Total PV Power Prediction over a Wide Area Using Forecasted Values of Solar Radiation at Multiple Observation Points

Ritsuya Watanabe
Department of EE and Bioscience
Waseda University
Tokyo, Japan
ritsuya.watanabe@akane.waseda.jp

Shinji Wakao
Department of EE and Bioscience
Waseda University
Tokyo, Japan
wakao@waseda.jp

Kei Morita
Central Technical Research Laboratory
JXTG Nippon Oil and Energy Corporation
Tokyo, Japan
morita.kei@jxtg.com

***Abstract*— Electric retailers need to predict spot market prices in the Japan Electric Power Exchange for efficient daily power procurement. And it is necessary to predict demand and PV output when predicting the price. Therefore, this paper deals with the simple and accurate estimation methods of total PV power over a wide area. We propose two kinds of simple methods to calculate the total PV output in the Kanto and Tohoku areas in Japan based on the predicted values of solar radiation at multiple observation points that are easy to obtain from the Japan Meteorological Agency.**

Keywords— MSM-GPV data, Solar radiation prediction, PV output prediction, Physical model, Regression model, JIT Modeling

I. INTRODUCTION

With the recent introduction of solar power generation, the apparent demand has decreased, and the daytime electricity market price has fallen. The influence of solar power generation on the market price is increasing year by year. In Japan, many electricity supplier rely their electricity procurement on JEPX. To easily and precisely predict the JEPX spot market price, we have developed a simple merit order model as in Fig.1. In the model, the JEPX price is regarded as the unit price of the generator at the intersection of the demand line and the curve created by stacking the amount of power generated from the generator in the order of lower unit price. The demand corresponds to the apparent demand obtained by subtracting the PV output from the actual demand. So, it is important to accurately predict the amount of solar power generation over a wide area controlled by an electric power company when predicting the market price. Otherwise, source at the intersection changes in the proposed merit order model, and the accuracy of the price forecast deteriorates.

With the above background, this paper deals with the estimation methods of total PV power over a wide area using forecasted values of solar radiation at multiple observation points. First, we predict solar radiation at multiple observation points with using just-in-time modeling based on the weather forecast values that are easily available from the Japan Meteorological Agency. And, we propose two kinds of simple methods to calculate the total PV power in the Kanto and Tohoku areas in Japan based on the predicted values of solar radiation at multiple observation points. Finally, some numerical examples, which demonstrate the validity of the proposed methods, will be presented.

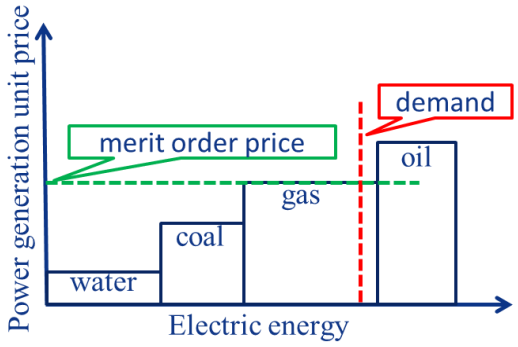

Fig. 1. Conceptual diagram of merit order model.

II. SOLAR RADIATION PREDICTION

In this chapter, the method of solar radiation prediction at a single location, which are used for wide area PV output prediction, is described. In this research, MSM-GPV data of meteorological forecasts distributed by the Japan Meteorological Agency is used [1-3]. The survey period is 1 year from April 1, 2018 to March 31, 2019, and the predicted value at a single point is calculated hourly for 5:00 to 19:00 at daytime. The verification point is the point where the measured value is observed (the point indicated by the red dot in Fig 2). In addition, the evaluation indicators of prediction accuracy are % MAE and RMSE expressed by the following formula.

$$\%MAE = \frac{\sum_{n=1}^{N} \left| y_{(n)} - \hat{y}_{(n)} \right|}{\sum_{n=1}^{N} y_{(n)}} \quad (1)$$

Identify applicable funding agency here. If none, delete this text box.

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}\left(y_{(n)} - \hat{y}_{(n)}\right)^2}{N}} \quad (2)$$

Where N is the number of data, $y_{(n)}$ is actual measured values, and $\hat{y}_{(n)}$ is predicted values.

A. MSM-GPV data

MSM-GPV data is the weather forecast value distributed by the Meteorological Agency. Since the MSM-GPV data includes the predicted value of solar radiation, this value can also be directly used for wide range PV power prediction. The MSM-GPV data consists of the predicted values of 12 components including this solar radiation. In this paper, Just-In-Time Modeling is constructed using these meteorological components, and the amount of solar radiation is predicted independently. In addition, the wide area PV power predicted value is calculated using the predicted solar radiation, and a comparative study is performed with the case where the solar radiation predicted value of MSM-GPV data is used.

MSM-GPV data forecasts up to 39 hours ahead are calculated every 3 hours from 00 UTC and will be delivered approximately 1.5 hours later. Predicted values are calculated every 5 km square. In this study, MSM-GPV data calculated at 21 UTC (09 JST) is used because PV power prediction must be done in the morning to predict JEPX spot market prices in the morning.

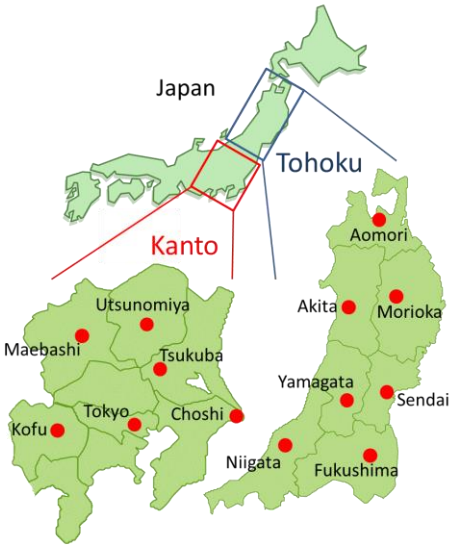

Fig. 2. Measurement points of solar radiation.

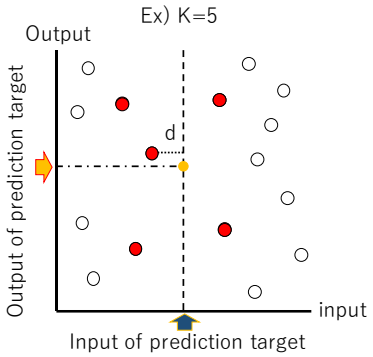

Fig. 3. Conceptual diagram of JIT Modeling.

B. Just-In-Time Modeling

Just-In-Time (JIT) Modeling is a type of black box model that searches for similar data from past input / output databases for the input to be predicted and calculates the predicted output based on the output of the similar data. It is considered that this model is suitable for the method of predicting solar radiation because of its strong adaptability to nonlinearity. The calculation method of JIT Modeling is summarized below, and a conceptual diagram is shown in Fig. 3.

1) Neighborhood selection: The input to be predicted is treated as an input vector (request point), and neighborhood data having an input vector similar to the request point is selected from the history database. Based on the K-NN method, the number of neighboring data to be selected is K in order of decreasing distance from the required point. The weighted Euclidean distance represented by (3) below is used as a measure of similarity. A partial regression coefficient was used for the weight w_i of the weighted Euclidean distance.

$$d_i = \sqrt{\sum_{n=1}^{P} w_n\left(x_{qn} - x_{in}\right)^2} \quad (3)$$

Where d_i is Weighted Euclidean distance between request point and i-th historical data, P is the number of input components, w_i is weigh, x_{qn} is n-th input component at the request point, x_{in} is the n-th input component in the i-th history data.

2) Output estimation: As a method of estimating the output using the selected neighborhood close to the required point, we adopted the weighted average method shown in (4) below. The weight of each neighborhood is determined according to the weighted Euclidean distance in (3). In this paper, the Tricube function in (5) is used as the weighting function T.

$$\hat{y} = \frac{\sum_{i=1}^{K} T_i y_i}{\sum_{i=1}^{K} T_i} \quad (4)$$

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.

$$T_i = \left(1 - \left(\frac{d_i}{d_K}\right)^3\right)^3 \quad (5)$$

Where $\hat{y}$ is estimated value, T_i is weights for the i-th neighbor, y_i is output of i-th neighbor.

C. Prediction accuracy of solar radiation

We will compare the accuracy of the solar radiation at a single point predicted by JIT Modeling with counterpart of MSM-GPV data. For the input of JIT Modeling, referring to [4], the relative humidity, lower cloud amount, middle cloud amount, upper cloud amount from MSM-GPV data and calculated solar altitude and daytime were used. A database was created for each prediction target time, and the data at the same time from 2017/4/1 to 2018/3/31 was used as historical data. The number of neighbors, which is one of the parameters of JIT Modeling, was simulated each from 10 to 150 at intervals of 10. As a result, the number of neighbors was 20 when %MAE is the smallest. And the number of neighbors was 30 when RMSE was the smallest. This time, from the viewpoint of the reduction of fatal error, we decided to use 30 in number of neighbors.

Fig. 6 and 7 show the accuracy of JIT Modeling insolation prediction for each number of neighbors. From Fig 6, it can be seen that the MSM-GPV data is slightly smaller in %MAE and has slightly better accuracy However, in Fig 7, the superiority and inferiority are reversed at some points compared to Fig 6. This suggests that JIT Modeling has a slightly larger average error compared to MSM-GPV, but is advantageous in avoiding fatal errors.

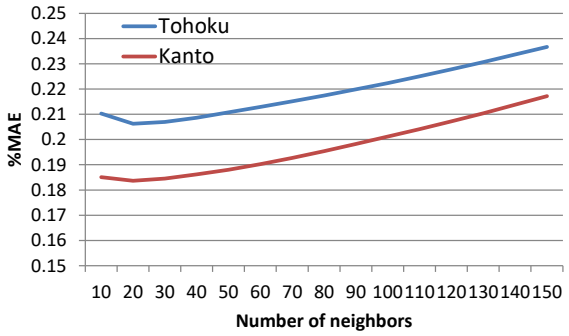

Fig. 4. Relationship between the number of neighbors and %MAE.

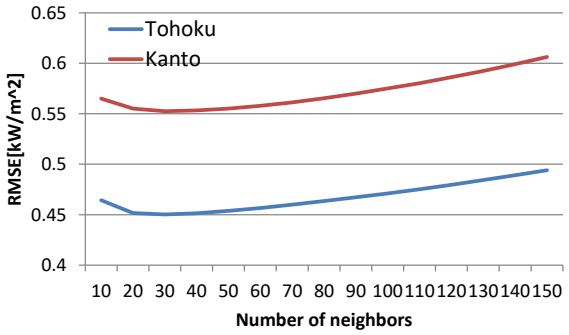

Fig. 5. Relationship between the number of neighbors and RMSE

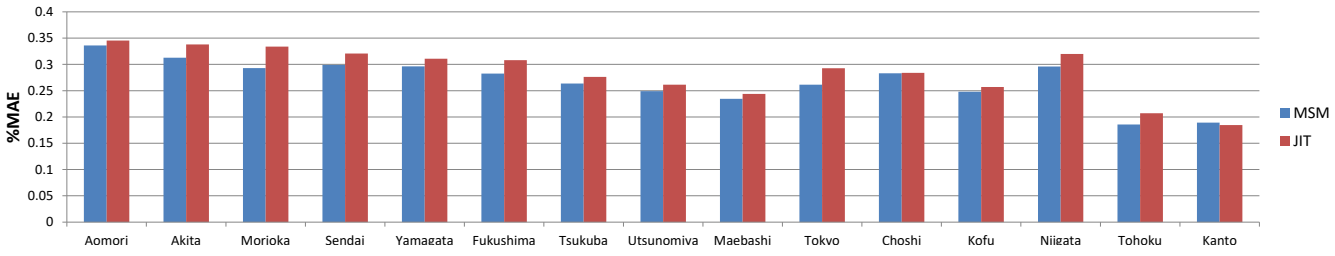

Fig. 6. %MAE of solar radiation forecast at each point.

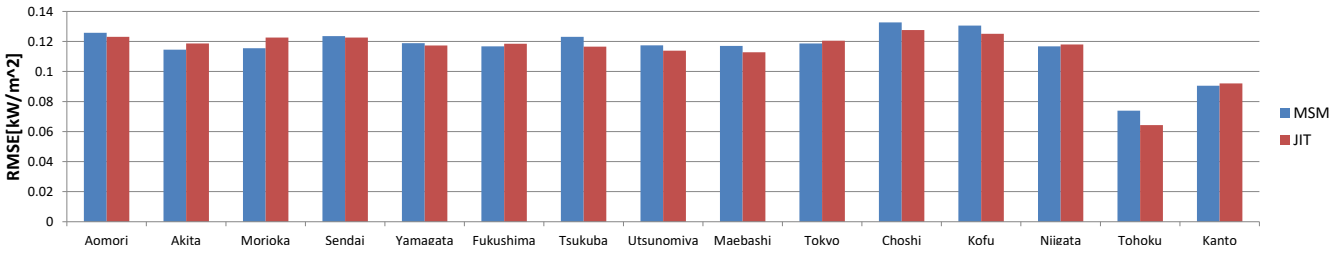

Fig. 7. RMSE of solar radiation forecast at each point.

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.

Fig. 8 shows the error distribution in the Kanto area, and Fig. 9 shows an enlarged view of the scale of the vertical axis of Fig. 8. Similarly, Fig. 10 shows the error distribution in the Tohoku area, and Fig. 11 shows an enlarged view of Fig. 10. In the histograms of Figs. 8 and 10, most of the MSM-GPV data and JIT Modeling prediction results are distributed near error 0 kw/m^2, and the numbers are similar. However, looking at Figures 9 and 11, it can be confirmed that the large error near -2 kw/m^2 tends to exist more double times in the predicted value of MSM-GPV. Therefore, JIT Modeling has the merit that is avoiding fatal errors.

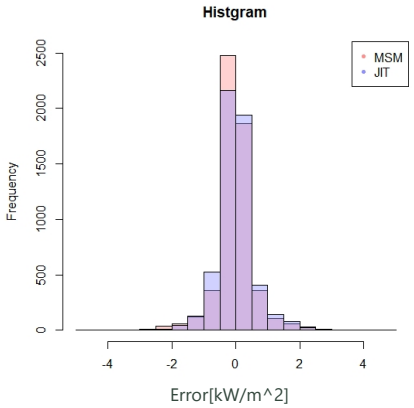

Fig. 8. Histgram of solar radiation forecast in Kanto.

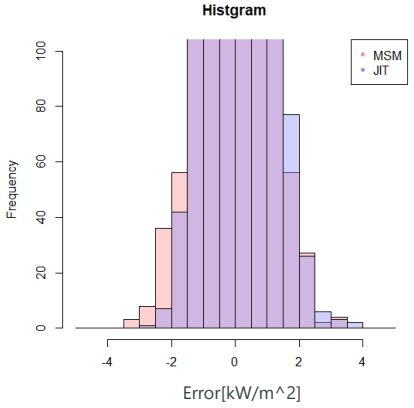

Fig. 9. Enlarged histgram of solar radiation forecast in Kanto.

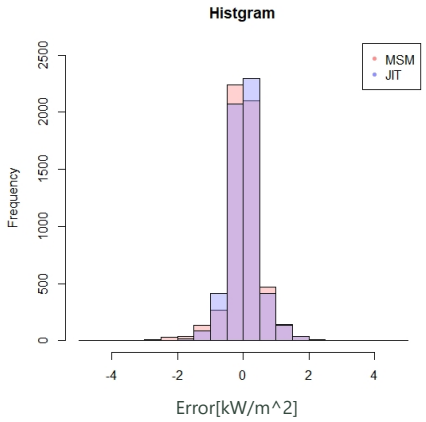

Fig. 10. Histgram of solar radiation forecast in Tohoku.

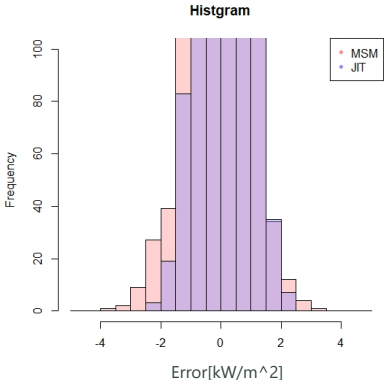

Fig. 11. Enlarged histgram of solar radiation forecast in Tohoku.

III. Prediction Total PV Power over a Wide Area

The hourly total PV power over a wide area in the next day is predicted with using the predicted solar radiation and the information published on the Internet, specifically measured total PV power and installed PV capacity in the area[5-7]. We propose the following two methods to predict PV power.

1) Physical method: Connverting the solar radiation forecast into PV power with using following (6) and (7) .

$$P_{(t)} = \alpha_{(m-1)} \sum_{n=1}^{N} C_{(n,m)} \times I_{(n,t)} \quad (6)$$

Where $P_{(t)}$ is the predicted PV powerat time t, $\alpha_{(m-1)}$ is conversion factor for the revious month on the forecast date, $C_{(n,m)}$ is PV installation capacity for a month m in the n-th prefecture, $I_{(n,t)}$ is the solar radiation forecast at time t at the representative point of the n-th prefecture.

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.

$$\alpha_{(m)} = \frac{G_{(m)}}{\sum_{n=1}^{N} C_{(n,m)} \times I_{(n,m)}} \quad (7)$$

Where $G_{(m)}$ is the accumulation of total PV power actual measurement for a month m in the area, $I_{(n,m)}$ is the accumulation of solar radiation forecast for a month m at the representative point of the n-th prefecture.

If the tentative PV power for a month is defined as the sum of the product of PV installation capacity of each prefecture and solar radiation at the representative point, the hourly total PV power the next day is predicted by multiplying the previous month's conversion factor by the sum of the tentative PV power of all prefectures in the area, as in (6). The conversion factor α is calculated from the ratio between the tentative PV power in the area and the actual measurement in the area as shown in (7).

2) Regression method: A method of collectively expressing PV introduction capacity C for each prefecture used in the physical method and the conversion coefficient α as a regression coefficient *a* by multiple regression analysis as in (8). In order to predicting for 2018, the data for 2017 is used as learning data. Since PV installation capacity is increasing in Japan, it is not possible to use the actual power generation results of the previous year while ignoring the PV installation capacity information. Therefore, multiply the PV introduction capacity increase rate from the same month of the previous year at each point.

$$P_{(t)} = \sum_{n=1}^{N} a_{(n)} \times I_{(n,t)} \times C'_{(n,m)} \quad (8)$$

Where $a_{(n)}$ is the multiple regression coefficient for the n-th prefecture, $C'_{(n,m)}$ is PV installed capacity ratio with the same month of the previous year.

The results of applying the proposed method to the Kanto area and Tohoku area are shown as the %MAE graph in Fig. 12 and the RMSE graph in Fig. 13, respectively. The physical model is more accurate in both areas, and its accuracy (%MAE) is about 22.5%. The reason why the accuracy of the regression model was relatively poor is that the points with short distances were similar in solar radiation, and one of them was regarded

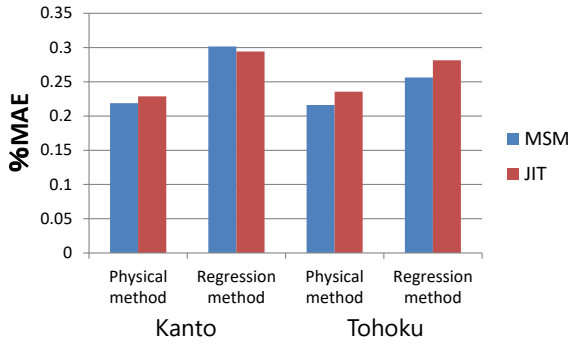

Fig. 12. %MAE of PV power prediction..

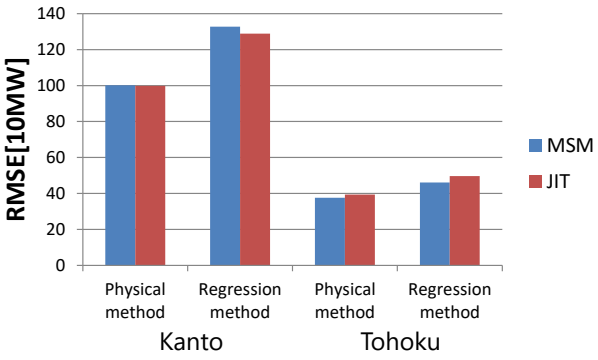

Fig. 13. RMSE of PV power prediction..

as a dependent variable and the regression equation deteriorated. For example, in this study, the regression coefficient of Choshi with a large PV introduction capacity became smaller, which indicates that the PV introduction capacity was not properly reflected in the regression coefficient.

The error distribution was investigated in more detail for the physical model with good accuracy. A histogram of the Kanto area is shown in Fig. 14, an enlarged version of the scale of the vertical axis is shown in Fig. 15, and the Tohoku area is also shown in Figs. 16 and 17. From Figs. 14 and 16, it can be seen that the error of most of the data are within ± 500 MW in the Kanto area where the annual averaged area power generation is 6150 MW at noon and ± 250 MW in the Tohoku area where the annual average is 2350 MW. The discussion in the previous chapter indicated that the accuracy of solar radiation forecast was better with MSM as for %MAE, and better with JIT Modeling as for fatal error. Figures 15 and 17 show the same tendency as in Figures 9 and 11 that PV power generation forecast with the inputs by JIT Modeling has less fatal error.

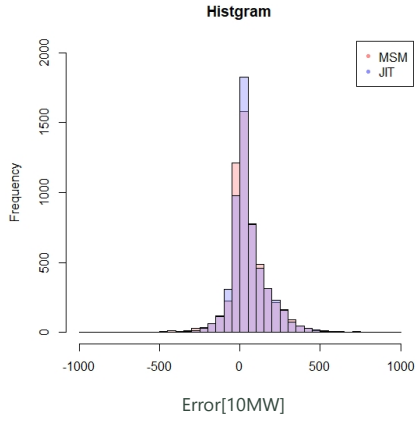

Fig. 14. Histgram of PV power predicted by physical method in Kanto.

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.

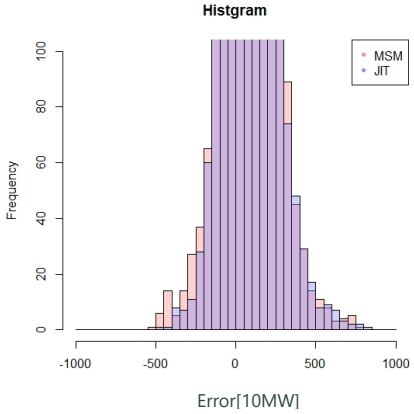

Fig. 15. Enlarged histgram of PV power predicted by physical method in Kanto.

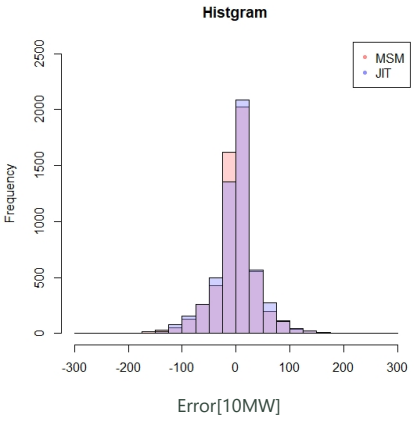

Fig. 16. Histgram of PV power predicted by physical method in Tohoku.

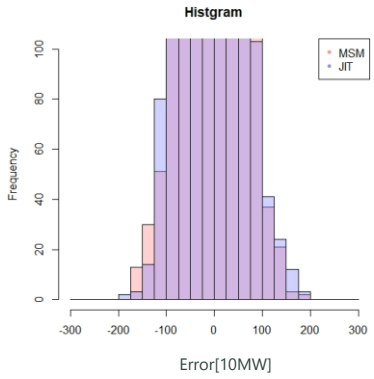

Fig. 17. Enlarged histgram of PV power predicted by physical method in Tohoku.

IV. Conclusion

Regarding the PV power over a wide area required for the JEPX spot market price forecast, this paper verified the method to convert the solar radiation forecast value at multiple points into the PV total power. Of the two proposed methods for area PV power prediction, the accuracy of the physical model was relatively good, with %MAE being about 22.5%. Considering that the accuracy of solar radiation prediction used for input is about 20%, it is considered that the accuracy degradation in the proposed PV power prediction method is small, so the proposed method is a relatively good conversion method.

As future work, we will apply the predicted value of PV total power by the proposed method to the merit order model, and investigate the effect of the accuracy of PV total power prediction on price prediction accuracy.

References

[1] JMA web site: http: //database.rish.kyoto-u.ac.jp/arch/jmadata/

[2] GPV DATA Archive: http://dias.tkl.iis.u-tokyo.ac.jp/gpv/

[3] Japan Meteorological Business Support Center: http://www.jmbsc.or.jp/

[4] Y. Kataoka, T. Kato, Y. Suzuoki, “A Study on Spatial Average Insolation Forecast in Electric Utility Service Areausing Meso-Scale Model Grid Point Value,” IEEJ Transactions on Power and Energy, Vol.133, No.6, pp.548–55. (Japanese)

[5] TEPCO: http://www.tepco.co.jp/forecast/html/area_data-j.html

[6] Tohoku Electric Power Co.,inc.: https://www.tohoku-epco.co.jp/comp/gaiyo/gaiyo.html .

[7] Agency for Natural Resources and Energy: . https://www.enecho.meti.go.jp/category/saving_and_new/saiene/statistics/past.html

Authorized licensed use limited to: Electricity Generating Authority of Thailand. Downloaded on August 23,2021 at 06:25:45 UTC from IEEE Xplore. Restrictions apply.