, x_{min} is the minimum value of x, and x_{max} is the maximum value of x. This function unifies the minimum and maximum values of the data to 0 and 1.

The R-LSTM was trained using training data. If the latest input data were less than 6 h, short-term wind power forecasting was continuously performed using the R-LSTM. If the latest input data were more than 6 h, the realized wind power output was added to the training data; that is, after wind power forecasting for 6 h, the actual wind power output data were included in the training data [15]. The forecast values were evaluated using the normalized mean absolute error (NMAE), defined as

$$NMAE = \frac{1}{N}\sum_{i=1}^{N}\frac{|p_i - \hat{p}_i|}{Installed\ Capacity} \times 100\,(\%), \tag{8}$$

where p_i denotes the measured value, $\hat{p}_i$ denotes the forecast value, and N denotes the number of forecast samples. The NMAE is suitable for comparing renewable energy resources with different installed capacities.



	Feature	Target
Time step 1	$WS_1, Hour_1$	WP_1
Time step 2	$WS_2, Hour_2$	WP_2
⋮	⋮	⋮
Time step t	$WS_t, Hour_t$	WP_t

FIGURE 6. Process and structure of LSTM model.

TABLE 3. Classification of time period.

Time Period	
Overnight	10 p.m. to 6 a.m.
Morning	6 a.m. to 1 p.m.
Afternoon	1 p.m. to 5 p.m.
Evening	5 p.m. to 10 p.m.

Fig. 6 shows the process by which the datasets of the feature and target variables are learned using the LSTM model. The feature variables are wind speed and hour. The number of time steps was 24 because the 6-hour forecasting was performed in 15-minute intervals; therefore, each node had 24 LSTM units.

IV. METHODOLOGY OF PROBABILISTIC TRANSMISSION CONGESTION ESTIMATION

To estimate transmission congestion based on wind power output, the annual data for wind farm A and power demand data for Jeju Island from 2021 were utilized to generate scenarios classified by season and time. South Korea has four seasons: spring (March to May), summer (June to August), fall (September to November), and winter (December to February). Table 3 lists the classification criteria for the morning, afternoon, evening, and overnight periods [24].

To estimate the probability of transmission congestion, scenarios for wind farm A and electricity demand were generated using the process shown in Fig. 7 to consider uncertainty and variability. Upon classifying the data into 16 types based on the season and time, each PDF was estimated using the KDE method. Wind farm A and power demand scenarios were generated using the Metropolis-Hasting algorithm

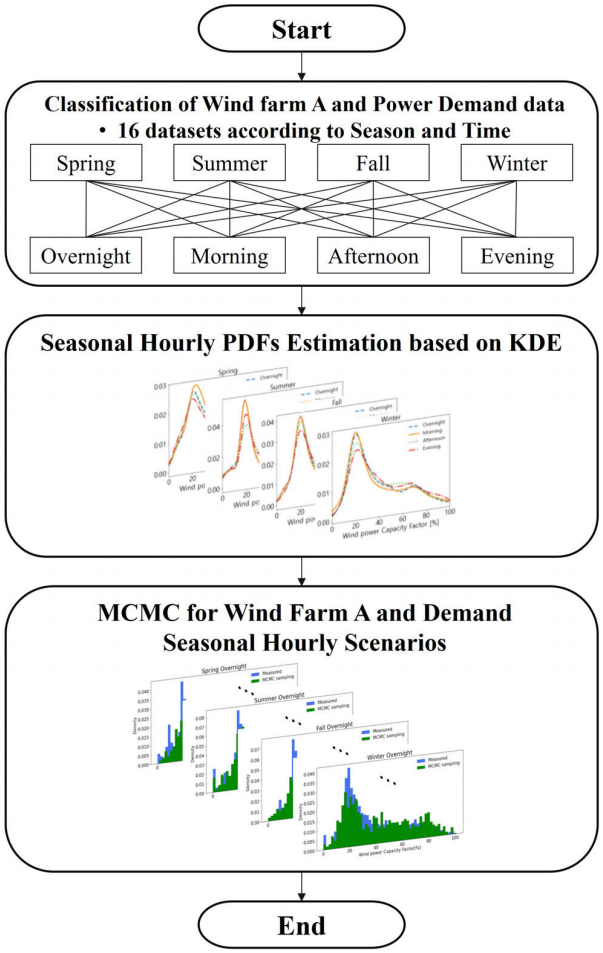


FIGURE 7. Flow chart for generating wind power output and demand seasonal hourly scenarios.

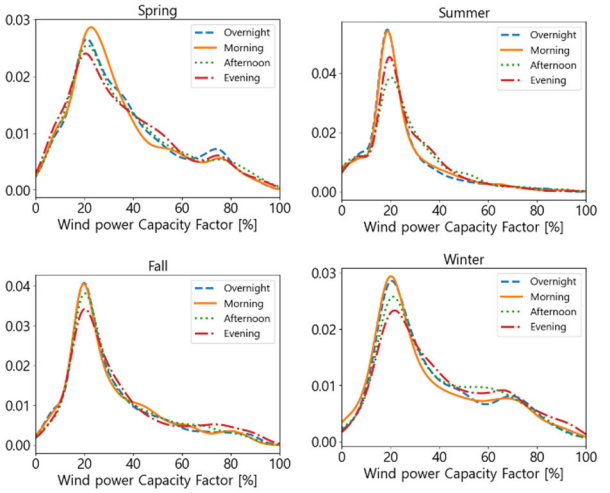
based on the estimated PDFs. The generated seasonal hourly scenarios were validated by comparing them with the measured data and using the NMAE to ensure their accuracy. Wind power density estimation and data sampling were conducted considering the capacity factor of wind farm A. The capacity factor was calculated to apply the wind power scenarios to the individual wind power plants in wind farm A.

A. KERNEL DENSITY ESTIMATION

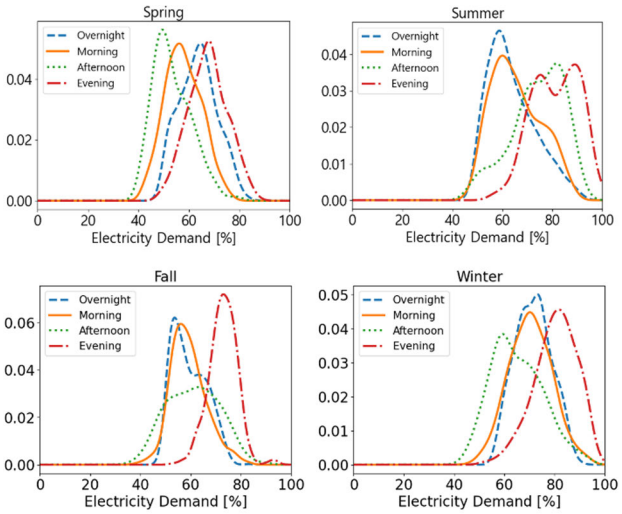
Nonparametric density estimation calculates the distribution directly from a given dataset. The method does not assume a specific distribution. Nonparametric density estimation provides flexibility in modeling complex and unknown distributions. The KDE performs better than the parametric approach in estimating the distribution of meteorological variables that change over different time periods because it does not assume a specific distribution [25].

KDE is a popular density-estimation method. The KDE is defined as:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{9}$$



(a) PDFs of wind farm A capacity factor during 2021



(b) PDFs of electricity demand during 2021

FIGURE 8. PDFs of wind farm A and electricity demand during 2021.

where x indicates a random variable; x_i represents the observed data; K denotes the kernel function; and h indicates the smoothing bandwidth. The kernel function is symmetric with an integral value of 1. KDE estimates the density distribution by averaging over a set of kernel functions $K(x - x_i)$ centered at each data point [26]. Bandwidth determines the width of the kernel and smoothing level.

In this study, a Gaussian function was applied as the kernel function and the optimal bandwidth was determined using Silverman's rule [27]. Sixteen PDFs of wind farm A and the power demand of the Jeju power system were estimated based on the season and time, as shown in Fig. 8. Based on the density estimation results, the estimated seasonal hourly PDFs of wind farm A exhibit a distribution similar to that of the boxplot statistics shown in Fig. 4. During summer, the wind power capacity factor was predominantly distributed below 40% compared with the other seasons. This implies that the wind power output of wind farm A is more likely to be lower during summer than during other seasons.

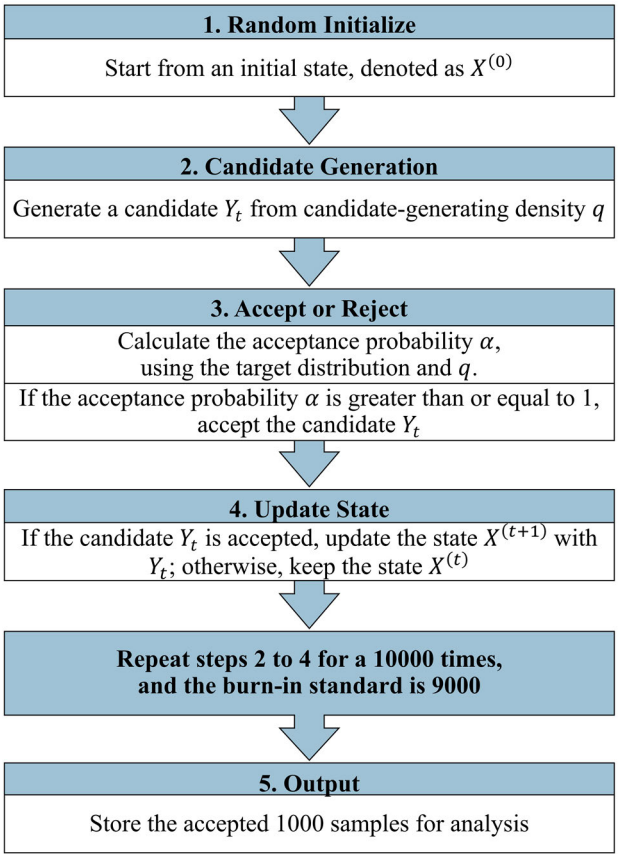


FIGURE 9. Metropolis-Hastings algorithm flow chart.

B. MARKOV CHAIN MONTE CARLO SAMPLING: METROPOLIS-HASTINGS ALGORITHM

MCMC methods generate sets of samples from PDFs estimated using KDE with a Gaussian kernel. These methods have Markov properties because they generate the next sample based on the current sample. In addition, they exhibit Monte Carlo properties because they estimate a large number of random samples from a distribution [28].

Among MCMC techniques, the Metropolis-Hastings algorithm is versatile and simple. A flowchart of the Metropolis-Hastings algorithm is shown in Fig. 9. This method starts with a random initialization $x^{(0)}$ of the initial state $X^{(0)}$ and generates samples of candidate Y_t from a candidate-generating density q.

The Metropolis-Hasting algorithm shows the transition from the value $x^{(t)}$ of $X^{(t)}$ to the candidate Y_t of $X^{(t+1)}$ based on the acceptance-rejection algorithm with acceptance probability α [29], [30]. The algorithm for the Metropolis-Hastings is as follows [29], [30]:

$$\begin{aligned} &Given\ X^{(t)} = x^{(t)} \\ &Generate\ Y_t \sim q\left(y \mid x^{(t)}\right) \\ &Take\ X^{(t+1)} = \begin{cases} Y_t\ with\ probability\ \alpha\left(x^{(t)}, Y_t\right) \\ x^{(t)}\ with\ probability\ 1-\alpha\left(x^{(t)}, Y_t\right) \end{cases} \end{aligned}$$

TABLE 4. NMAE of Metropolis-Hastings algorithm.

Period	NMAE (%)			
	Overnight	Morning	Afternoon	Evening
Spring	3.36	5.91	2.92	1.96
Summer	0.97	0.59	2.29	3.84
Fall	2.26	1.12	0.53	4.54
Winter	0.13	2.08	0.73	2.10

$$where$$

$$\alpha(x, y) = \min\left[\frac{\pi(y)\, q(x \mid y)}{\pi(x)\, q(y \mid x)}, 1\right] \tag{10}$$

In this study, the random initialization $x^{(0)}$ was 0, and a normal distribution was used as the candidate-generating density q. π is a target distribution, which is a PDF estimated using the KDE.

Sample generation was performed 10000 times, and the burn-in standard was 9000. A total of 1000 samples from 9000 – 10000 samples were analyzed. Consequently, 1000 samples were generated for each of the 16 datasets based on season and time, and a transmission congestion estimation was conducted with 16000 samples, as shown in Fig. 10. In addition, comparing the samples with the measured data, the resulting NMAE values, as listed in Table 4, indicate high sampling accuracy using the Metropolis-Hasting algorithm across the 16 probability distributions.

C. PROBABILISTIC TRANSMISSION CONGESTION ESTIMATION

The probability of transmission congestion was calculated using the generated seasonal hourly scenarios, as shown in Fig. 11. Historical data for wind farm A and power-demand data for Jeju Island from 2021 were analyzed. Seasonal hourly scenarios were generated based on KDE-MCMC sampling to consider the characteristics of wind power and demand, such as variability and uncertainty. Power system data, demand, and wind power scenarios were loaded to calculate steady-state power flow and conduct an N-1 contingency analysis. Specific lines were determined when their line load rates exceeded their thermal limits, based on seasonal hourly scenarios [12]. The probability of transmission congestion for specific lines was calculated by determining the number of line overloads.

V. CASE STUDY: FORECASTING AND TRANSMISSION CONGESTION ESTIMATION RESULTS

A. A SHORT-TERM WIND POWER OUTPUT FORECASTING FOR JEJU ISLAND'S WIND FARM

R-LSTM was used to conduct short-term wind power output forecasting. The R-LSTM model was evaluated using Keras and TensorFlow, and its performance varied based on the number of time steps, hidden layers, nodes, and epochs [5], [31]. The model was fixed at 24-time steps for 20 epochs.

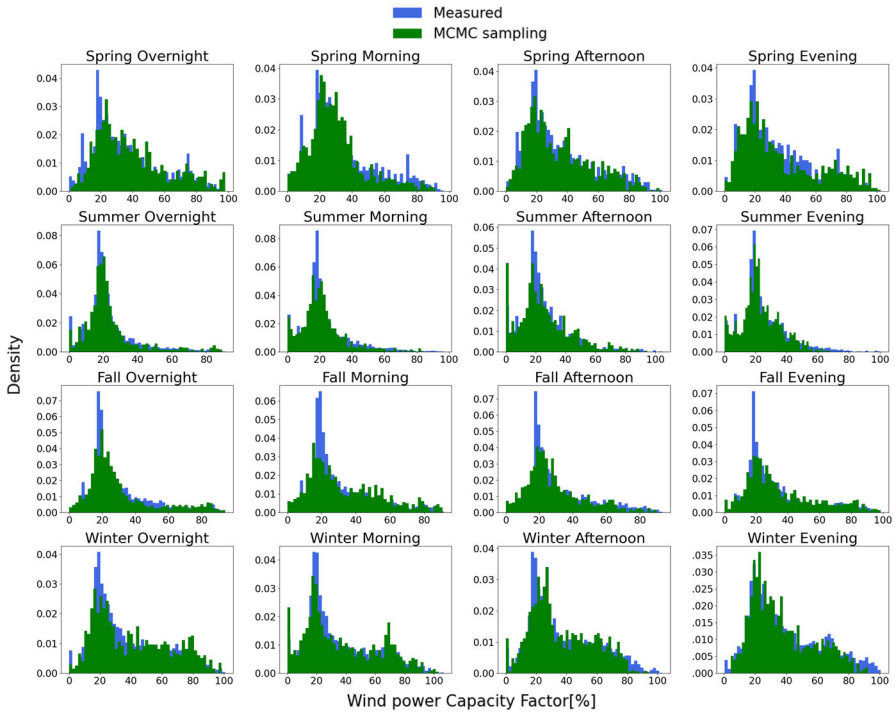


FIGURE 10. Comparison of wind power output scenarios using the Metropolis-Hastings algorithm and measured data in 2021.

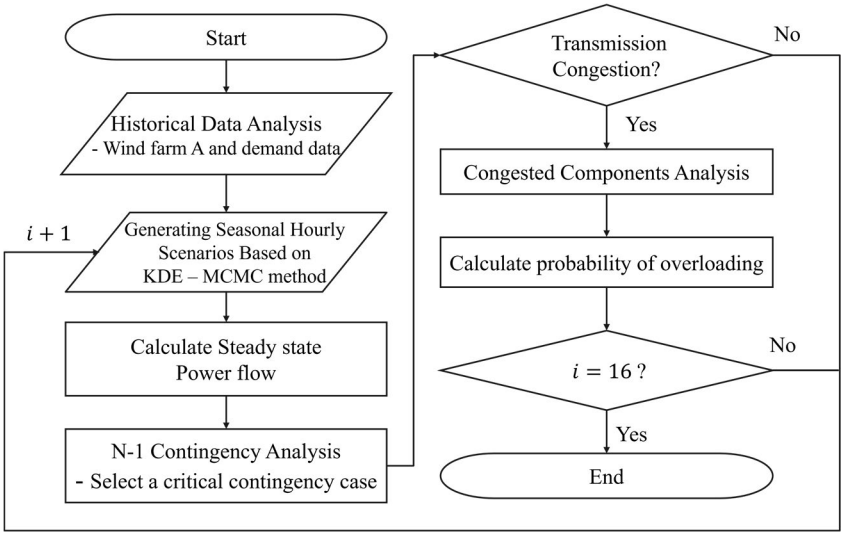
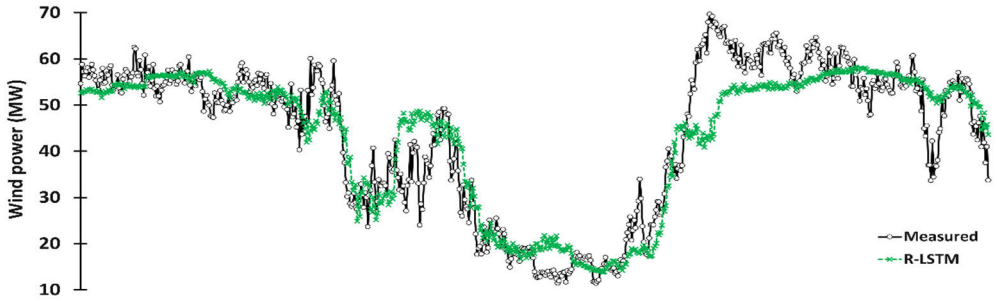


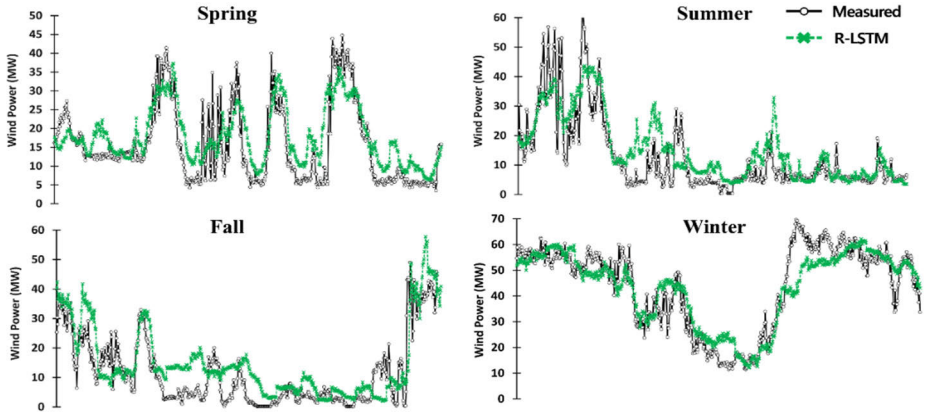
FIGURE 11. Algorithm for calculating the probability of transmission congestion using the seasonal hourly scenarios based on KDE-MCMC sampling.

To compare the performance of R-LSTM based on different hyperparameters, data from wind farm A from January 1 to December 24, 2021, were used as training data, and data from December 25 to 31, 2021 were used as test data. The NMAE values were compared by varying the number of hidden layers and nodes, as listed in Table 5. For three hidden layers, the NMAE value was 8.39%, indicating the highest performance (Fig. 12 (a)).

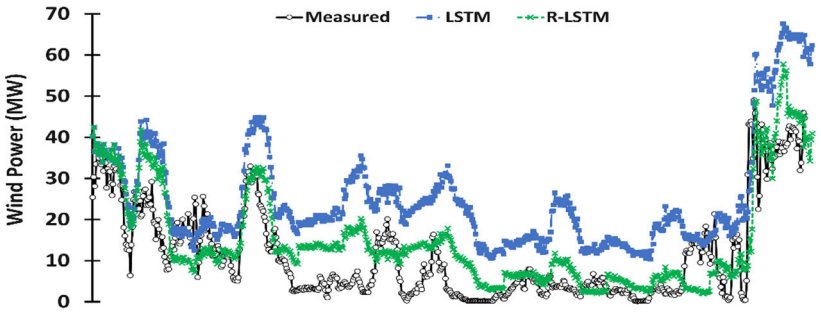
In addition, the performances in each season were compared. For spring, data from March 1 to May 24 were used as training data, and data from May 25 to 31 were used as test data. Similarly, for the other seasons, the last week of the month was used as test data. The R-LSTM and LSTM models conduct short-term wind power forecasting based on the same structure with three hidden layers. The results of the R-LSTM and LSTM were compared and analyzed.



(a) Wind power forecasting of R-LSTM model for Dec 25–31, 2021 (3 hidden layers)



(b) Seasonal wind power forecasting of R-LSTM model during 2021



(C) Wind power forecasting of R-LSTM and LSTM models during the fall of 2021

FIGURE 12. Wind power forecasting of R-LSTM and LSTM models during 2021.

TABLE 5. NMAE of R-LSTM model by changing parameter.

Hidden layer	Node	NMAE (%)
1	24	10.29
2	[48, 24]	9.95
3	[48, 24, 24]	8.39

TABLE 6. Seasonal NMAE of R-LSTM model.

Season		Spring	Summer	Fall	Winter
NMAE (%)	LSTM	10.73	14.84	22.86	9.34
	R-LSTM	8.60	9.27	10.59	9.14

Fig. 12(b) presents the measured data and forecast results of the R-LSTM model for each season. Fig. 12(c) illustrates a graph of the forecasting results from November 24 to 31 in the fall, when the NMAE difference between the R-LSTM and LSTM was the greatest.

In Fig. 12(a), discrepancies are observed between the measured and forecasted wind power data. When analyzing the measured data, an average wind power output

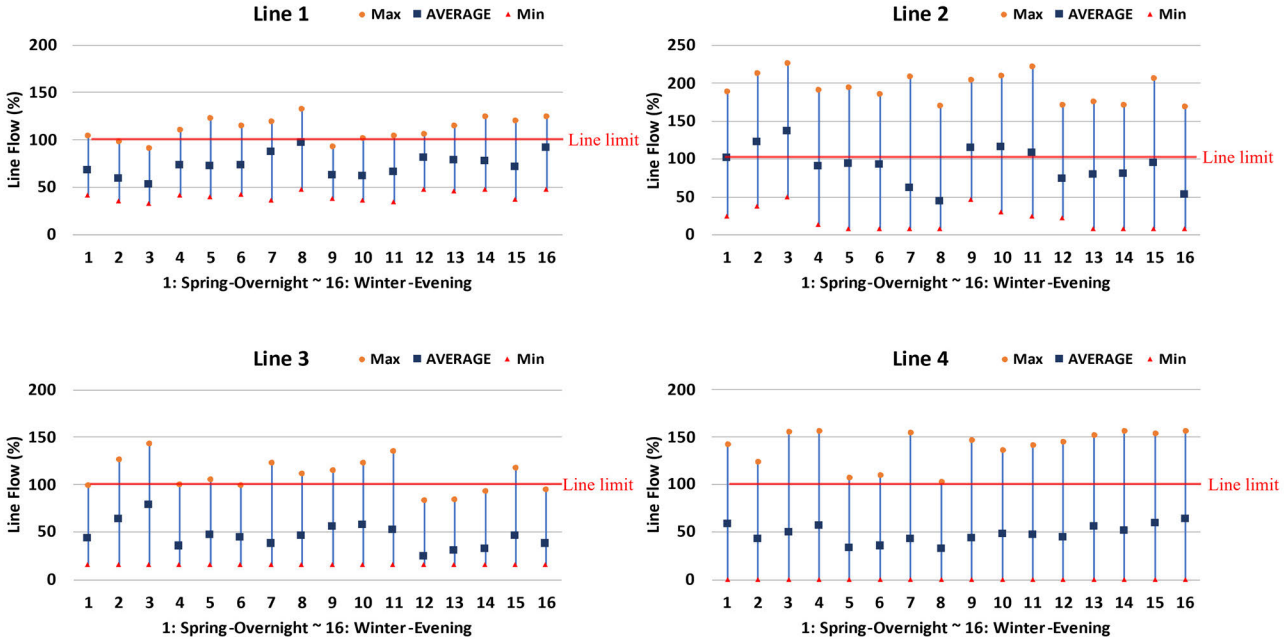


FIGURE 13. Min, Max, and Average of line loads based on KDE-MCMC wind farm A and demand scenarios.

TABLE 7. Probability of transmission congestion caused by KDE-MCMC wind farm A and demand scenarios.

Classification		Probability of overloading (%)					
		Line 1	Line 2	Line 3	Line 4	Line 5	Line 6
Spring	Overnight	0.17	49.30	0.00	15.07	29.63	24.87
	Morning	0.00	78.17	6.00	2.47	11.70	8.10
	Afternoon	0.00	**86.17**	**24.7**	8.47	19.07	15.13
	Evening	1.90	35.60	0.1	13.73	27.33	23.40
Summer	Overnight	7.80	46.20	1.00	0.70	4.80	3.33
	Morning	5.87	43.40	0.03	1.00	6.13	3.30
	Afternoon	23.77	15.23	1.00	5.20	13.03	9.60
	Evening	**48.90**	6.33	0.60	0.20	2.00	1.00
Fall	Overnight	0.00	68.43	0.5	5.57	11.73	10.17
	Morning	0.13	71.93	1.2	8.77	19.60	16.00
	Afternoon	0.63	56.87	8.43	10.17	17.90	16.57
	Evening	1.50	15.90	0.00	4.30	11.77	9.80
Winter	Overnight	4.57	24.67	0.00	15.10	25.57	22.37
	Morning	4.47	26.00	0.00	11.40	23.03	20.17
	Afternoon	5.80	46.23	2.70	16.77	32.20	26.87
	Evening	29.40	8.13	0.00	**21.53**	**34.17**	**31.87**

of 36 MW was recorded at this range of wind speed. The 25th percentile was 28.59 MW, the 50th percentile was 35.88 MW, and the 75th percentile was 43.86 MW. In other words, for this range of wind speeds, 50% of the data fell within the range of 28.59 MW to 43.86 MW. However, in the interval where inconsistencies were observed, an average power output of 66.46 MW was recorded at this wind speed. When using R-LSTM to forecast this

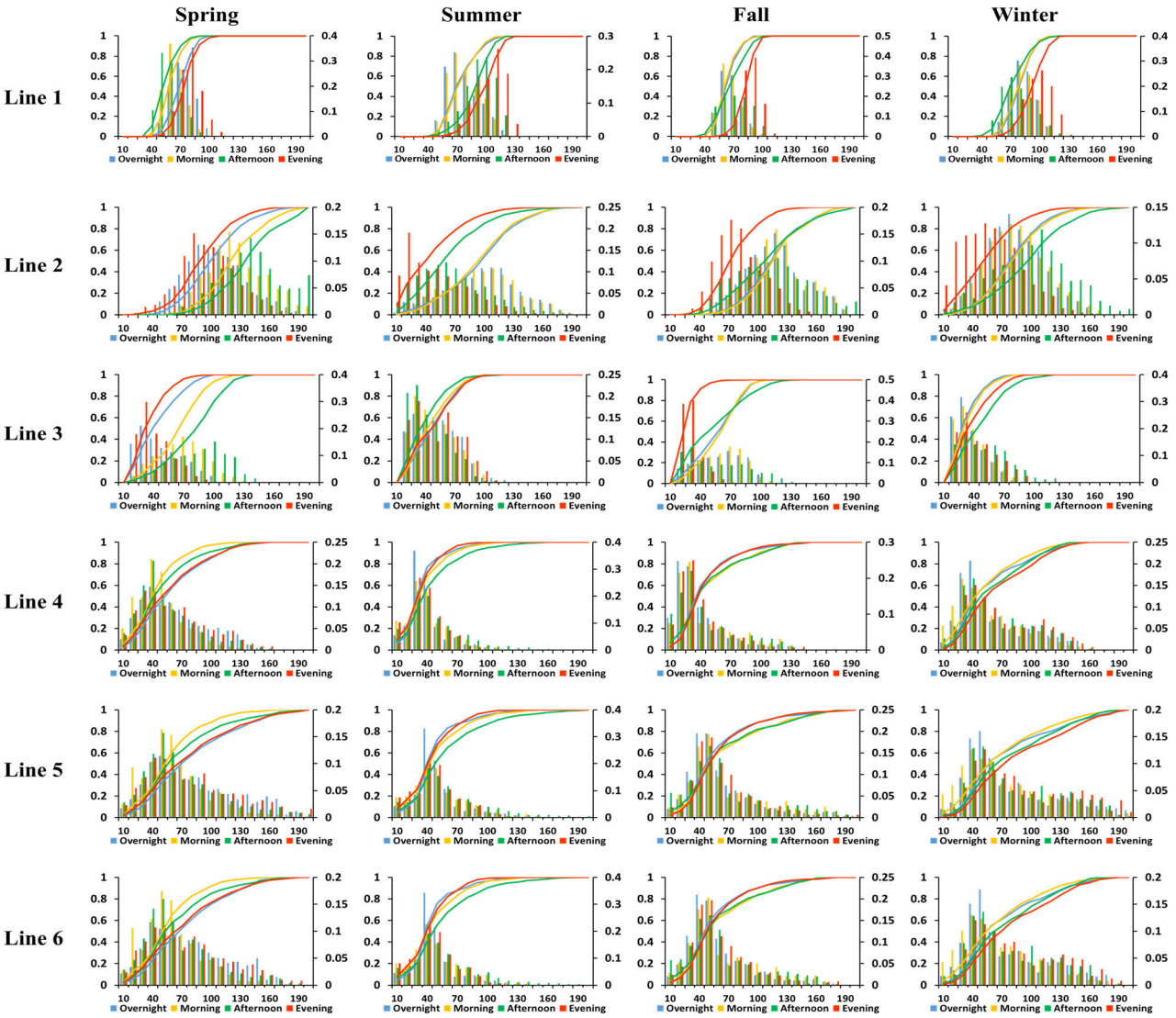


FIGURE 14. Histogram and CDF of line flow based on KDE-MCMC wind power scenarios.

interval, such abnormal relationships between wind speed and wind power were not present in the previous period, resulting in no performance difference between R-LSTM and LSTM.

Table 6 lists the NMAE of the R-LSTM and LSTM models for each season. The NMAE value was lowest in spring and highest in the fall. We observed a significant performance improvement in the R-LSTM model, which learned a new pattern between wind speed and power. Unlike the discrepancies in Fig. 12(a), a new relationship between wind power and wind speed emerged from the measured values in the previous period. R-LSTM improved its accuracy by retraining with this new pattern between wind power and wind speed. Consequently, it could learn and adapt to the new conditions, leading to enhanced accuracy. The R-LSTM based forecasting results confirmed that the NMAE values were within 10%, and the model performance varied with the hyperparameter.

B. ENHANCED TRANSMISSION CONGESTION ESTIMATION IN JEJU ISLAND'S POWER SYSTEMS

The steady-state power flow was calculated using the wind power output and power demand scenarios generated using the KDE-MCMC method. Critical contingency analysis was also conducted. As a result of the power flow calculation and contingency analysis, six transmission lines, labeled Lines 1–6, were identified as congested components because their line flows exceeded the thermal limit based on the scenarios [12]. The probability of transmission congestion was calculated according to the line, season, and time.

The probability of transmission congestion was calculated using the number of line overloads. The results for the

FIGURE 15. Line flow of 24-hour each season based on R-LSTM wind power forecasting results.

probabilities of the congested components are presented in Table 7 and Figs. 13 and 14. Fig. 13 shows the minimum, average, and maximum line flows for each season and period. The x-axis values from 1 to 16 represent spring overnights to winter evenings, respectively. Fig. 14 shows the histogram and cumulative distribution function of the transmission line load ratio.

The probability estimation results for transmission congestion indicate that Line 1 exhibits the highest probability of transmission congestion during summer evening, while the probability is lowest during spring morning and afternoon. Throughout all seasons, the evening time period consistently shows the highest probability of transmission congestion. For Line 2, the probability of transmission congestion is highest during spring afternoon and lowest during summer evening. During all seasons, the evening time period consistently has the lowest probability of transmission congestion, with an average of 51.05% during other time periods. Line 3 exhibits the highest probability of transmission congestion during spring afternoon, with highest probabilities during the afternoon time period across all seasons. Line 4, 5, and 6 show the highest probability of transmission congestion during winter evening and the lowest during summer evening.

In summer, it can be observed that the probability of transmission congestion differs from other seasons because the wind power output of wind farm A is more likely to be lower compared to other seasons. These results were derived by considering seasonal hourly scenarios that can occur, and it is necessary to consider these findings when selecting a site for a wind farm and establishing a transmission reinforcement plan. In addition, the congestion can be mitigated by efficiently utilizing flexible resources based on the results of the probabilistic estimation of transmission congestion, potentially mitigating wind power curtailment.

Fig. 15 shows the 24-hour transmission congestion estimation using the R-LSTM forecasting results with a time interval of 15 min. The x-axis represents a 24-hour period with a 15-minute interval and the y-axis represents the line flow (%). The short-term wind power forecasting results for the first day of the last week of each season were applied to the power system, and steady-state power flow was calculated. The results indicate that the frequency of transmission congestion during the 24-hour is the highest in winter compared to other seasons. Accurate wind power output forecasting can optimize congestion management.

VI. CONCLUSION

To achieve carbon neutrality, the use of renewable energy sources, such as wind turbines, is increasing. However, because wind power generators use wind as an energy source to output electricity, they exhibit characteristics such as variability and uncertainty. Such characteristics make it difficult to maintain power system stability and reliability. Additionally, an increase in wind power resources has led to an increase in curtailment. Transmission congestion is a major cause of this curtailment. Therefore, accurate wind power output forecasting and transmission expansion are required to manage the curtailment caused by transmission congestion. When expanding a transmission network, it is necessary to estimate the transmission congestion by assuming wind power scenarios.

In this study, the R-LSTM was evaluated to improve the accuracy of short-term wind power forecasting. The R-LSTM, which incorporates recursive strategies, exhibited higher accuracy than the LSTM using wind farm A data from the Jeju power system. As a result, the performance of R-LSTM changes from an NMAE of 10.29% to 8.39% based on the number of hidden layers and nodes. In addition, the seasonal forecast results of the R-LSTM showed a higher accuracy than those of the LSTM. In particular, in the fall, the NMAE decreased from 22.86% for LSTM to 10.59% for R-LSTM. Moreover, the forecasting results were used for 24-hour transmission congestion estimation.

Furthermore, this study was conducted to estimate the probability of transmission congestion for mitigating wind power curtailment. Wind power and electricity demand scenarios were generated using KDE-MCMC method to calculate the probability of transmission congestion. This study used the data of 1-year from wind farm A and power demand data from Jeju Island and generated 16,000 scenarios to consider variability and uncertainty. The KDE-MCMC sampling showed high accuracy. Based on power flow calculation and N-1 contingency analysis results, the probability of line overload based on congested components were calculated. The probability estimation results of transmission congestion identified the season and time of day with the highest and lowest probability of transmission congestion occurrence for each congestion components. When expanding a transmission network to accommodate the available wind power output, it is necessary to consider the seasonal hourly congestion probability of the congested components.

In future, we will estimate the curtailment based on the results of the estimation of transmission congestion and perform a study on various wind farms and other areas.

REFERENCES

[1] *The 10th Basic Plan of Long-Term Electricity Supply and Demand*, Ministry Trade, Ind. Energy, Jan. 2023.

[2] *CFI 2030 Plan Amendment and Supplementation Service*, Jeju Special Self-Governing Province, Jun. 2019.

[3] L. Bird, J. Cochran, and X. Wang, "Wind and solar energy curtailment: Experience and practices in the United States," Nat. Renew. Energy Lab. (NREL), Golden, CO, USA, Tech. Rep. TP-6A20-60983, Mar. 2014.

[4] J. Jorgenson, T. Mai, and G. Brinkman, "Reducing wind curtailment through transmission expansion in a wind vision future," Nat. Renew. Energy Lab. (NREL), Golden, CO, USA, Tech. Rep. TP-6A20-67240, Jan. 2017.

[5] E. Mora, J. Cifuentes, and G. Marulanda, "Short-term forecasting of wind energy: A comparison of deep learning frameworks," *Energies*, vol. 14, no. 23, p. 7943, Nov. 2021.

[6] L. Han, H. Jing, R. Zhang, and Z. Gao, "Wind power forecast based on improved long short term memory network," *Energy*, vol. 189, Dec. 2019, Art. no. 116300.

[7] X. Liu, J. Zhou, and H. Qian, "Short-term wind power forecasting by stacked recurrent neural networks with parametric sine activation function," *Electr. Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 107011.

[8] J. Ren, Z. Yu, G. Gao, G. Yu, and J. Yu, "A CNN-LSTM-LightGBM based short-term wind power prediction method based on attention mechanism," *Energy Rep.*, vol. 8, pp. 437–443, Aug. 2022.

[9] W.-H. Lin, P. Wang, K.-M. Chao, H.-C. Lin, Z.-Y. Yang, and Y.-H. Lai, "Wind power forecasting with deep learning networks: time-series forecasting," *Appl. Sci.*, vol. 11, no. 21, p. 10335, Nov. 2021.

[10] D. Liu, H. Cheng, J. Lv, Y. Fu, and J. Zhang, "Probability constrained optimisation model for transmission expansion planning considering the curtailment of wind power," *J. Eng.*, vol. 2019, no. 18, pp. 5340–5344, Jun. 2019.

[11] T. Mai, A. Lopez, M. Mowers, and E. Lantz, "Interactions of wind energy project siting, wind resource potential, and the evolution of the U.S. power system," *Energy*, vol. 223, May 2021, Art. no. 119998.

[12] E. Memmel, S. Schlüters, R. Völker, F. Schuldt, K. Von Maydell, and C. Agert, "Forecast of renewable curtailment in distribution grids considering uncertainties," *IEEE Access*, vol. 9, pp. 60828–60840, 2021.

[13] P. Beiter, T. Mai, M. Mowers, and J. Bistline, "Expanded modelling scenarios to understand the role of offshore wind in decarbonizing the United States," *Nature Energy*, vol. 8, no. 11, pp. 1240–1249, Oct. 2023.

[14] M. J. Mayer, B. Biró, B. Szücs, and A. Aszódi, "Probabilistic modeling of future electricity systems with high renewable energy penetration using machine learning," *Appl. Energy*, vol. 336, Apr. 2023, Art. no. 120801.

[15] A. ShobanaDevi, G. Maragatham, M. R. Prabu, and K. Boopathi, "Short-term wind power forecasting using R-LSTM," *Int. J. Renew. Energy Res.*, vol. 11, no. 1, pp. 392–406, Mar. 2021.

[16] Electric Power Statistics Information System (EPSIS). *Installed Capacity by Generating Resources*. Accessed: Aug. 5, 2022. [Online]. Available: https://epsis.kpx.or.kr/epsisnew/selectEkmaGcpBftChart.do?menuId=040301

[17] Korea Power Exchange (KPX). *Hourly Electricity Demand in Jeju Power System*. Accessed: Aug. 5, 2022. [Online]. Available: https://www.data.go.kr/data/15065239/fileData.do#tab-layer-file

[18] Korea Power Exchange (KPX). *PV and Wind Curtailment in Jeju Power System*. Accessed: Aug. 2, 2022. [Online]. Available: https://www.data.go.kr/data/15099624/fileData.do

[19] T. Lee and Y. Lee, "Renewable energy expansion and direction of stable operation of the power system in Jeju Island," Energy Focus, Korea Energy Econ. Inst., Tech. Rep., Jan. 2021, pp. 48–63.

[20] Korea Power Exchange (KPX). *Hourly Wind Curtailment in Jeju Power System in 2021*. Accessed: Jul. 20, 2022. [Online]. Available: https://www.data.go.kr/data/15102485/fileData.do

[21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[22] H. Liu, X.-W. Mi, and Y.-F. Li, "Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network," *Energy Convers. Manage.*, vol. 156, pp. 498–514, Jan. 2018.

[23] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Orlando, FL, USA, Dec. 2018, pp. 1394–1401.

[24] J. Ho et al., "Regional energy deployment system (ReEDS) model documentation: Version 2020," Nat. Renew. Energy Lab. (NREL), Golden, CO, USA, Tech. Rep. TP-6A20-78195, Jun. 2021.

[25] A. Benseddik, A. Azzi, F. Chellali, R. Khanniche, and K. Allaf, "An analysis of meteorological parameters influencing solar drying systems in Algeria using the isopleth chart technique," *Renew. Energy*, vol. 122, pp. 173–183, Jul. 2018.

[26] F. Kamalov, "Kernel density estimation based sampling for imbalanced class distribution," *Inf. Sci.*, vol. 512, pp. 1192–1201, Feb. 2020.
[27] B. W. Silverman, "Using kernel density estimates to investigate multimodality," *J. Roy. Stat. Soc., B, Methodol.*, vol. 43, no. 1, pp. 97–99, Sep. 1981.
[28] K. Sahlin, "Estimating convergence of Markov chain Monte Carlo simulations," M.S. thesis, Dept. Math. Statist., Stockholm Univ., Stockholm, Sweden, 2011.
[29] C. P. Robert and G. Casella, *The Metropolis-Hastings Algorithms, Introducing Monte Carlo Methods With R*, vol. 18. New York, NY, USA: Springer, 2010, pp. 167–197.
[30] S. Chib and E. Greenberg, "Understanding the metropolis-hastings algorithm," *Amer. Statistician*, vol. 49, no. 4, p. 327, Nov. 1995.
[31] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," *Procedia Comput. Sci.*, vol. 167, pp. 2091–2100, Jan. 2020.

MINJU LEE (Student Member, IEEE) received the B.S. degree in climate and energy systems engineering from Ewha Womans University, Seoul, South Korea, in 2022, where she is currently pursuing the degree with the Department of Climate and Energy Systems Engineering. Her research interests include short-term wind power forecasting and the probabilistic estimation of transmission congestion for grid integration.

JIN HUR (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, South Korea, in 1997 and 1999, respectively, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, in 2012. He is currently an Associate Professor with the Department of Climate and Energy Systems Engineering, Ewha Womans University. His research interests include all areas related to integrating high-level renewable energy into electric power systems.