Progress Of Electrical And Electronic Engineering
2023 Volume 1, Issue 1
ISSN: 2630-4643

Comparison of Solar Power Prediction Model Based on Statistical and Artificial Intelligence Model and Analysis of Revenue for Forecasting Policy

Jeong-In Lee[1,2], Wan-Ki Park[1], Il-Woo Lee[1], Sang-Ha Kim[2]*

[1]Energy ICT Research Section, Electronics and Telecommunications Research Institute, Yuseong, Daejeon 34129, Korea

[2]Department of Computer Engineering, Chugnam National University, Yuseong, Daejeon 34134, Korea

***Corresponding author:** Sang-Ha Kim, shkim@cnu.ac.kr

Abstract

Korea is pursuing a plan to switch and expand energy sources with a focus on renewable energy to become carbon neutral by 2050. As the instability of energy supply increases due to the intermittent nature of renewable energy, accurate prediction of the amount of renewable energy generation is becoming more important. Therefore, the government has opened a small-scale power brokerage market and is implementing a system that pays settlements according to the accuracy of renewable energy prediction. In this paper, a prediction model was implemented using a statistical model and an artificial intelligence model for the prediction of solar power generation. In addition, the results of prediction accuracy were compared and analyzed, and the revenue from the settlement amount of the renewable energy generation forecasting system was estimated.

Keywords

Solar forecast
Deep learning
Electricity brokerage market
Accuracy incentives

1. Introduction

In response to the global climate crisis caused by worldwide climate change, major countries have declared carbon neutrality, and global companies are also pursuing Renewable Energy 100 (RE100) declarations and carbon neutrality management strategies. In line with this trend, Korea has also declared a national vision to achieve carbon neutrality by 2050. In addition, Korea has increased its national greenhouse gas (GHG) reduction, also known as Nationally Determined Contribution (NDC), target for 2030 from the previous target of 26.3% to 40%. The NDC requires a 40% reduction in GHG emissions by 2030 (436.6 million tons) compared to the peak of GHG emissions in 2018 (727.6 million tons) [1]. The sector with the highest proportion of GHG emissions is the energy sector, specifically power generation, which needs to be reduced by 44% from 2018 (269.6 million tons) to 2030 (149.9 million tons) [2]. **Figure 1** shows changes in the proportion of renewable energy sources in the power generation sector for the 2030 GHG reduction target. The proportion of renewable energy sources presented in NDC is 30.2%, which is more than 10% higher than the proportion proposed in 'Renewable Energy 3020'. Therefore, a supporting system is being developed for expanding renewable energy and enhancing adaptability [3].

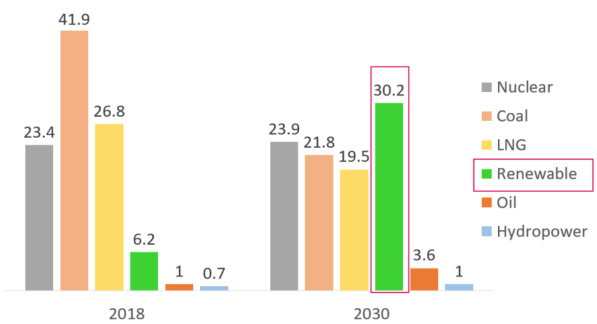


Figure 1. Percentage of energy mix configuration

Under the Energy 3020 policy, more than 95% of new installations are planned to be supplied by solar power (57%) and wind power (28%), leading to a rapid increase in supply. As renewable energy capacity increases, there are challenges in the operation of the power grid, including increased variability in electricity demand and the need for reserve capacity to address intermittent renewable energy generation [4,5]. The importance of accurate forecasting of renewable energy generation is growing.

Table 1 shows the cumulative installation status of domestic solar power facilities in Korea in 2020, with most of them being solar power generation facilities of 1 MW or less. Solar power generation facilities with 1 MW or less have power purchase agreement (PPA) contracts with Korea Electric Power Corporation (KEPCO) or entrust them to power brokers to generate revenues through power brokerage market transactions.

Table 1. Current status of solar power

Cumulative solar supply by capacity in 2020 (kW)		Percentage (%)
1 kW or less	31,333	
1–3 kW or less	1,221,898	
3–10 kW or less	280,528	
10–50 kW or less	814,768	82.6
50–100 kW or less	4,886,736	
100–500 kW or less	3,749,380	
500–1,000 kW or less	3,323,023	
1,000–5,000 kW or less	2,057,141	
5,000–10,000 kW or less	273,204	15.1
1,0000–2,0000k W or less	282,028	
Over 20,000 kW	402,664	2.3
Total	17,322,703	100

Source: Korea Energy Agency, Renewable Energy Supply Statistics (2020)

To address the variability in the output of renewable energy sources with intermittent characteristics, KEPCO introduced a renewable energy generation forecasting system in October 2021 for participants in the electricity brokerage market. If the prediction error rate is 8% or less, a settlement of 3 KRW (4 KRW/kWh for errors of 6% or less) was paid [6]. The capacity criteria for collective resources participating in the electricity brokerage market were raised to 20 MW or less from 1 MW in 2021 [7], which is expected to activate participation in the electricity brokerage market and the renewable energy generation forecasting system.

$$MAE = \frac{1}{N}\sum |\hat{y} - y| \quad (1)$$

$$NMAE = \frac{MAE}{P} \quad (2)$$

$$MSE = \frac{1}{N}\sum (\hat{y} - y)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum (\hat{y} - y)^2} \quad (4)$$

To evaluate the renewable energy generation forecasting accuracy, KPX used the normalized mean absolute error (NMAE) as a performance metric. Formula (1) represents the mean absolute error (MAE), where N is the number of data, $\hat{y}$ represents the predicted value, and y represents the actual value. Formula (2) represents NMAE, which is advantageous for comparing prediction error values of different power plants by normalizing them based on the power plant's capacity. Additionally, mean square error (MSE) and root mean square error (RMSE) are sometimes used for accuracy assessment, as shown in formula (3) for MSE and formula (4) for RMSE.

Figure 2 illustrates the structure of a small-scale electricity brokerage market, where intermediaries can sell electricity generated from renewable energy sources on behalf of power producers and generate revenue through the accuracy-based settlement system of the renewable energy generation forecasting system. Moreover, they can facilitate renewable energy certificate (REC) transactions through the REC market.

As of 2021, 59 companies and 602 MW of resources are registered in the electricity brokerage market, and the status of participation in the renewable energy generation forecasting system is presented in **Table 2**. The number of new resources continues to increase for participation in the renewable energy generation forecasting system.

In this paper, various forecasting models were implemented and compared to analyze the revenue of intermediaries based on the accuracy of generation forecasts. The structure of this paper includes an explanation of dataset composition in Section 2, a comparison and analysis of implemented forecasting models in Section 3, and finally, Section 4 concludes the paper and outlines future research directions.

2. Data analysis of solar power generation forecasting models

In this section, an analysis of the meteorological data and solar power generation data required for training solar power generation forecasting models is conducted. Data from November 2015 to December 2020 was collected and analyzed. All models implemented in this paper were developed using Python and Python libraries.

2.1. Solar power generation data

Figure 3 shows the empirical site information of the solar power generation facility used in this paper. Solar data was extracted from a 30-kW capacity solar facility installed in a building parking lot, with data including solar radiation, module temperature, and power generation recorded at hourly intervals.

Figure 4 visualizes the solar power generation and solar radiation data from the empirical solar power plant site to be used in this paper, both on an hourly and daily basis. Patterns in solar radiation and power generation were identified to repeat seasonally, forming patterns

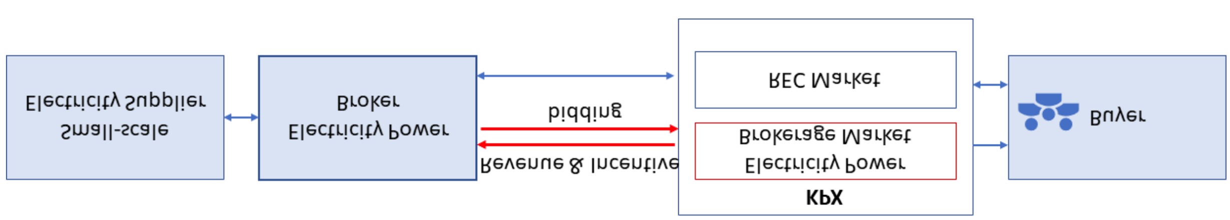


Figure 2. Configuration of small-scale electricity brokerage market

Table 2. Participation in renewable energy generation forecasting system

Sort	Resource	2021	
		Number of businesses	**Capacity (MW)**
Power brokerage	Collective power resources	8	190
Power generation company	Individual power resources	1	96

Source: Korea Power Exchange (KPX), Electricity Market Statistics (2021)

Figure 3. Information on solar power test site

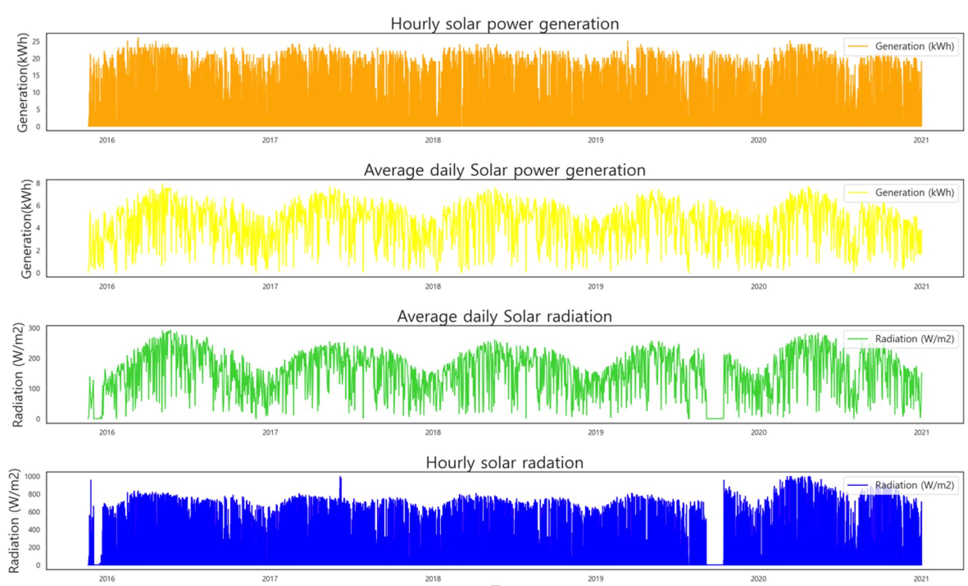
(a) Solar power test site

현 장 명	한국전자통신연구원
설비명(용량)	전체보기 (30.2 kW)
조회일자	2017-06-01

년 - 월 - 일 시	일사량 (W/m²)	온도 (℃)	DC 전력 (kWh)	AC 전력 (kWh)	발전효율 (%)	발전량 (kWh)	누적발전량 (kWh)
2017-06-01 06	3.3	18.4	0.1	0.1	0.0	0.0	55373.0
2017-06-01 07	65.9	20.2	2.2	2.0	6.6	2.0	55375.0
2017-06-01 08	150.5	25.2	4.8	4.7	13.2	4.0	55379.0
2017-06-01 09	173.7	28.3	5.3	5.2	19.8	6.0	55385.0

(b) Solar power generation status report

Figure 4. Information on solar power generation and radiation



that vary with the seasons.

2.2. Local forecast data

In order to create power generation forecasting models, meteorological information from short-term (daily) forecasts such as temperature and humidity is required in addition to solar power generation history data. Short-term forecast data for the location where the solar power generation facility is installed was obtained from the Korea Meteorological Administration's Meteorological Data Open Portal, as shown in **Figure 5**, and was analyzed [8]. **Table 3** presents statistical information for the short-term forecast data in the corresponding area.

Table 3. Statistical information of daily weather forecasts

Sort	Mean	Std	Min	25%	50%	75%	Max
Temperature	13.5	10.4	-17.0	5.0	14.0	22.0	38.0
Humidity	68.6	18.9	5.0	55.0	70.0	85.0	100.0
Wind speed	1.9	1.4	0.0	1.0	1.6	2.4	26.9
Wind direction	202.2	103.4	0.0	104.0	223.0	297.0	360.0
Snow cover	0.04	0.37	0.0	0.0	0.0	0.0	5.0
Precipitation	1.5	6.1	0.0	0.0	0.0	0.0	100.0
Sky condition	2.55	1.13	1.0	1.0	3.0	3.8	4.0
Precipitation type	0.17	0.52	0.0	0.0	0.0	0.0	4.0
Precipitation probability	20.6	20.9	0.0	0.0	20.0	30.0	100.0

Figure 5. Korea Meteorological Administration Weather Data Service

Figure 6 visualizes the correlation between solar power generation and short-term forecasts. Temperature showed the highest positive correlation with power generation, while wind speed and wind direction also exhibited positive correlations. Other short-term forecast variables demonstrated negative correlations with power generation, with humidity, sky conditions, precipitation probability, and precipitation type showing significant negative correlation values.

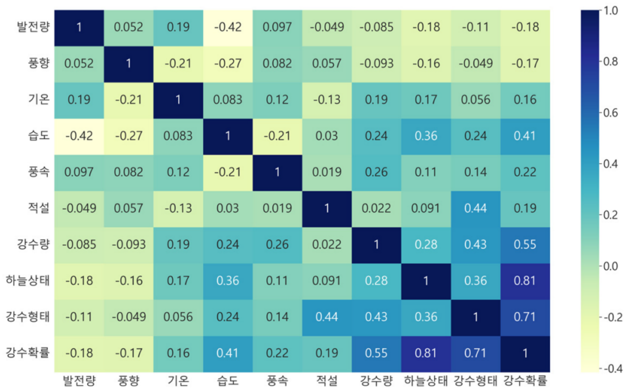


Figure 6. Visualization of the correlation between weather and solar power generation

3. Solar power generation forecasting models and estimation of renewable energy forecasting system settlement fees

In this section, forecasting models for solar power generation were implemented using both solar power generation history data and short-term forecast data. **Figure 7** illustrates the classification of time series forecasting models [9], and based on this, solar power generation forecasting models were explained and implemented using statistical and machine learning approaches.

3.1. Statistical model-based solar power generation forecasting

3.1.1. Moving average method

The moving average method calculates the moving average of time series data over a certain period and uses it to estimate the next period's value based on trends. It utilizes equal weights for the past n observed values to predict the value for the next period. Equation (5) represents the formula for the moving average method, where n is the number of data observations, Z_t is actual value at time t, and F_{n+1} represents the predicted value for the next time point (t+1).

$$F_{n+1} = \frac{1}{n}(Z_t + Z_{t-1} + \cdots Z_{t-n+1}) \qquad (5)$$

Figure 8 displays graphs of 6-hour and 24-hour moving averages. Predictions made using the moving average are based solely on the most recent estimated value and do not estimate the trend. Therefore, it is suitable for short-term predictions up to 6 hours, as seen in **Figure 8(a)**, while 24-hour power generation predictions are more challenging, as shown in **Figure 8(b)**.

3.1.2. Autoregressive integrated moving average (ARIMA) model

The autoregressive integrated moving average (ARIMA) model is a time series forecasting model expressed as ARIMA(p, q, d), where 'p' represents the order of autoregression (AR), 'q' represents the order of the moving average (MA) model, and 'd' represents the

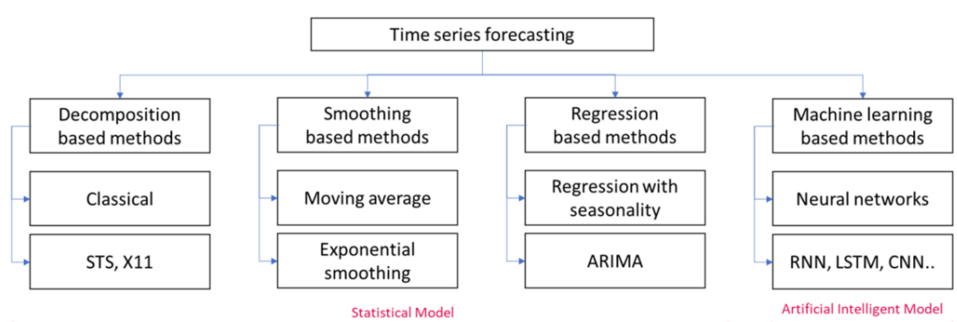
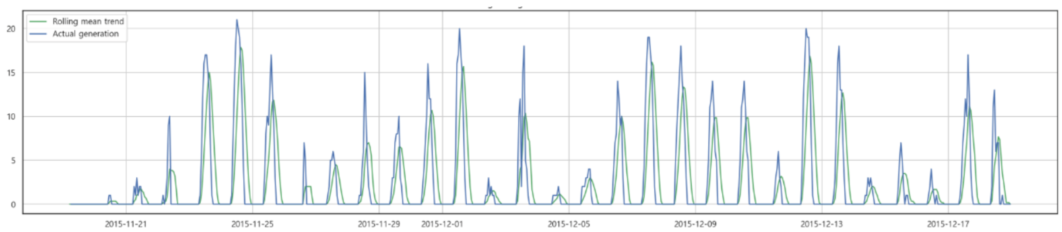


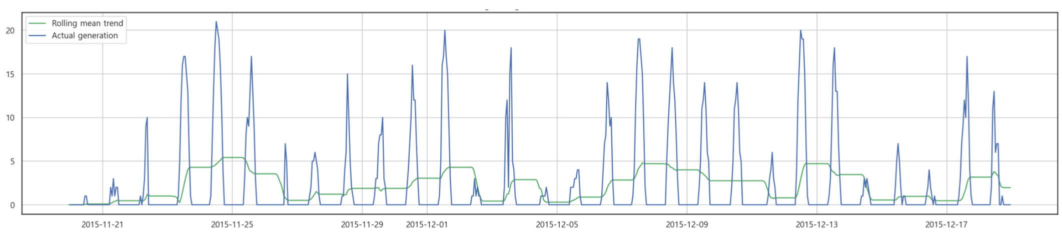
Figure 7. Taxonomy of time series forecasting techniques

differencing order [10].

In **Figure 9**, the ARIMA model was used to compare actual and predicted values of solar power generation at hourly, daily, and weekly intervals. To participate in the electricity brokerage market and the renewable energy power generation forecasting system, a mechanism for predicting 24-hour power generation in advance and receiving settlement fees based on the comparison with actual generation is used. By applying ARIMA, the NMAE for hourly power generation prediction

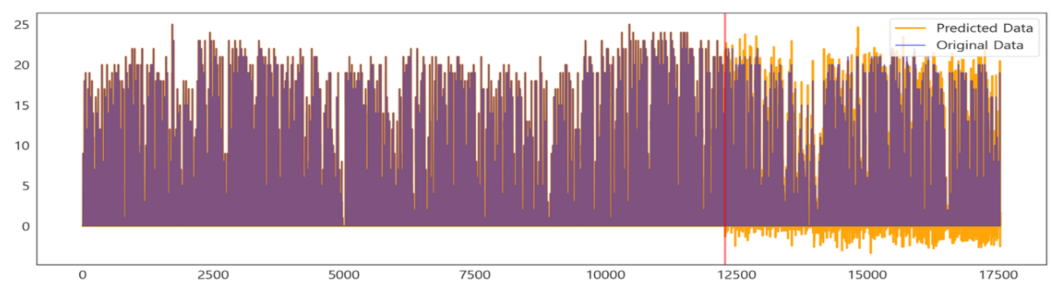


(a) Moving average with window size = 6

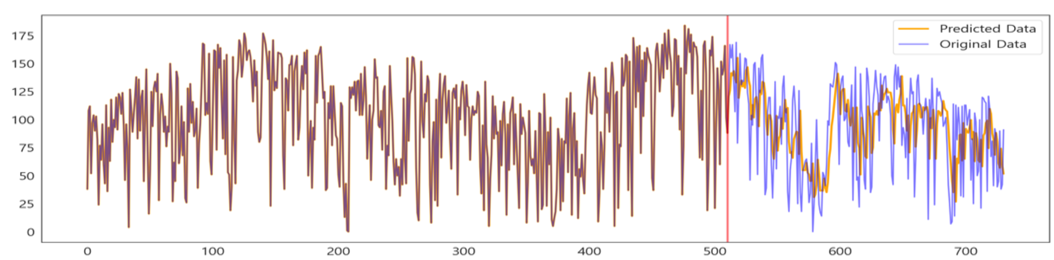


(b) Moving average with window size = 24

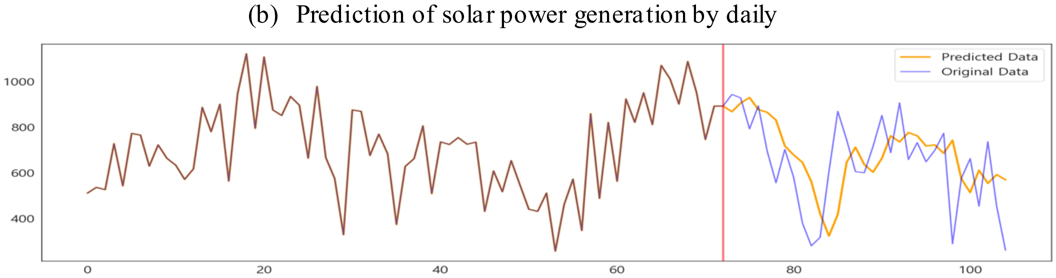
Figure 8. Prediction of solar power generation by moving average method



(a) Prediction of solar power generation by hourly



(b) Prediction of solar power generation by daily



(c) Prediction of solar power generation by weekly

Figure 9. Prediction of solar power generation by ARIMA

errors was calculated as 7.934, indicating that there are segments where settlement fees can be received.

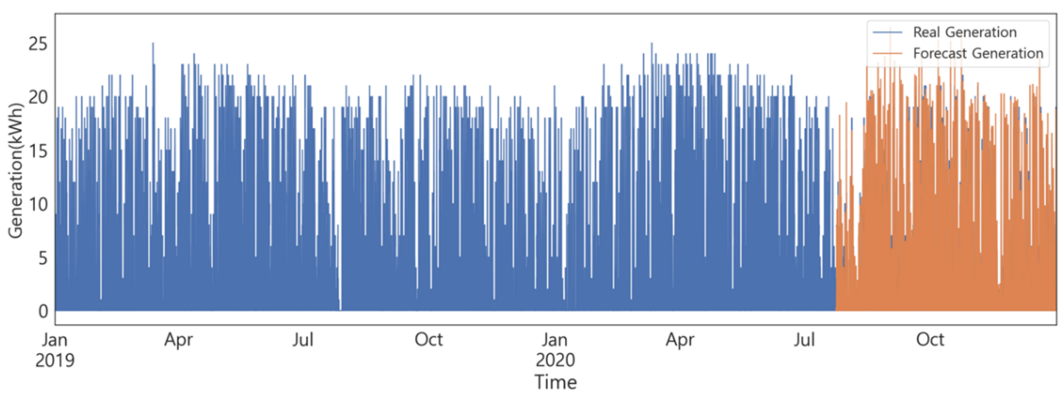
3.1.3. Seasonal autoregressive integrated moving average (SARIMA) model

While ARIMA considers trend differencing, the seasonal ARIMA (SARIMA) model incorporates seasonality by adding seasonal autoregressive (SAR) terms, seasonal moving average (SMA) terms, and seasonal differencing to the ARIMA model [11]. It is expressed as SARIMA(p, q, d) (P, Q, D, S), where ‘P’ represents the SAR order, ‘D’ represents the seasonal differing order, ‘Q’ denotes the SMA order, and ‘S’ represents the seasonal period. **Figure 10** compares predictions for hourly and daily solar power generation. The NMAE for SARIMA was approximately 4.16, indicating lower prediction errors compared to the ARIMA model.

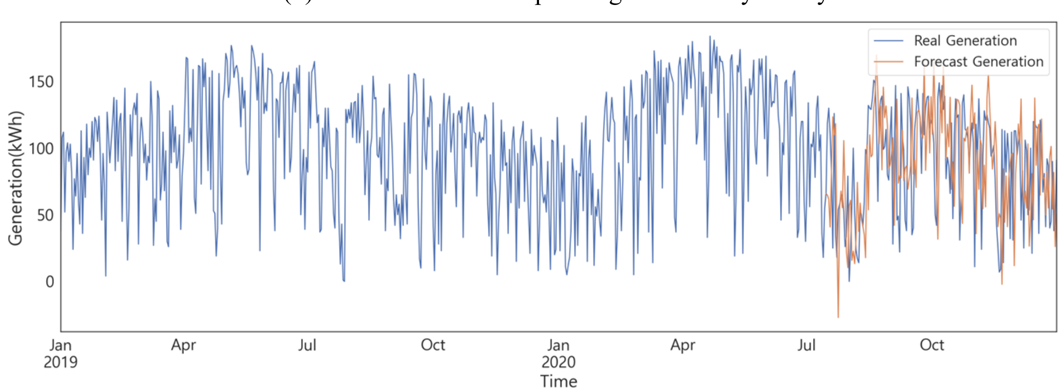
3.2. Artificial intelligence model-based solar power generation forecasting models

Settlement fee earnings in the forecasting system vary based on the accuracy of the power generation prediction technology used by intermediary businesses. Therefore, intermediaries are researching to improve the accuracy of their power generation prediction technology by applying various algorithms. Recent research has seen a surge in the use of artificial intelligence (AI) models for power generation forecasting, alongside traditional statistical models. In this section, solar power generation forecasting models were implemented using artificial intelligence models, and a comparative analysis of the results was provided.

Figure 11 presents the results using specialized recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU)



(a) Prediction of solar power generation by hourly

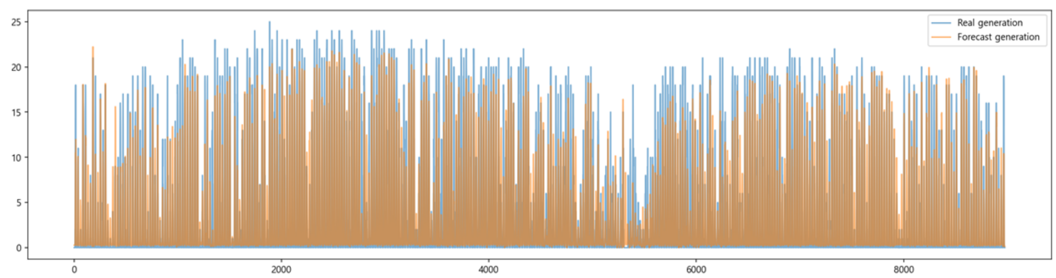


(b) Prediction of solar power generation by daily

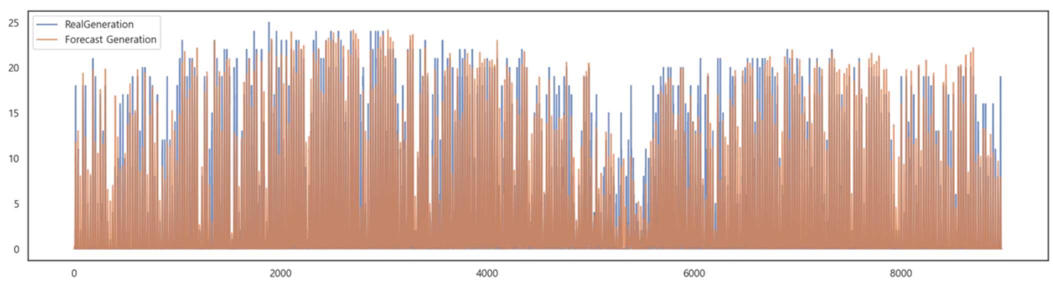
Figure 10. Prediction of solar power generation by SARIMA

models, and a combination of a convolutional neural network (CNN) and LSTM model known as convolutional long short-term memory neural networks (CNN-LSTM). **Figure 12** compares the monthly prediction accuracy of the four AI models, showing that the RNN model exhibits a larger deviation compared to the other models.

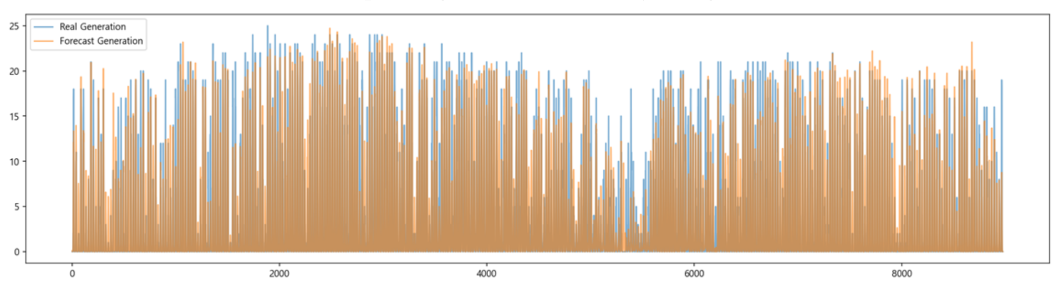
Table 4 compares the performance of the mentioned models in predicting solar power generation using various techniques. All five models have prediction error rates of less than 5% NMAE annually, confirming that all comparison models are eligible to receive settlement fees. Furthermore, the SARIMA model, which accounts for seasonal characteristics among the statistical models, exhibits higher prediction accuracy than the RNN model. This indicates that accurate solar power generation prediction can be achieved using statistical models alone, considering the seasonality and periodic characteristics



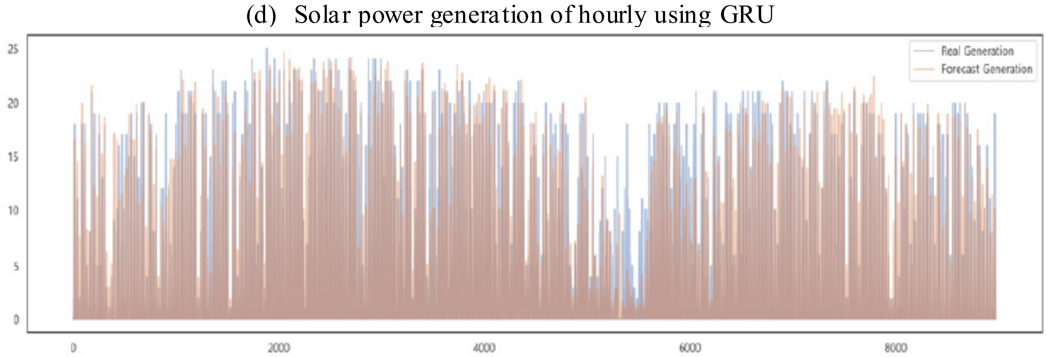
(a) Solar power generation of hourly using RNN



(c) Solar power generation of hourly using LSTM

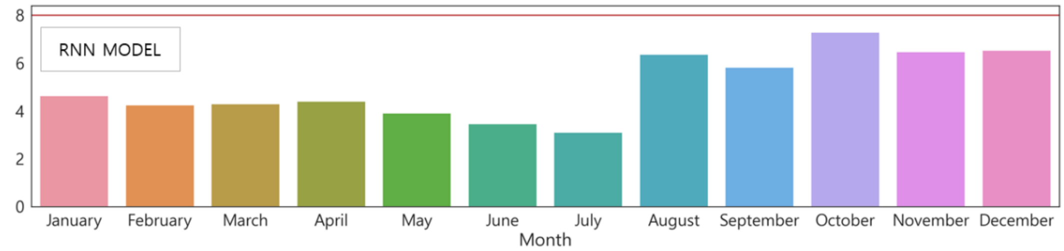


(d) Solar power generation of hourly using GRU

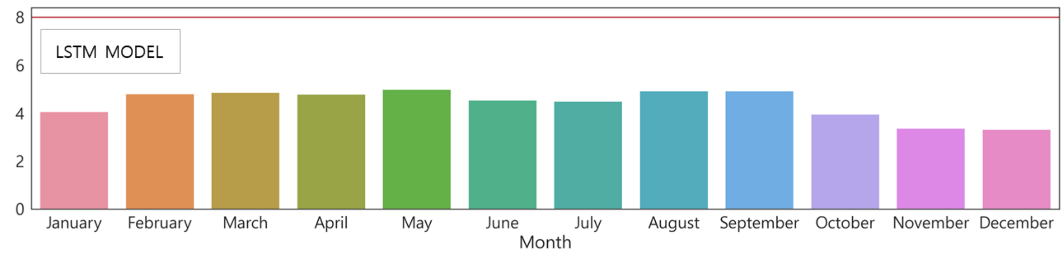


(b) Solar power generation of hourly using CNN-LSTM

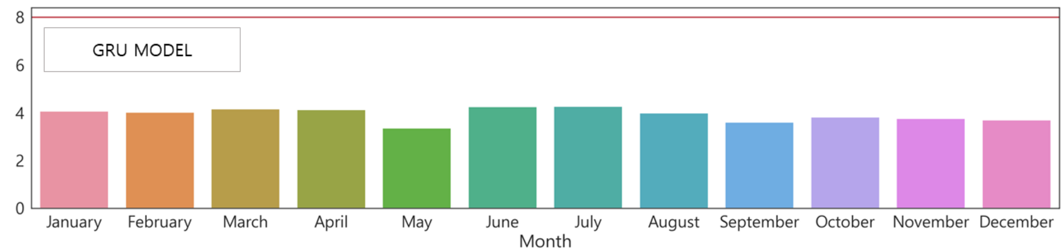
Figure 11. Prediction of solar power generation by artificial intelligence neural network



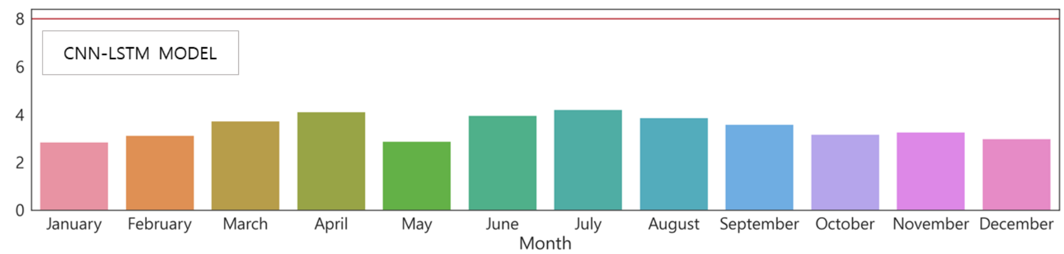
(a) Monthly NMAE using RNN



(b) Monthly NMAE using LSTM



(c) Monthly NMAE using GRU



(d) Monthly NMAE using CMM-LSTM

Figure 12. Solar power generation forecast NMAE by artificial intelligence neural network

Table 4. Metrics related to solar power generation dataset

Models	MAE	MSE	RMSE	NMAE
RNN	1.40225	6.04756	2.45918	4.67418
SARIMA	1.23770	4.77762	2.18577	4.15902
LSTM	1.24720	5.49212	2.34353	4.15732
GRU	1.16986	5.02075	2.24070	3.89952
CNN-LSTM	1.10522	4.82530	2.19666	3.65958

of solar generation and meteorological data. Moreover, using statistical models implies potential savings in computing power, GPU, and calculation time compared to AI models. Among the AI models, the CNN-LSTM model demonstrates the highest prediction accuracy. CNN-LSTM treats time series data as one-dimensional images in the CNN layer, extracting significant features from the time series data. With LSTM applied, it maintains long-term dependencies in the data, resulting in improved prediction accuracy compared to conventional LSTM and GRU models.

3.3. Estimation of settlement fees in the renewable energy generation forecasting system based on prediction accuracy

Table 5 provides an estimation of annual settlement fee earnings based on the prediction models described earlier and solar power generation data from the year 2020. It confirms that the earnings for intermediaries can vary depending on the level of prediction accuracy. It is noted that the scale of earnings can differ by approximately 24% based on the level of prediction accuracy.

Figure 13 visualizes the average NMAE and average settlement fee earnings by time of day. Significant NMAE errors occur during the daytime. However, since power generation is high during these hours due to abundant sunlight, it results in the highest earnings from settlement fees. **Figure 14** illustrates the monthly average of NMAE and settlement dee earnings. In January 2020, there was heavy snowfall, leading to significant prediction errors. Additionally, an unusual monsoon season extended until August 2020, causing substantial prediction errors [12]. Further analysis is required for September, and it is observed that settlement fee earnings are higher in the spring and autumn months when there is more sunlight.

Table 5. Estimation of the annual revenues from the incentives

Models	RNN	SARIMA	LSTM	GRU	CNN-LSTM
Revenues (KRW)	65,764	75,352	81,762	84,299	86,035

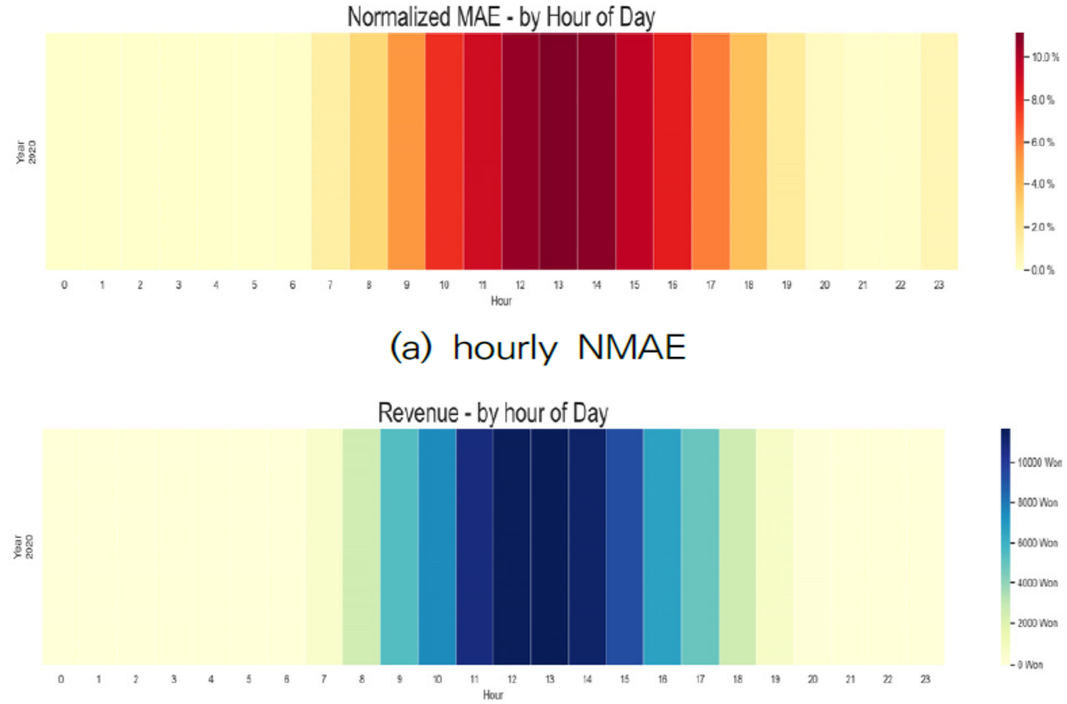
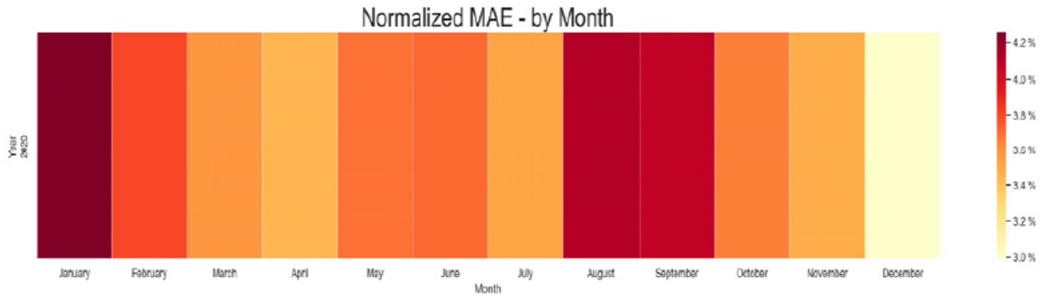
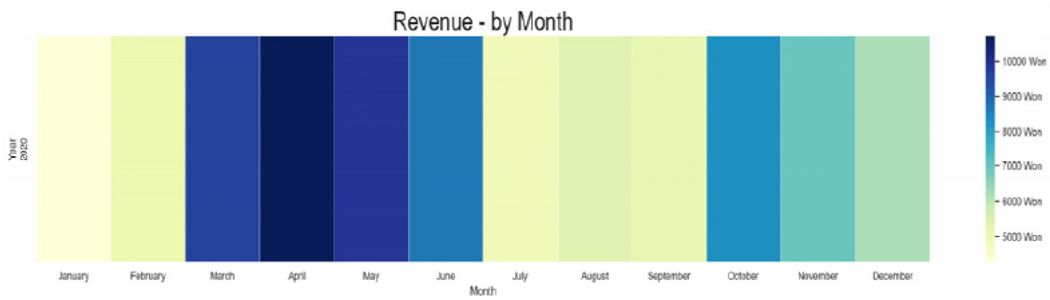


Figure 13. Hourly average NMAE and revenue



(a) Monthly NMAE



(b) Monthly revenue

Figure 14. Monthly average NMAE and revenue

4. Conclusion

In this paper, a 24-hour solar power generation forecasting model for participation in the electricity brokerage market using both statistical and AI models was implemented. The results were compared and analyzed. Among the statistical models, the SRIMA model, which incorporates seasonal characteristics, has a high performance. The AI models were confirmed to generally have higher prediction accuracy compared to statistical models. Among the AI models, CNN-LSTM exhibited the highest performance. Furthermore, the annual settlement fee earnings for intermediaries were estimated based on different prediction models in the context of the renewable energy generation forecasting system. It was observed that the accuracy of prediction technology can result in more than a 24% difference in earnings for intermediaries. Since the performance of AI models can vary based on user hyperparameter tuning, the level of prediction technology for intermediaries can become a crucial factor affecting prediction accuracy. In future research, additional analyses are planned to be conducted for times and seasons with significant prediction errors, as well as proposing methods to enhance prediction accuracy.

Disclosure statement

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the KETEP and the Ministry of Trade, Industry and Energy of the Republic of Korea.

References

[1] The Presidential Commission on Carbon Neutrality and Green Growth. 2050 Carbon Neutrality of the Republic of Korea. https://www.2050cnc.go.kr

[2] Joint Office of Korea Government, 2021, 2030 National Greenhouse Gas Reduction Target (NDC) Upgrade Plan, p. 5–6.

[3] Korean Ministry of Commerce, Industry and Energy, 2017, Renewable Energy 3020 Implementation Plan, p. 2–10.

[4] Kim Y-H, Myung H-S, Kang N-H, et al., 2018, Operation Plan of ESS for Increase of Acceptable Product of Renewable Energy to Power System. The Transactions of the Korean Institute of Electrical Engineers (KIEE), 67(11): 1401–1407. https://doi.org/10.5370/KIEE.2018.67.11. 1401

[5] KEMRI. Renewable Intermittent Response Trend in Major Countries. http://www.keaj.kr

[6] Korean Ministry of Commerce, Industry and Energy, 2020, Introduction of Renewable Energy Generation Prediction System.

[7] KPX. Brokerage Market Operation Rules. https://der.kmos.kr/info/fr_notice_view.do

[8] Korean Meteorological Agency. Short-Term Forecast. https://data.kma.go.kr

[9] Gonzalez-Longatt F, Acosta MN, Chamorro HR, et al. 2020 International Conference on Smart Systems and Technologies, October 14–16, 2020: Short-Term Kinetic Energy Forecast Using a Structural Time Series Model: Study Case of Nordic Power System. 2020, Osijek, 173–178. https://doi.org/10.1109/SST49455.2020.9264087

[10] Jeong H-Y, Hong S-H, Jeon J-S, et al., 2022, A Research of Prediction of Photovoltaic Power Using SARIMA Model. Journal of Korea Multimedia Society, 25(1): 82–91. https://doi.org/10.9717/kmms.2022.25.1.082

[11] Min K-C, Ha H-K, 2014, Forecasting the Korea's Port Container Volumes with SARIMA Model. Journal of Korean Society of Transportation, 32(6): 600–614. https://doi.org/ 10.7470/jkst.2014.32.6.600

[12] Korea Meteorological Administration, 2020, [2020 Abnormal Climate Report].

Publisher's note

Art & Technology Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.