

Research Article

An Improved Method for Sizing Standalone Photovoltaic Systems Using Generalized Regression Neural Network

Tamer Khatib and Wilfried Elmenreich

Institute of Networked and Embedded Systems and Lakeside Labs, University of Klagenfurt, 9020 Klagenfurt, Austria

Correspondence should be addressed to Tamer Khatib; tamer.khat@hotmail.com

Received 14 May 2014; Revised 9 July 2014; Accepted 23 July 2014; Published 11 August 2014

Academic Editor: Dimitrios Karamanis

Copyright © 2014 T. Khatib and W. Elmenreich. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this research an improved approach for sizing standalone PV system (SAPV) is presented. This work is an improved work developed previously by the authors. The previous work is based on the analytical method which faced some concerns regarding the difficulty of finding the model's coefficients. Therefore, the proposed approach in this research is based on a combination of an analytical method and a machine learning approach for a generalized artificial neural network (GRNN). The GRNN assists to predict the optimal size of a PV system using the geographical coordinates of the targeted site instead of using mathematical formulas. Employing the GRNN facilitates the use of a previously developed method by the authors and avoids some of its drawbacks. The approach has been tested using data from five Malaysian sites. According to the results, the proposed method can be efficiently used for SAPV sizing whereas the proposed GRNN based model predicts the sizing curves of the PV system accurately with a prediction error of 0.6%. Moreover, hourly meteorological and load demand data are used in this research in order to consider the uncertainty of the solar energy and the load demand.

1. Introduction

Photovoltaic systems are environment-friendly energy systems. Thus, PV system installation has been given a big concern in the last three decades. However, PV systems' high capital cost is considered one of the most important challenges to this technology, especially when it is compared with conventional power systems. Therefore, many research works are being conducted in order to propose methods for optimization of PV systems so as to provide reliable systems with minimal capital cost. In this context, Sharma et al. in [1] define PV system optimization as "the process for determining the cheapest combination of PV array and battery that will meet the load requirement with an acceptable availability level over the expected lifetime." As a fact, PV system performance depends on available solar energy and ambient temperature and, therefore, meteorological variables must be extensively studied in order to optimally size a PV system [2].

In the literature, works related to PV system sizing can be categorized into intuitive, numerical, and analytical methods. The intuitive method is defined according to [3] as conducting simplified calculation of the system size without

taking into account the random nature of solar radiation or establishing a relationship between the different subsystems. The method considers the use of solar energy data such as the lowest monthly average, average annual, or monthly solar energy. A major disadvantage of the intuitive method is that it may result in an over- or undersizing of the designed system which results in low reliability of the system or high cost of produced energy [3]. In the meanwhile, numerical method is defined as implementing system simulation for each time period considered. In this method, daily or hourly data are used and based on that the energy balance and flow of the system are calculated. The simulation method offers the advantage of being more accurate, and the evaluation of system availability can be approached in a quantitative manner. System availability in this case is defined as the load percentage satisfied by the PV system for long period of time [3]. The simulation method allows optimizing the energy and economic cost of the system. Simulation methods can be divided into two, namely, stochastic and deterministic methods. In the stochastic method, the uncertainty in solar energy and load demand is considered by simulating a random process modelling hourly solar radiation and load demand

records. Due to the difficulties in finding the hourly solar energy models [3–5], the deterministic simulation method considers the use of predefined load and meteorological data. As for the analytical method, the equations describing the PV system size as a function of reliability have to be developed. This method has the advantage of providing simple calculation of the PV system size, while the disadvantage is the difficulty in finding the coefficients of the equations which are location dependent [6, 7].

Due to the difficulty in calculating the optimum PV size by the simulation and analytical methods, artificial neural networks (ANN) are employed to overcome these limitations. For optimizing PV systems in many regions in Algeria, a combined numerical and ANN method has been used [8, 9]. In this method, the optimal PV sizing factors of the targeted sites are first calculated using the numerical method first. After that the ANN based model is utilized for predicting these factors using the geographical location coordinates. The developed ANN model has two input variables, namely, latitude and longitude, and two outputs, namely, the sizing factors for the PV array (C_A) and the storage battery (C_s), respectively. The ANN model helps in simplifying the calculation of the sizing factors but its limitation is that it can only determine the optimum size of PV system at one reliability level or loss of load probability (LLP). In [10], the analytical method is used to obtain a large data set for optimum sizes of a PV system at different LLPs and this data set is used to train an ANN to predict the optimum size of the PV array in terms of the optimum storage battery, LLP, and “yearly cleanses index.” The main drawback of this method is that the ANN model predicts the optimum PV array size depending on the optimum storage battery capacity which is not explained as to how it is calculated. In [9], the same combined numerical and ANN method presented in [11] has been used but this time it is used for generating the sizing curves at different LLPs for certain sites in Algeria. Two ANN models have been developed in which the first ANN model has four inputs which are latitude, longitude, altitude, and LLP and thirty outputs representing the thirty possible C_A values. After predicting the C_A , the C_s is calculated and then the sizing curve is predicted by the second ANN model. In this method, the sizing curve is predicted by ANN only after determining the optimum pair of C_A and C_s , thus making the procedure laborious and impractical.

To overcome the limitations of the abovementioned methods in determining optimal sizing of PV systems, we propose an improved approach using a general regression neural network (GRNN) to predict the PV array and battery capacities in terms of LLP, latitude, and longitude. By using the GRNN model, the calculation for the optimum PV size of a standalone PV (SAPV) system can be automated and improved without the need for extensive mathematical calculations or graphical analysis techniques.

2. Analytical Method for Sizing SAPV System

The background theory and formulation used in the analytical method for sizing SAPV system as proposed in [7]

are described in this section. The analytical method is based on a PV system energy model and a long term of meteorological data such as solar energy and ambient temperature. In general, a typical SAPV system consists of a PV module/array, power conditioner comprising charge controller and maximum power point tracking controller, storage battery, inverter, and load. A PV module collects energy from the sun and converts it to DC power which then is handled by a power conditioner to supply loads. The energy produced by a PV array is given by

$$E_{PV} = A_{PV} \cdot E_{\text{sun}} \cdot \eta_{PV} \cdot \eta_{\text{inv}} \cdot \eta_{\text{wire}}, \quad (1)$$

where A_{PV} is the area of the PV module/array and E_{sun} is daily solar irradiation. η_{PV} , η_{inv} , and η_{wire} are efficiencies of PV module, inverter, and conductors, respectively.

PV module efficiency depends on cell temperature and it can be given as function of reference efficiency (η_{ref}) which is provided in the data sheet, cell temperature, (T_C) and standard testing temperature (T_{ref}) as illustrated below

$$\eta_{PV} = \eta_{\text{ref}} [1 - \beta_{\text{ref}} (T_C - T_{\text{ref}})]. \quad (2)$$

β_{ref} is a factor and it can be given by

$$\beta_{\text{ref}} = \frac{1}{T_o - T_{\text{ref}}}. \quad (3)$$

As for cell temperature, it can be calculated using ambient temperature as follows:

$$T_C = T_a + \frac{\text{NOCT} - 20}{800} G_{\text{ref}}, \quad (4)$$

where T_a is the ambient temperature, G_{ref} is the standard testing solar radiation, and NOCT is the nominal operation cell temperature.

The difference between the energy at the front end of a PV system E_{PV} and at the load side is given by

$$E_D = \sum_{i=1}^{8760} (E_{PV} - E_L), \quad (5)$$

where E_L is the load energy demand.

E_D may have a positive value ($E_{PV} > E_L$) or a negative value ($E_{PV} < E_L$). In case of having E_D positive, there is an excess in energy in the system. In the meanwhile, if E_D is negative then there will be an energy deficit. Excess energy is usually stored in batteries in order to be used in the deficit time. The energy deficit is defined as the disability of the PV system to fulfil the load demand at a specific time.

System availability (reliability) is an important issue to be considered in designing of PV system. 100% availability of a PV system means that the system is able to cover the load demand all the year time without shortages. Consequently, 99% availability means that the system is not able to cover the load demand in 88 hours during one year time. This means that high PV system availability leads to high reliability and vice versa. However, high reliable PV system results high initial cost and, thus, it is not feasible to consider very high

availability rates in designing PV system. The availability of a PV system can be as a loss of load probability (LLP) index. LLP is defined as the ratio of annual energy deficit to annual load demand and it is given by

$$LLP = \frac{\sum_{i=1}^{8760} \text{Energy deficits}_i}{\sum_{j=1}^{8760} \text{Energy demand}_j}. \quad (6)$$

In [7], an analytical method is presented for optimal sizing of a standalone PV system. This method is represented by two empirical equations between PV array sizing ratio (C_A) and battery sizing ratio (C_s) and system reliability LLP. In [7], the authors assumed that the relation between C_A and LLP can be expressed by two exponential terms while the relation between C_A and C_s is linear as follows:

$$\begin{aligned} \text{optimum } C_A &= c_1 e^{c_2 LLP} + c_3 e^{c_4 LLP}, \\ \text{optimum } C_s &= c_5 C_A + c_6. \end{aligned} \quad (7)$$

The sizing ratios C_A and C_s are calculated as follows:

$$\begin{aligned} C_A &= \frac{C_{PV}}{C_L}, \\ C_s &= \frac{C_B}{C_L}, \end{aligned} \quad (8)$$

where C_B and C_{PV} are battery capacity and PV array capacity at specific load, respectively, and C_L is the load demand.

The optimization process presented in [7] starts by defining some initial values for load demand, PV efficiency, charging efficiency, inverter efficiency, and conductor efficiency. After that, daily solar irradiation for the targeted site is utilized in order to calculate the expected output power of the system. After that, a design space that contains a set of PV array area values is initiated. Based on each PV array area value and the defined load demand, C_A is then calculated. Then, E_D is calculated using (5). Subsequently, arrays of deficit and excess energies are constructed. At this point, at each specific PV array area, LLP, C_s , and C_A are calculated and stored in arrays. Consequently this loop is repeated until the maximum value of PV array area is reached. Finally, plots of LLP versus C_A and C_s versus C_A are constructed and from these plots, curve fitting equations are derived using the MATLAB fitting toolbox to find the coefficients of (7).

3. GRNN for Sizing SAPV System

Artificial neural networks (ANNs) are nonalgorithmic information processing systems which are able to learn and generalize the relationship between input and output variables from the recorded data. In this work we apply a GRNN model to improve the method presented in [7]. The aim of the proposed GRNN model is to predict the sizing curves directly without the need to run any iterative simulation and to abandon the need for calculating models coefficient. A schematic diagram of the basic architecture of a GRNN is shown in Figure 1. The network has several layers: the input,

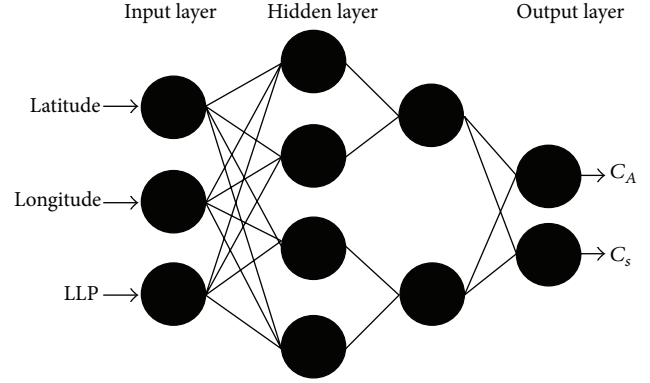


FIGURE 1: Topology of the GRNN used to predict the optimum size of a SAPV.

hidden, and output layers. Each layer is interconnected by connection strengths, called weights.

A generalized regression neural network (GRNN) is a probabilistic neural network consisting of an input layer, a hidden layer, a pattern/summation layer, and a decision node. Each predictor variable has a corresponding input neuron. The input values standardize the input values by subtracting the median and scaling the value to the interquartile range. The input layer feeds the hidden neuron layers where each training pattern is represented by a hidden neuron. In the pattern layer, there are only two neurons, a denominator summation unit and a numerator summation unit. The denominator summation unit adds up the weights of the values coming from each of the hidden neurons. The numerator summation unit adds up the weights of the values multiplied by the actual target value for each hidden neuron. The decision node divides the values accumulated by the numerator summation unit by the value in the denominator summation unit and produces the predicted target value of the GRNN. The advantage of GRNNs is simplicity, fast training, good approximation also with smaller training sets, and, thus, high efficiency in comparison to other networks [12].

In general, there is no rule to determine the optimum number of hidden nodes in the hidden layer without training several networks and estimating the generalization error of each one. Large number of hidden nodes resulted in high generalization error due to the overfitting and high variance. In the meanwhile, low number of hidden units causes large training and generalization error due to underfitting and high statistical bias [13]. Anyways, there are some “rules of thumb” for selecting the number of the hidden nodes in the literature. Blum in [14] suggests that the number of neurons in the hidden layer is supposed to be somewhere between the input layer size and the output layer size. Swingler in [15] and Berry and Linoffin [16] claim that one does not require more than twice the number of the inputs. In addition, Boger and Guterman in [17] suggest that the number of the hidden nodes can be 70–90% of the number of the input nodes. Caudill and Butler in [18] recommended that the number of hidden nodes should be equal to the number of the inputs

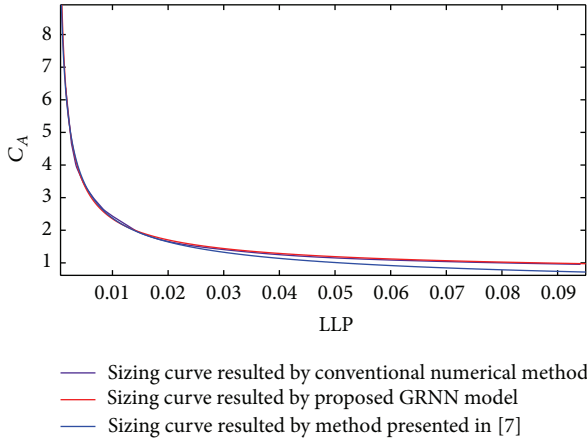


FIGURE 2: PV array sizing curve.

plus the number of the outputs multiplied by $(2/3)$. Based on these recommendations the recommended number of the hidden layer for our model is in the range of 2 to 4 hidden neurons. In this paper we used 4 hidden nodes.

In this research, three variables are used as input parameters for the input nodes of the input layer, latitude, longitude, and LLP. Two nodes are at the output layer, namely, C_A , C_s , which are the optimum sizing ratios of the PV array and battery. We used the data set from [7] to train the ANN with the Levenberg-Marquardt backpropagation algorithm. The analytical method proposed in [7] is used to design PV system at 11 reliability levels (LLP (0–10%)) for each year of each station. The training of the proposed model is done utilizing MATLAB “nftool” and using 75% of the provided data set. These data were divided into three parts 70% for training, 15% for internal validation, and 15% for internal training.

4. Results and Discussion

As mentioned before, in [7], five sites are considered in the conducted optimization. In this research, the developed GRNN is trained by using four of these sites, namely, Johor Baharu, Kuching, Ipoh, and Alor Setar, while the fifth site which is Kuala Lumpur is used for testing the developed model. Figure 2 shows a comparison of sizing curves for the PV array obtained by numerical simulation, equations presented in [7], and the developed GRNN.

Figure 2 indicates that the GRNN model is more accurate in predicting the size of the PV array than the results based on the equations presented in [7]. The average mean absolute percentage error values were 1.2% and 5.1%, respectively. Using the GRNN, the difficulty of calculating the coefficients in the equations proposed by Khatib et al. [7] is avoided. Figure 3 shows the predicted storage battery by GRNN compared to the one calculated by the numerical simulation. The prediction accuracy was very high which indicates the utility of the use of GRNN in sizing PV systems.

From these figures, it is concluded that the proposed model is able to predict system sizing curve using only the

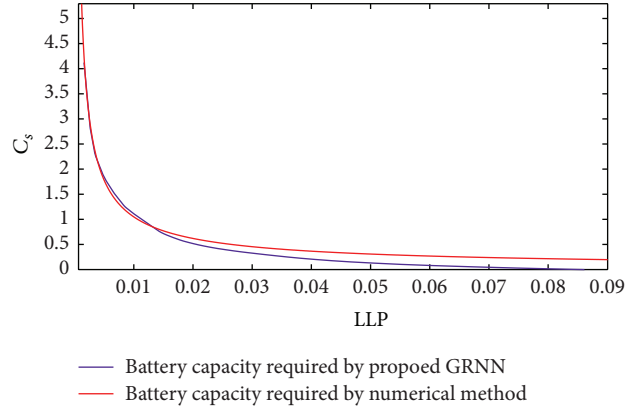


FIGURE 3: Battery capacity at different LLP.

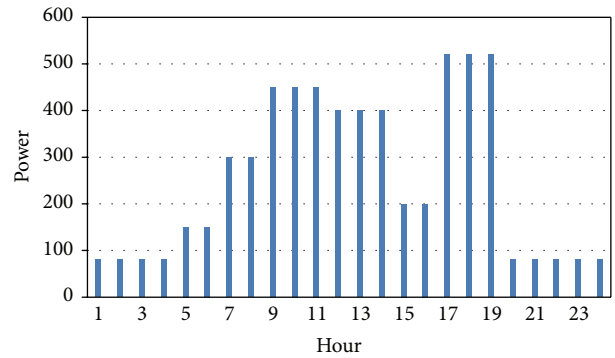


FIGURE 4: Simulated load demand.

location coordinates as well as the loss of load probability which can be considered as an advantage when it is compared to other ANN based models such as the models presented in [8–11]. Anyway, to validate the proposed method in terms of system reliability a design example is conducted in this paper considering the proposed method and some previously published methods. In [5, 7] optimization of PV system is done for Malaysia. However to validate the optimization results both authors designed a PV system supplying a 2.215 kWh/day load being located at Kuala Lumpur at 0.01 LLP. However, Shen in [5] did not generalize his results for other load demands while Khatib et al. in [7] did not consider the uncertainty of solar energy and the variation of the load demand. Therefore, to prove the validity of the proposed procedure and to avoid the limitations of [5, 7] an hourly load demand occurring at 24 hours and hourly solar energy data are used. In order to model the respective load demand of a typical remote area, we consider a larger load of 6.130 kWh/day in this paper. Figure 4 shows the temporal distribution of the simulated load demand. In this paper the hourly model of PV system in [19] is used to validate the proposed method.

Based on the developed ANN, the optimum sizing ratios for the considered site (Kuala Lumpur) are $C_A = 2.02$ and $C_s = 0.793$. We calculate the size of the PV based on the assumption that the used PV module has 200 Wp

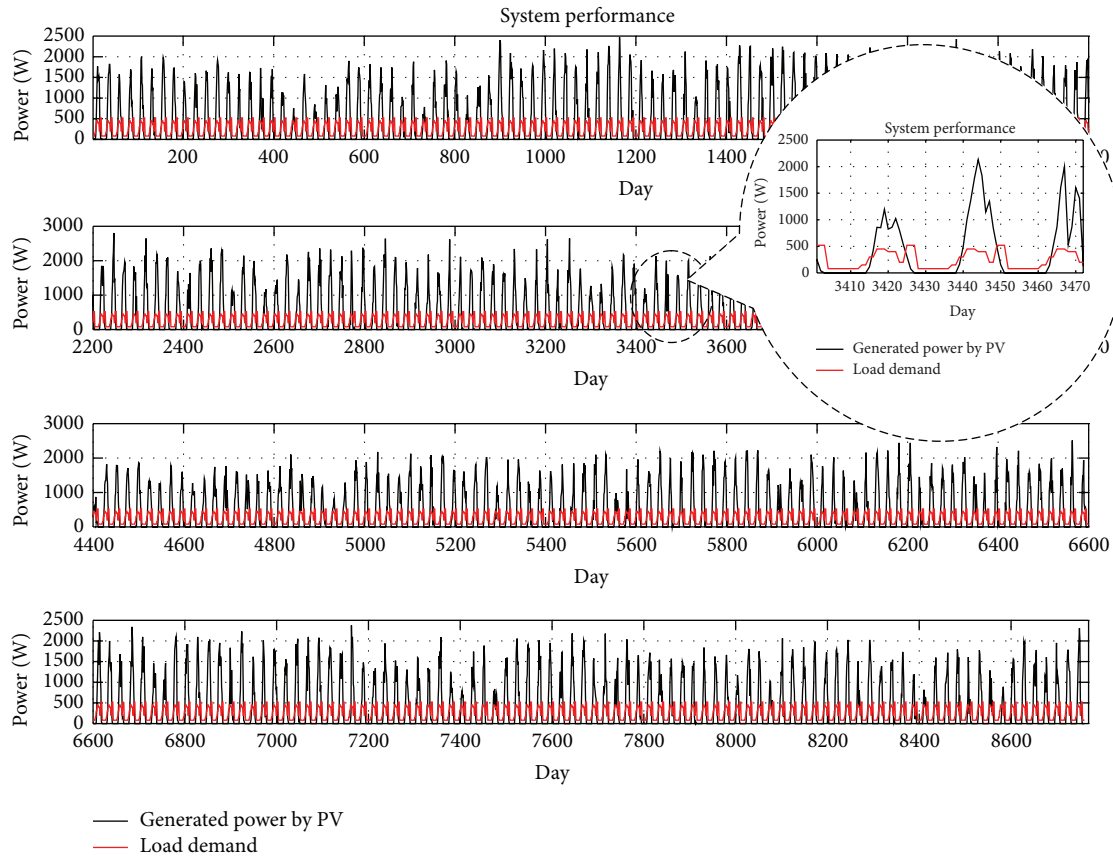


FIGURE 5: Generated power by the designed system.

rated power, 1.4 m² area, and 16% conversion efficiency (as reference value) and the rated battery voltage operates on 12 volts with a charging efficiency of 80% and an inverter conversion efficiency of 90%. The required PV array and battery capacities are 2.5 kWp and 324 Ah/12 V, respectively. The power generated by the proposed photovoltaic system is calculated with respect to the load (see Figure 5). The average power generated by the designed PV array is 1.075 kW. This power is supposed to cover the load demand while the excess power is used to charge the battery. In the case of fully charged batteries, the excess power will be dumped using a dumping load. Figure 6 shows the power balance of the system in which negative power indicates that the battery is used to supply the load, while the positive net power indicates that this power has to be used in charging the battery or to be dumped. Figure 7 shows the net power which needs to be dumped. The sum of the dumped energy per year is 1757 kWh.

Figure 8 shows the state of charge (SOC) of the battery storage for a year (1–8760 hours). A SOC value of 1.0 indicates that the battery is not used while SOC value of less than 1.0 means that the battery is used. As indicated in the figure, the battery periodically supplies power to cover the load demand during the night time. Through one year, the battery reaches its allowable minimum SOC (0.2) 87 times. Figure 9 shows the loss of load days. The figure shows the percentage of covered load demand during loss of load days. During

a year, the load is lost for 87 hours which corresponds to 99% availability during one year. Note that most of load loss incidences happen in January, February, and December. From the figure, the loss of load probability is 0.5%.

In [7], the authors used the same sizing ratios and location to design a PV system to supply a 2.215 kWh/day load considering a 1% LLP. The authors used to simulate this designed system daily solar radiation values and daily averages of load demand. According to [7], the LLP of the designed system is 0.95%, while in this research the use of hourly solar radiation and load demand resulted in a 0.5% LLP despite that we aimed to design at it 1%. This slight difference can be due to the differences in the used metrological data. From these points two conclusions can be stated. First, the proposed sizing ratio for Malaysia in [5, 7] and in this paper is proven even by considering the uncertainty in solar energy and variation in the load demand. In addition, the using of hourly meteorological data and load demand definitely yields more accurate optimization but it could cause a slight oversizing in certain cases.

5. Conclusion

An ANN model is used to facilitate the use of a developed method for sizing PV system for Malaysia. The proposed

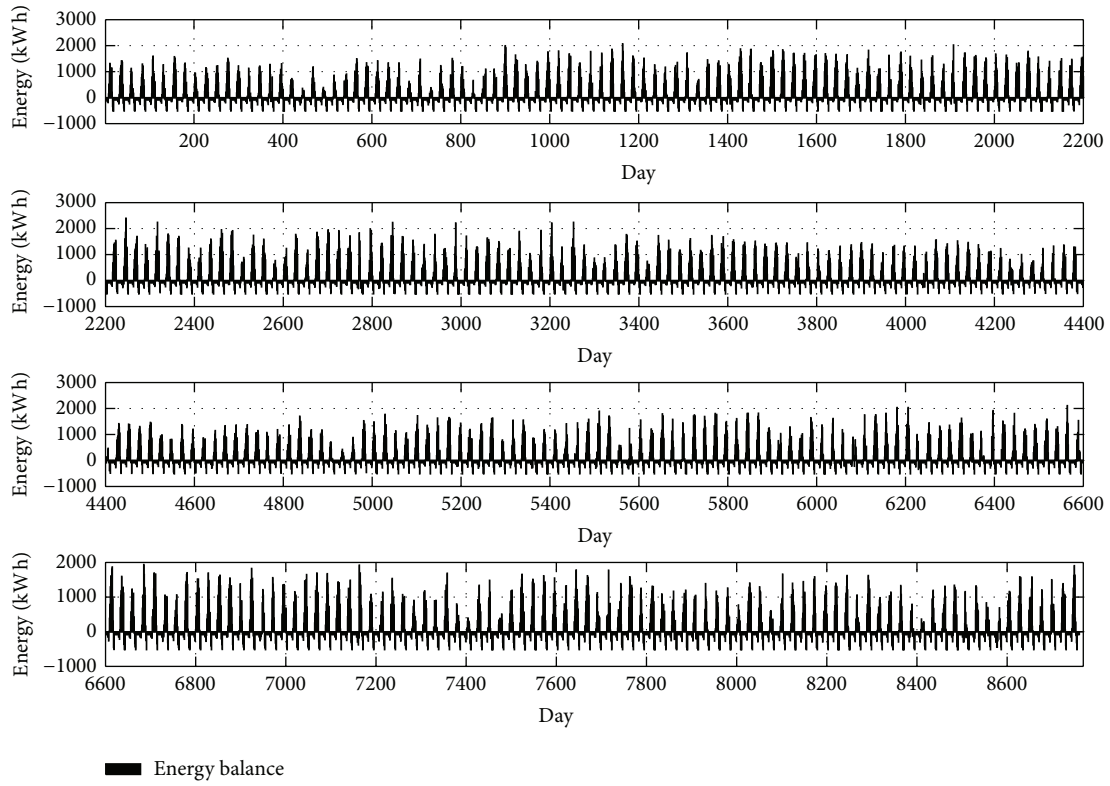


FIGURE 6: Energy balance for designed PV system.

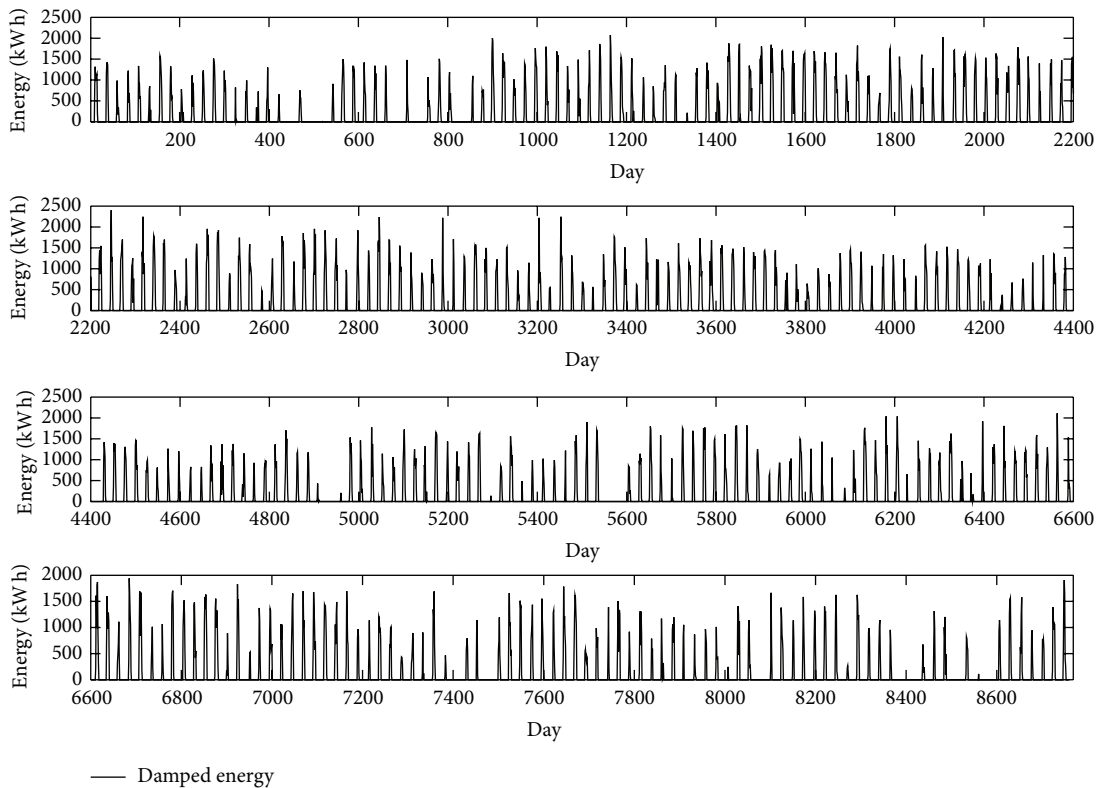


FIGURE 7: Dumped power for the designed PV system.

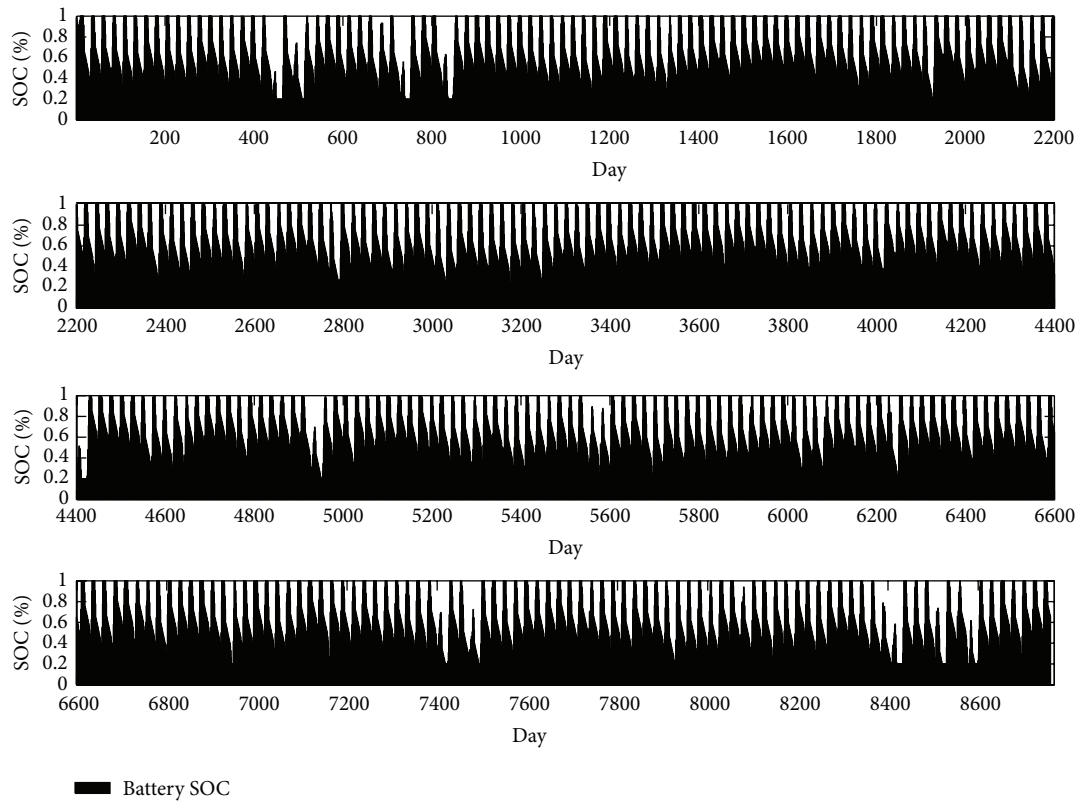


FIGURE 8: Battery storage SOC for the designed PV system.

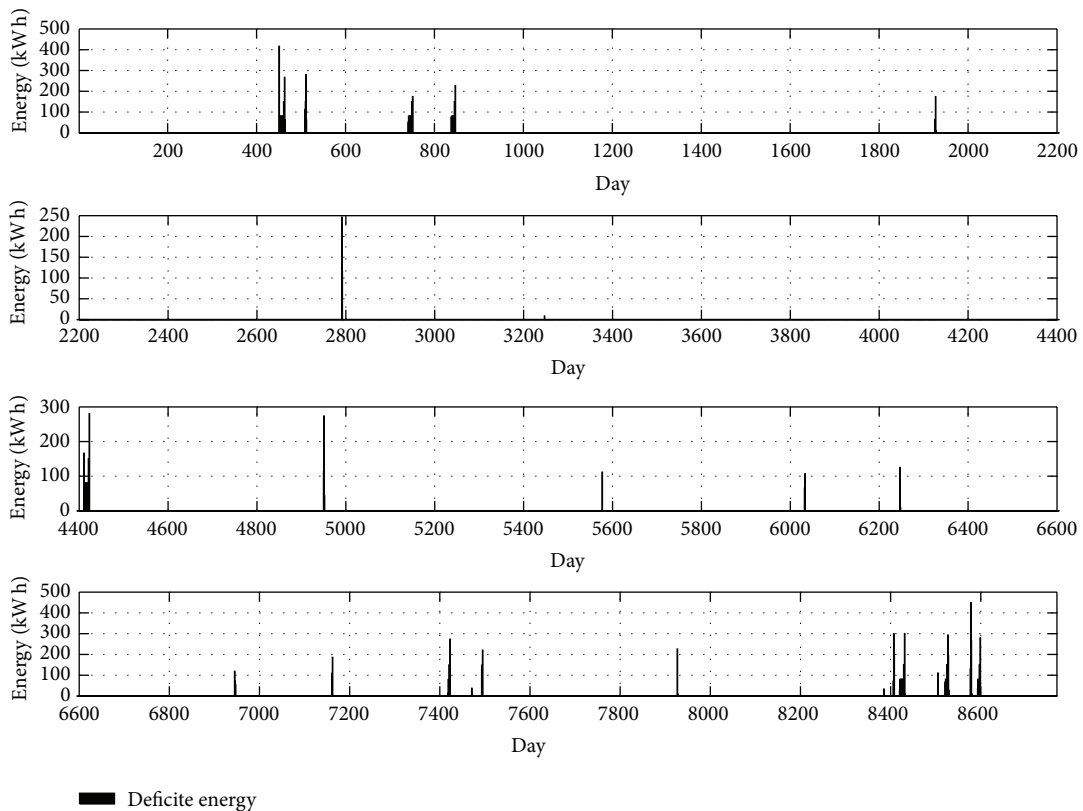


FIGURE 9: Loss of load days for the designed PV system.

ANN model predicts the size of the PV system in terms of LLP, latitude, and longitude. The developed ANN model showed high accuracy in predicting the PV system size whereas the MAPE is 0.6%. However, to ensure the validity of the proposed method, a designed example for a specific load is conducted considering the uncertainty in the solar radiation and the variation of the load demand. To validate the designed system we used a simulation based on hourly solar radiation and load demand. As a result, the LLP of the designed system is found to be 0.5% which indicates a sufficiently high reliability of the designed system.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by Lakeside Labs, Klagenfurt, Austria, and funded by the European Regional Development Fund (ERDF) and the Carinthian Economic Promotion Fund (KWF) under Grants KWF 20214|23743|35470 (Project MONERGY) and 20214|22935|34445 (Project Smart Micro-grid).

References

- [1] V. K. Sharma, A. Colangelo, and G. Spagna, "Photovoltaic technology: basic concepts, sizing of a stand alone photovoltaic system for domestic applications and preliminary economic analysis," *Energy Conversion and Management*, vol. 36, no. 3, pp. 161–174, 1995.
- [2] T. Khatib, A. Mohamed, and K. Sopian, "A review of photovoltaic systems size optimization techniques," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 454–465, 2013.
- [3] M. Sidrach-de-Cardona and L. M. López, "A simple model for sizing stand alone photovoltaic systems," *Solar Energy Materials and Solar Cells*, vol. 55, no. 3, pp. 199–214, 1998.
- [4] A. Hadj Arab, B. Ait Driss, R. Amimeur, and E. Lorenzo, "Photovoltaic systems sizing for Algeria," *Solar Energy*, vol. 54, no. 2, pp. 99–104, 1995.
- [5] W. X. Shen, "Optimally sizing of solar array and battery in a standalone photovoltaic system in Malaysia," *Renewable Energy*, vol. 34, no. 1, pp. 348–352, 2009.
- [6] T. Markvart, A. Fragaki, and J. N. Ross, "PV system sizing using observed time series of solar radiation," *Solar Energy*, vol. 80, no. 1, pp. 46–50, 2006.
- [7] T. Khatib, A. Mohamed, K. Sopian, and M. Mahmoud, "A new approach for optimal sizing of standalone photovoltaic systems," *International Journal of Photoenergy*, vol. 2012, Article ID 391213, 9 pages, 2012.
- [8] A. Mellit, M. Benghanem, A. H. Arab, and A. Guessoum, "An adaptive artificial neural network model for sizing stand-alone photovoltaic systems: application for isolated sites in Algeria," *Renewable Energy*, vol. 30, no. 10, pp. 1501–1524, 2005.
- [9] A. Mellit, "ANN-based GA for generating the sizing curve of stand-alone photovoltaic systems," *Advances in Engineering Software*, vol. 41, no. 5, pp. 687–693, 2010.
- [10] L. Hontoria, J. Aguilera, and P. Zufiria, "A new approach for sizing stand alone photovoltaic systems based in neural networks," *Solar Energy*, vol. 78, no. 2, pp. 313–319, 2005.
- [11] A. Mellita and M. Benghanem, "Sizing of stand-alone photovoltaic systems using neural network adaptive model," *Desalination*, vol. 209, no. 1–3, pp. 64–72, 2007.
- [12] C. H. Dagli, A. L. Buczak, J. Ghosh, M. J. Embrechts, O. Ersoy, and S. Kercel, *Intelligent Engineering Systems Through Artificial Neural Network*, ASME, 2000.
- [13] S. Geman, E. Bienenstock, and R. Doursat, *Neural Networks and the Bias/Variance*, 1992.
- [14] A. Blum, *Neural Networks in C++*, John Wiley & Sons, New York, NY, USA, 1992.
- [15] K. Swingler, *Applying Neural Networks: A Practical Guide*, Academic Press, London, UK, 1996.
- [16] M. Berry and G. Linoff, *Data Mining Techniques*, John Wiley & Sons, New York, NY, USA, 1997.
- [17] Z. Boger and H. Guterman, "Knowledge extraction from artificial neural networks models," in *Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics. Part 3 (of 5)*, pp. 3030–3035, Orlando, Fla, USA, October 1997.
- [18] M. Caudill and C. Butler, *Understanding Neural Networks: Computer Explorations*, vol. 1 of *Basic Networks*, The MIT Press, Cambridge, Mass, USA, 1993.
- [19] T. Khatib and W. Elemenreich, "Novel simplified hourly energy flow models for photovoltaic power systems," *Energy Conversion and Management*, vol. 79, pp. 441–448, 2014.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

