Prediction model for day-ahead solar insolation using meteorological data for smart grid

Min Hee Chung^{1,*}

¹School of Architecture and Building Science, Chung-Ang University, 06974 Seoul, Republic of Korea

Abstract. In the overseas market, power generation and energy service companies have been engaged in the business of providing personalized trading services for the production of electric power through the Internet platform. This is, so that the electric power sharing system between individuals is being developed through the Internet platform. The prediction of insolation is essential for the prediction of power generation for photovoltaic systems. In this study, we present a prediction model for insolation from data observed at the Meteorological Administration. We also present basic data for the development of the insolation prediction model through meteorological parameters provided in future weather forecasts. The prediction model presented is for five years of observation of weather data in the Seoul area. The proposed model was trained by using the feed-forward neural networks, taking into account the daily climatic elements. To validate the reliability of the model, the root mean square error (RMSE), mean bias error (MBE), and mean absolute error (MAE) were used for estimation. The results of this study can be used to predict the solar power generation system and to provide basic information for trading generated output by photovoltaic systems.

1 Introduction

Currently, the paradigm shifts from methods of transferring power generated by constructing a largescale power plant that was promoted in the past industrialization period through the ultra-high voltage power grid to a small-scale power generation using distributed power [1]. The application of smart grid is being actively promoted around the world. The smart grid will allow users to use the energy generated by renewable energy system or sell them on the market [2].

Photovoltaic systems are one of the renewable energy systems that have high applicability in urban area. Numerous buildings are installing and using these types of systems [3]. For peer to peer energy trading, it is necessary to decide in advance whether to use or sell the produced energy through the photovoltaic systems [4]. The prediction of insolation is essential to predict the energy produced by the photovoltaic system [5].

In this study, the insolation predictive model was proposed using the weather conditions experienced the day before. In order to determine the range of data for an optimal prediction model, we evaluated the prediction accuracy according to the range of the prediction data.

2 Method

We collected the meteorological data for the past 5 years from Seoul, South Korea. We excluded the missing data and selected the available data. The summary of data and prediction model are listed on Table 1. The artificial neural network (ANN), known as multi-layer feedforward neural network (MLF) was used for the estimation of total daily insolation using meteorological data such as maximum temperature (T_{max}), minimum temperature (T_{min}), humidity (RH), wind speed (WS), sunshine duration (SD), continued sunshine duration (CSD), precipitation (PR), and precipitation duration (PD). The MLF was implemented in MATLAB. The MLF consists of activation functions, bias and neurons. The data was divided into every other day for training and checking. The architecture of this model is shown in Fig.1.

Table 1. The summary of data and the prediction model.

Data location	Seoul, South Korea	
Year of data	2013~2017	
Number of days for training and checking	182, 541, and 901 days	
Input variables	T _{max} , T _{min} , RH, WS, SD, CSD, PR, PD	
Training method	Levenberg-Marquardt	
Number of hidden layers	1~5	
Number of hidden neurons	15~25	

In this research, several statistical indexes such as the root mean square error (RMSE), mean bias error (MBE), and mean absolute error (MAE). Equations used for computing the indexes are mentioned below $(1) \sim (3)$.

$$RMSE = \{1/n\sum (P-M)^2\}^{-2}$$
(1)

$$MBE = 1/n\Sigma(P - M)$$
(2)

^{*} Corresponding author: mhloveu@cau.ac.kr

[©] The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

$$MAE = 1/n\sum |P - M| \tag{3}$$

where, P is the predicted value, M is measured value, and n is the total number of measured days, repectively.



Fig. 1. A multilayer feedforward neural network architecture.

3 Result

In this study, the range of data was divided into 1 year, 3 years, and 5 years. The data in each range was divided into trains and checks in half. A total of 55 calculations were performed for each data set, with 1 to 5 hidden layers and 15 to 25 hidden neurons.

Table 2 lists the accuracy result of the optimal prediction model. The shorter the range of the analysis period, the shorter the calculation time and convergence to the result. However, the accuracy of the calculated values was low. The optimal model structure for 1 year, 3 year, and 5 year was 1 layer and 21 neurons, 2 layers and 20 neurons, and 3 layers and 20 neurons, respectively. This is clearly as shown on Table 2. As the number of processed data increases, the structure of ANN tends to become complicated.

Table 2. The result of the optimal prediction model

Scope of data	Number of hidden layer and neuron	RMSE (kWh/m²day)	MBE (kWh/m²day)	MAE (kWh/m²day)
1 year	1/21	1.716	-0.615	1.337
3year	2/20	1.643	-0.519	1.254
5year	3/20	1.615	-0.576	1.255

Based on the RMSE, the accuracy improved from 1.716 kWh/m²/day to 1.643 kWh/m²/day when the analysis data was increased from one year to three years, and improved from 1.643 kWh/m²/day to 1.615 kWh/m²/day when it was increased from three years to five years. Moreover, when the range of data increased from one year to three years, the values of MBE and MAE improved. This means that the accuracy of the prediction model was generally improved. On the other hand, there was no significant difference between MBE and MAE results at years three and five. However, when the number of analytical data was increased from three years to five years, it was concluded that the RMSE value had decreased, and thus the predicted value with a

large error was reduced. Fig. 2 plotted the comparison predicted value with the measured value for only one year out of the predicted value. Following the seasonal accuracy analysis, it was found that there was a large error in the summer especially when the insolation variation was severe.



Fig. 2. The example of the predicted and the measured daily insolation; (a) trained by 3 years and (b) trained by 5 years

4 Conclusion

When the smart grid is activated, peer to peer energy trading will take place. A photovoltaic system requires production amount according to the solar insolation. Therefore, it is necessary to predict solar insolation according to weather conditions. This study presents a model for predicting the solar insolation of the next day with the weather conditions of the previous day. At least three years of data is required to predict the optimum values of insolation. Moreover, it is necessary to correct the seasonal insolation in order to increase the prediction accuracy.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(NRF-2017R1C1B20117911)

References

- M. Manfren, P. Caputo, G. Costa, Appl. Energ., 88, 4 (2011)
- 2. C. Park, T. Yong, Enrgy. Proced., 128, (2017)
- 3. J. Hofierka, J. Kaňuk, Renew. Energ., 34, 10, (2009)
- 4. C. Zhang, J. Wu, C. Long, M. Cheng, Enrgy. Proced., **105**, (2017)
- 5. P. Mathiesen, J. Kleissl, Sol. Energy, 85(5) (2011)