VOL. 11, NO. 19, OCTOBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

CLIMATE EFFECT ANALYSIS ON SOLAR ENERGY GENERATION IN JEJU CITY

Junghoon Lee and Gyung-Leen Park
Department of Computer Science, Jeju National University, Jeju-City, Republic of Korea
E-Mail: jhlee@jejunu.ac.kr

ABSTRACT

This paper develops and assesses an area-specific prediction model for solar energy generation according to climate parameters on Jeju City, taking advantage of artificial neural networks. The hour-by-hour climate information, open to public in Rep. of Korea, are retrieved from the national weather administration, while the log file from a solar panel monitoring system is interpreted and systematically processed in the MySQL database. The neural network takes temperature, humidity, insolation, and sunlight duration as input, and the amount of solar energy generation is defined for output. Currently, 250 daily records are used for training, with 42 taken for evaluating. The developed model catches the day-by-day change of the power generation and its root mean square error is 16.4 *kwh*. After all, this model will estimate next day power generation based on the weather forecast, making it possible to build a power generation plan for base energy facilities or an electricity purchase plan from the main power system.

Keywords: smart grid, solar energy, prediction model, artificial neural network, open data.

1. INTRODUCTION

Modern power networks, called smart grids, are trying to extend the coverage of renewable energies as much as possible to pursue its final goal, namely, system-wide energy efficiency [1]. As the renewable energy sources, such as wind, sunlight, and the like, have enormous benefits, especially in eco-friendliness, many countries are accelerating the development of their own energies. However, they are commonly suffering from unpredictability. Thus, we cannot generate electricity when we want but only when sources allow. That is, electricity can be created when sufficient wind or bright sunlight is present. In power grids, energy generation can be planned and adjusted. Planned production is decided based on the prediction on next energy demand and available renewable energy. Dynamic energy production is mainly conducted by burning fossil fuels. Here, such expensive and pollutional generation is controlled by the regulation signal led by the continuous observation of real-time demand changes. Hence, it is necessary to precisely predict the availability of renewable energy sources and make an efficient power generation plan [2].

As the renewable energy is deeply dependent on weather conditions, the prediction must be built on top of the climate data. Most countries provide climate forecast services and increasingly open the past climate records to public. The effect of climate conditions to a specific renewable energy system will be different area by area, while the massive volume of history data makes it possible to build a prediction model by means of a variety of machine learning techniques. It is worth mentioning that the main motivation of the smart grid lies in the integration of such computational intelligence to the power system [3]. We can obtain the climate records of interest for a specific energy sources on a specific region and associate them with a power plant monitoring system, which generally keeps 1accumulating its operation status.

In the meantime, Jeju City, Rep. of Korea, is vigorously installing wind and solar energy generation facilities under its ambitious enterprise called *Carbon-Free Island 2030*. It is planning to replace all gasoline-powered vehicles with EVs (Electric Vehicles), and EVs will be charged by renewable energies as much as possible [4]. Moreover, the over-charged electricity can be sold back to the grid [5]. Recent power generators are essentially designed to periodically record their operation status. Accordingly, if we integrate the climate data provided from KMA (Korean Meteorological Administration) and the generator operation records, we can build a prediction model which estimates the amount of next day electricity from the solar panel. There are a variety of software components available for this process, including database engines for massive data handling, open libraries for machine learning, and statistics packages for enhanced data processing. The log file conversion is the last building block for the prediction model development.

In this regard, this paper collects the necessary data for the power generation of a solar energy system in Jeju City and builds an ANN (Artificial Neural Network)-based prediction model so as to efficiently integrate renewable energy sources into the main grid. It employs open software components such as MySQL, R, FANN (Fast ANN). The main focus is given to the climate parameters consisting of temperature, humidity, insolation, and sunlight hours. A new software application can be integrated or replace an existing one in this framework. It is also possible to design a new heuristic for the given smart grid service [6].

This paper is organized as follows: After issuing the problem in Section 1, Section 2 reviews related work. Section 3 explains the system model and how to process the raw data, while Section 4 tracks the effect of each parameter to the solar energy generation. Section 5 builds the prediction model and measures its accuracy. Finally, Section 6 concludes this paper with a brief introduction of future work.

2. RELATED WORK

A monitoring system is used for various purposes in solar energy production. Basically, it is necessary to continuously monitor and detect the malfunction of the target equipment, such as the electrical discontinuity between modules or unjustified rapid variation in the system status [7]. Moreover, for better operation efficiency, environmental parameters are collected, including temperature, solar radiation, and power production. The acquired information is stored locally and sent to a remote server for sophisticated statistical analysis. The monitoring equipment is also essential for the solar equipment to be integrated by electric vehicles (EVs), a new promising consumer of renewable energy [8]. Here, the monitoring unit is installed on the roof of a vehicle and manages the connection of a PV (Photovoltaic) cell to the EV battery. With the management of this connection built upon the solar energy monitor, the driving range of EVs has been extended.

W. Ya¨ ıci *et al*. investigates the accuracy and the robustness of the baseline ANN method for a solar energy system working in Ottawa, Canada, specifically, at Conmet Energy Research Centre [9]. The authors take the input parameters of time-of-day, 6 previous tank temperatures as well as ambient system temperature along with both horizontal and inclined pyranometers. As output variables, their model picks 6 next tank temperatures and solar fraction. Solar fraction means the portion of solar energy to total energy delivered to a tank. Their experiment finds out that the optimal prediction model comes from the combination of the Levenberg-Marquardt learning algorithm, 9/8/7 inputs, and 20 hidden layers. The authors claim that the configuration suppresses the maximum error below 19 % for solar fraction prediction. Here, the log file is processed to extract daily insolation, collector thermal efficiency, and energy consumption from the propane-fired storage tank.

M. Detyniecki *et al*. focus on how to decide the operation of solar energy systems, specifically, whether to consume the stored energy now or keep saving for later use [10]. Their main purpose lies in the combination of a solar panel with a maximum power point tracker (MPPT), which maintains the generation system on the optimal performance. It is based on the weather condition given from the National Oceanic and Atmosphere Administration. Weather conditions are symbolically classified into sunny, fair, partly cloudy, mostly cloudy, shower, and the like. In addition, an Arduino module acquires the data from the target solar panel and communicates with a dedicated laptop computer. The amount of energy according to the weather condition is trained on fuzzy decision trees. At each expansion of the tree, an attribute is selected by the Shanon entropy. The authors claim that this scheme reduces the energy prediction error by 22 % compared with a constant prediction mechanism.

3. DATA PROCESSING

Figure-1 illustrates our data acquisition and processing platform. To begin with, through the KMA website open to public (*http://sts.kma.go.kr*), we can download climate archives. In this site, the record is provided on an hourly basis since KMA has begun its service. A record consists of wind speed, wind direction, atmospheric pressure, cloudiness, humidity, precipitation, insolation and sunlight duration. Some of them are relevant to wind, and some to solar energy generation. The web site allows us to select the target area. Jeju City is located on an island, which has a bunch of tourist attractions and is famous recently for a variety of its smart grid challenges. After hour-by-hour records are downloaded as text files, we develop a C language program to convert the text to SQL statements, which combine insert commands and retrieved text fields. By running them, our MySQL running on a Linux machine, stores the hourly climate records of Jeju City for the last 25 years. Here, even though this server supports a Hadoop-based big data analysis [11], the amount of data to handle in this paper is not so much to employ the Hadoop family yet.

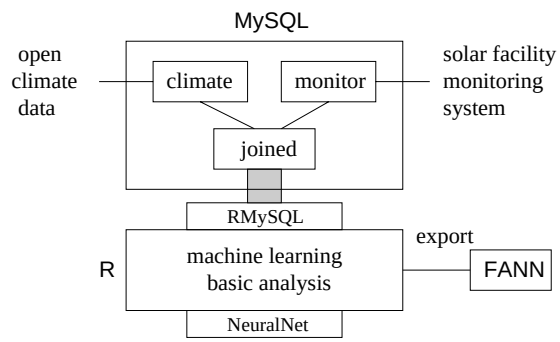

Figure-1. Data processing framework.

A local company sells building-scale solar energy systems and has developed their monitoring system. On each device, a microcontroller board captures several sensor readings and reports to a remote server on every 10 minutes, during the period when the solar system produces electricity. Each record includes voltage, current, PV generation, accumulated generation amount, and system status. We develop again a C language program to read the log file and convert to SQL statements. Here, in the monitoring equipment, a record can be made only after a sequence of multiple polls. Even a single erroneous response from the apparatus should invalidate the entire record. With this program, we can make a database table for daily power generation in a target solar system, and now the table includes the records from 2015-05-08 to 2016-02-26. Then, both climate and power generation tables are joined to create a new table consisting of average temperature, average humidity, total insolation, total sunlight duration, and the total amount of produced power for each day.

For a sophisticated data analysis, the jointed table is downloaded to the R space, which is enriched by abundant sets of libraries, especially in visualizing, mining, and integrating other applications [12]. RMySQL enables us to develop an R script which selects the remote MySQL

server, table, and issues an SQL statement. We can easily test the dependency between parameters with R and reissue the SQL statement to alter the field of interest. Here, our framework also combines the FANN utility for fast and vigorous data analysis specialized in ANN [13].

4. PER-PARAMETER ANALYSIS

This section measures the effect of individual climate parameters to the amount of solar power generation based on the database table obtained by the procedure described in the previous section. To begin with, the daily power generation is plotted in Figure-2. Point 0 in the x-axis corresponds to 2015-05-08. This curve indicates that the amount of solar generation is dependent on the seasonal effect quite much. During the winter, the overall generation is cut down. Even in the same season, severe porcupine-style fluctuations are observed. Moreover, two adjacent days show quite different power generation, the largest gap being 159 *kwh* for a single target generator. The drop-down is very sharp considering that the daily maximum is 195 *kwh*. The solar energy seems to be more affected by climate conditions compared with the wind energy. Actually, it is well known that insolation and sunlight duration are two main parameters for solar energy generation. Those climate parameters have less interdependency between two consecutive days, unlike the wind speed. In cloudy days, the power generation is stuck below 35 *kwh*.

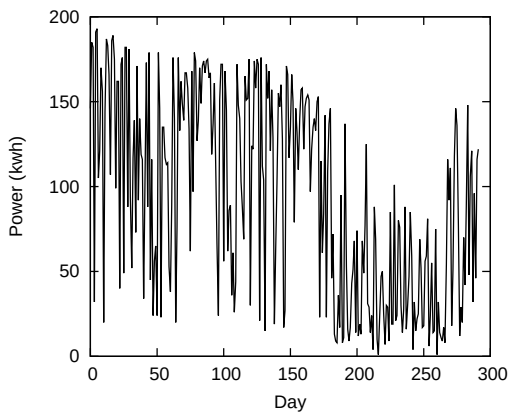

Figure-2. Daily amount of solar power generation.

Figures from Figure-3 to Figure-6 plot the effect of each climate parameter one by one. First, Figure-3 plots the effect of temperature. The correlation between temperature and generated power is calculated as 0.577. This value implies average-level interdependency. In spite of a relatively high density when the temperature exceeds 30 °*C* and the power generation is more than 120 *kwh*, dots distribute over the whole rectangular area. In addition, there are many days generating less than 35 *kwh*. They are not influenced by temperature as they appear across the whole temperature range. Next, Figure-4 plots the effect of humidity. The correlation between humidity and generated power is calculated as -0.264, which indicates quite low interdependency between two variables. Here, most days have humidity higher than 40 % in Jeju City. In the rectangular area where most humidity levels reside, the dot distribution looks almost uniform.

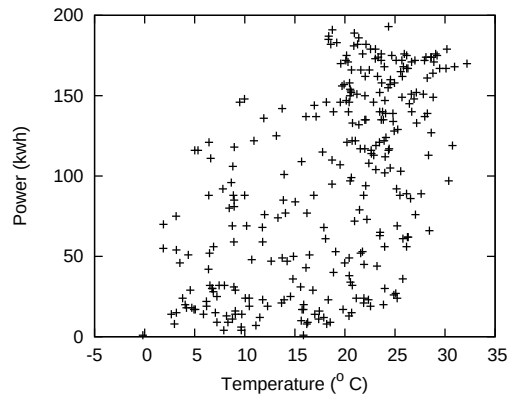

Figure-3. Effect of temperature.

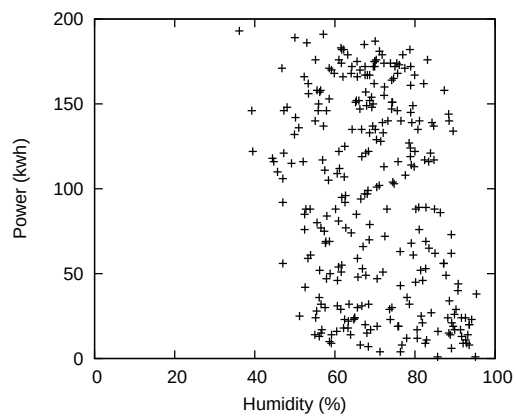

Figure-4. Effect of humidity.

Now, Figure-5 plots the effect of insolation. The correlation between insolation and generated power is calculated as 0.963, and this is a very strong interrelation. As can be seen in the figure, the dependency seems to be linear. When the insolation is less than 5 MJ/m^2 or greater than 25 MJ/m^2, dots line up almost straight. In the middle area, the thickness, namely, the gap between the maximum and the minimum on the same insolation gets larger. Figure-6 plots the effect of sunlight duration. The correlation between sunlight duration and generated power is calculated as 0.931. It also indicates high interdependency. The dot distribution looks linear, but a little bit less strong than the case of insolation. In Figure-5 and Figure-6, relatively many dots are placed on the smaller amount of solar generation. On 31 days, the daily sum of sunlight durations is observed to be 0.

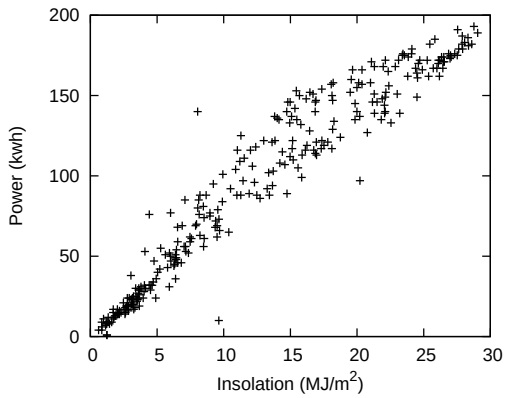

Figure-5. Effect of insolation.

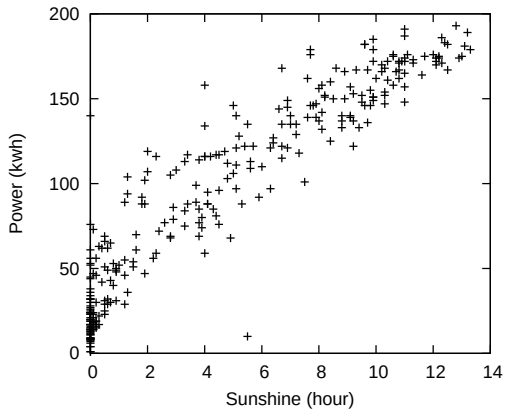

Figure-6. Effect of sunlight duration.

5. PREDICTION MODEL DEVELOPMENT

This section builds a prediction model for the amount of solar energy generation according to the climate parameters. As shown previously, some parameters seem to have almost linear dependency, while other parameter dependencies are hard to specify. ANN can model the complex nonlinear behaviour of target objects, mainly based on the learn-by example principle, and makes it unnecessary to prove an intricate mathematical formulation [14]. It is also exploited in the performance prediction of solar systems [15]. Actually, even though many statistical packages provide ANN-based modelling mechanisms, our experiment discovers that FANN converges fast and reduces the fitting error, or residuals, for the solar energy prediction [13]. Moreover, its library and provide an enriched set of API functions to C or Java applications.

With 292 day-by-day records we use 250 for training and the others for assessing the accuracy. A C language program, implemented on top of the FANN library, makes an ANN model learn 250 patterns. Here, the database table is retrieved and converted to the set of learning patterns according to the FANN file specification. Definitely, the amount of daily generation is the output variable, while the input layer includes temperature, humidity, insolation, and sunlight duration. The number of hidden layers is set to 20 by trial-and-error. Figure-7 compares the fitted solar power generation with the actual amount, the maximum error reaching 60.9 *kwh*, which corresponds to 43.5 %. The root mean square error (RMSE) for fitting is calculated as 8.58 *kwh*.

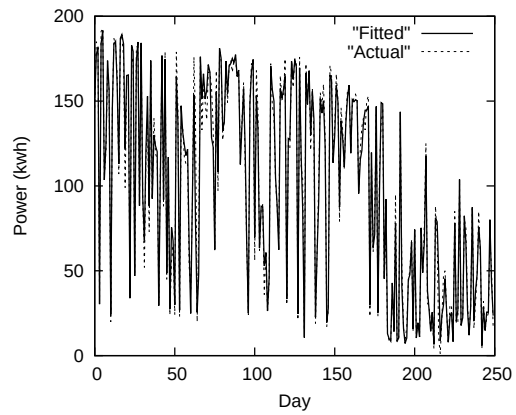

Figure-7. Fitting result.

Next, Figure-8 plots the prediction result for those days which are not used in the training phase. The figure indicates that the ANN model catches the flow of the time series quite well, intuitively. The maximum prediction error reaches 51.6 *kwh* while its RMSE is estimated as 16.4 *kwh*. In Figure-7 and Figure-8, those days having a large prediction error enlarge RMSE. Figure-9 demonstrates that for 73 % of test days, the prediction error is less than 12 *kwh*. Here, each error is mapped to one of 3 *kwh*-intervals to plot the error distribution. After all, the climate parameters have a straightforward effect on the amount of solar energy generation and the accuracy of climate forecast is the critical prerequisite for the accuracy of our model. Exact prediction will help the grid to make an efficient electricity production plan for base power plant facilities.

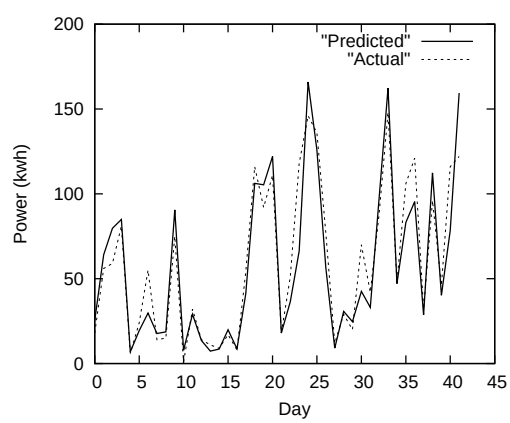

Figure-8. Prediction result.

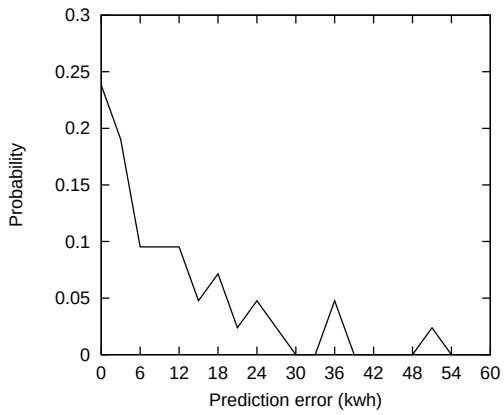

Figure-9. Probability distribution of prediction errors.

6. CONCLUSIONS

In this paper, we have built a data processing framework for the solar energy generation in Jeju City, Rep. of Korea. Log files accumulated in a solar panel monitoring system are parsed to calculate the daily amount of electricity production, while hour-by-hour climate parameters are retrieved from KMA. An SQL program joins both tables in MySQL, and an R script downloads the joined records in its workspace. Then, a FANN-based program makes its ANN learn by the given training data set to build a solar energy prediction model, taking temperature, humidity, insolation, and sunlight duration as input. The last two have a very strong correlation, over 0.93, to the solar energy generation, as expected. The experiment has found out that the ANN model catches the flow of the time series quite well, while RMSEs for fitting and prediction are estimated as 8.48 *kwh* and 16.4 *kwh*, respectively, where the maximum and the average generation amounts are 195 *kwh* and 94.7 *kwh*. This prediction will make it possible to build a next day power generation plan for base energy facilities such as nuclear or heat power plants.

As future work, we are planning to acquire more solar energy generation records to refine the prediction model as well as to embrace more monitoring systems over Jeju City to look into hidden geographical factors. In addition, our database has now collected many smart grid-related spatiotemporal streams such as SoC (State-of-Charge) dynamics for EV driving, climate condition changes including wind speed, charger operation monitors, and the like [16]. Hence, we will synthesize them to create more promising smart grid business models and investigate their interrelations.

ACKNOWLEDGEMENTS

Prof. Gyung-Leen Park is the corresponding author.

This research was supported by the 2016 scientific promotion program funded by Jeju National University.

REFERENCES

[1] Chakraborty C., Ho-Ching Iu H., Dah-Chuan Lu D. 2015. Power converters, control, and energy management for distributed generation. IEEE Transactions on Industrial Electronics. 62(7): 4466-4470.

[2] Passow K., Ngan L., Littmann B., Lee M., Panchula A. 2015. Accuracy of energy assessment in utility scale PV power plant using Plant Predict IEEE 42nd Photovoltaic Specialist Conference.

[3] Yaici W., Entchev E. Prediction of the performance of a solar thermal energy system using adaptive neuro-fuzzy inference system. 2014. 3rd International Conference on Renewable Energy Research and Applications. pp. 601-604.

[4] Bhattarai B., Levesque M., Maier M., Bak-Jensen B., Pllai J. 2015. Optimizing electric vehicle coordination over a heterogeneous mesh network in a scaled-down smart grid testbed. IEEE Transactions on Smart Grid. 6(2):784-794.

[5] Bayram I., Shakir M., Abdallah M., Qaraqe K. 2014. A survey on energy trading in smart grid. IEEE Global Conference on Signal and Information Processing. pp. 258-262.

[6] Lee J., Park G. 2015. A heuristic-based electricity trade coordination for microgrid-level V2G services. International Journal of Vehicle Design. 69(1/2/3/4):2018-223.

[7] Visconti P., Cavalera G. 2015. Intelligent system for monitoring and control of photovoltaic plants for optimization of solar energy production. IEEE 15th International Conference on Environment and Electrical Engineering. pp. 1933-1938.

[8] Schuss C., Eichberger B., Rahkonen T. 2012. A monitoring system for the use of solar energy in electric and hybrid electric vehicles. IEEE International Instrumentation and Measurement Technology Conference. pp. 524-527.

[9] Ya¨ıci W., Entchev E., Longo M., Brenna M., Foiadelli F. 2015. Artificial neural network modelling for performance prediction of solar energy system. 4thInternational Conference on Renewable Energy Research and Applications. pp. 1147-1151.

[10] Detyniecki M., Marsala C., Krishnan A., Siegel M. 2012. Weather-based solar energy prediction. IEEE World Congress on Computational Intelligence.

[11] Goiri I., Le K., Nguyen T. D., Guitart J., Torres J., Bianchini R. 2012. Green Hadoop: Leveraging green energy in data-processing frameworks. Proceedings of EuroSys.

[12] Brunsdon C., Comber L. 2015. An Introduction to R for Spatial Analysis & Mapping. SAGE Publication Ltd.

[13] Nissen S. 2005. Neural Network Made Simple. http://fann.sourceforge.net/fann_en.pdf Software 2.0.

[14] Methaprayoon K., Yingvivatanapong C., Lee W., Liao J. 2007. An integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty. IEEE Transactions on Industry Applications. 43: 1441-1448.

[15] Kalogirou S., Mathioulakis E., Belessiotis V. 2014. Artificial neural networks for the performance prediction of large solar systems. Renewable Energy. 63: 90-97.

[16] Lee J., Park G., Cho Y., Kim S., Jung J. 2015. Spatio-temporal analysis of state-of-charge streams for electric vehicles. 4th ACM/IEEE International Conference on Information Processing in Sensor Networks. pp. 368-369.