Research article

Integration of the radial basis functional network and sliding mode control for the sunshine radiation forecast

Ming-Tang Tsai*, Chih-Jung Huang

Department of Electrical Engineering, Cheng-Shiu University, Kaohsiung, Taiwan

* Correspondence: Email: k0217@gcloud.csu.edu.tw; Tel: +188677310606; Fax: +188677315367.

Abstract: In this paper, we propose a forecasting system of sunshine radiation for planners to quickly and accurately predict the output of solar power. The field data, including observation time, temperature, relational humidity, wind speed and global radiation, were collected, and the data clusters were embedded in the Excel Database. To improve the computational performance, the data selection technique was used in the stage of data cleaning, data integration and data reduction. Using the Integration of the Radial Basis Function Network (RBFN) and Sliding Mode Control (SMC), a Sliding Mode Radial Basis Function Network (SMRBFN) was proposed to solve this forecasting problem. Since the Sliding Mode Control has the design’s sense of optimal parameters, three parameters in the SMRBFN were dynamically adjusted to promote the accurate and reliability of forecasting system. Linking the SMRBFN and Excel database, the learning stage and testing stage of SMRBFN retrieved the input data from Excel Database to perform and analyze the forecasting system. The proposed algorithm was tested on Kaohsiung district in summer and winter. The average prediction error of MAPE and RMSE obtained from the forecasting results are about 9% and 0.223, respectively. It can be proved that SMRBFN can efficiently forecast the sunshine radiation and accurately provide the output of solar power in an uncertainty environment.

Keywords: forecasting system; sunshine radiation; data selection; sliding mode control; radial basis function network; solar power; mean absolute percentage error; root mean square error

1. Introduction

The CO₂ emission of power sectors in Taiwan is about 1/3 of total CO₂ emission, indicating the
significance of power sectors in the global warming issue. In 2022, Taiwan’s energy import
dependency was about 97% and the proportion of electricity consumption accounts for 51.2% of the
final energy consumption electricity consumption [1]. In Taiwan, as an isolated island, the power grid
is not connected to other national grids. Thus, 100% of the electricity must be produced by itself, and
stable energy and power supply are very important for Taiwan. When the electricity supply is
insufficient or the system fails, it will be accompanied by serious economic losses. Whether electric
energy can be fully supplied has become the deep expectation of people. Renewable Energy (RE) is needed
for local energy markets to provide an important option of energy production in the near future [2]. Due to
advances in solar energy technologies, solar power is currently considered one of the most rapidly
increasing resources [3]. It is no doubt that the benefit of solar power is attracting many utilities in the
electricity market [4].

Solar energy is called to play an important role in the future energy supply of the Taiwan. The
development of solar power has become one of the main alternative energy sources. It not only
provides alternatives to new energy sources, but also reduces environmental pollution caused by fossil
fuels. The biggest challenge comes from the intermittent nature of the photovoltaic, which is the
unpredictable nature and dependence on weather and climate conditions [5,6]. Due to the instability of
the natural climate, the output of solar power is also discontinuous and intermittent. The instability of
the solar power generation is easily affected by factors such as sunshine, clouds and rainfall changes.
It leads to the increase of power supply instability in the power grid. Therefore, how to accurately
forecast the output of solar power in the uncertainty environment is a very important issue for power
management, demand planning and power security [7,8].

As it is an important research field of solar power forecasting, several approaches have been
reported in the past. Ref [9] evaluated the solar forecast quality improvement for a single-point forecast
model by incorporating these expert variables. Ref [10] developed a probabilistic solar power
forecasting method based on weather scenario generation considering inherent correlation among
different weather variables. Ref [11] proposed a statistical method to solve the problem of stationarity
of photovoltaic production data, and then predicted the power output model of solar power. Ref [12]
analyzed the prediction of three different variables of photovoltaic, ambient temperature and wind
speed, constructed a conversion model of photovoltaic and estimated the power generated at the connection
terminal of the grid. Ref [13] presented a variety of time-series methods including deep-learning algorithm
and machine learning algorithms to predict the photovoltaic power generation output for quick respond
to equipment and panel defects. Ref [14] proposed an interpretable probabilistic model for short-term
solar power forecasting using natural gradient boosting. Ref [15] proposed the hybrid cascaded forecasters
network model to enable accurate Solar Power Generation Forecasting (SPGF) for the next 24 h. To help
resolve the weather forecast error, Ref [16] introduced a SolarPredictor model to train genuine weather
forecasts. Ref [17,18] utilized the historical database to predict the solar power using the machine learning
and Support Vector Machine (SVM). Integrational Multilayer feedforward Neural Network (MFFNN) and
Multiverse Optimization (MVO) was proposed to efficiently predict the PV output power [19].
However, the solar power forecasting is not easy to use due to various influences, such as climate
factors and seasonal factors. Solar power is heavily affected by the weather, so the solar power
forecasting is strongly volatile with the changing weather [20]. Solar power forecasting could rise up
hundred times the normal value to reflect the volatility. It is complicated to perform solar power
forecasting, especially when finding the best strategy in a huge amount of data. Therefore, we use data
selection technology [21,22] to calculate the difference of the meteorological database in the same
period, so as to reduce the amount of training data, thereby shortening the calculation time and
improving the accuracy.
We propose a forecasting system of sunshine radiation for planner to quickly and accurately predict the output of solar power. The output of solar power is based on the prediction of sunshine radiation as the research goal. The data, including observation time, temperature, relational humidity and sunshine radiation in the Kaohsiung Weather Bureau [23], were collected and the data cluster are embedded in the Excel Database dependent. To improve the computational performance, the data selection technique was used at the stage of data cleaning, data integration and data reduction.

Bad data are handled in the data selection technique. Integration of the Radial Basis Function Network (RBFN) [24,25] and Sliding Mode Control (SMC) [26,27], a Sliding Mode Radial Basis Function Network (SMRBFN) is proposed to solve the forecasting problem. Since the Sliding Mode Control has the design’s sense of optimal parameters, the parameters in the SMRBFN were dynamically regulated to promote the accurate and reliability of prediction. Linking the SMRBFN and Excel database, the learning stage and testing stage of SMRBFN retrieved the input data from Excel Database to perform and analyze the efficiency and accuracy of the forecasting system. The proposed algorithm was tested on Kaohsiung district in summer and winter. Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) were calculated in the forecasting procedure. It is demonstrated that SMRBFN can efficiently forecast the sunshine radiation and accurately provide the output of solar power.

2. Data selection

The output of solar power output will change with time and weather information. The dominant factor affecting is the solar irradiance intensity. The output of solar power from a Photovoltaics (PV) module can be calculated as Eq (1) [28]:

\[ P_s(t) = K_{py} \times A_{py} \times P_G(t) / 3600 \]  

Eq (1)

\( P_s(t) \) represents PV output power at time \( t \) (\( w \)); \( P_G(t) \) is sunshine radiation at time \( t \) (\( j / m^2 \)); \( A_{py} \) is the area of the PV array (\( m^2 \)); \( K_{py} \) is the efficiency of PV.

The predicted data of sunshine radiation related attributes include time, temperature, relative humidity and the hour of sunshine. The sunshine radiation will change with time and weather information, as it is a very huge meteorological database. In this paper, a data selection technique was used to find meaningful information in the huge database. The data selection is to find the data group with the most similar temperature, relative humidity, sunshine hours and sunshine radiation, and identify the similarity of the data as shown in Eq (2). With chosen reference days, the similarity sequencing is carried out for the database according to the reference data, and the sequenced data were then integrated into a data set.

\[ \Delta H_i = \sqrt{(H_{f}^{T} - H_{D,i}^{T})^2 + (H_{f}^{R} - H_{D,i}^{R})^2 + (H_{f}^{S} - H_{D,i}^{S})^2 + (H_{f}^{G} - H_{D,i}^{G})^2} \]

Eq (2)

\( H_f^T \) : Temperature of current weather

\( H_f^R \) : The relative humidity of the current weather

\( H_f^S \) : The sunshine hours of current weather

\( H_f^G \) : The global radiation of current weather
The degree of similarity of the $i-$th data

The data for training is selected after sorting and calculation as shown in Eq (3).

$$H_n = \text{sort}\{\Delta H_i\} \quad i = 1, 2, ..., m \quad \& \quad n = 1, 2, ..., m$$

$H_n$: Data similarity calculated after sorting

$n$: Ranking of the similarity of the data

In this paper, data selection is used as the preprocessing of the meteorological database, and more meaningful data are selected from the historical database, as shown in Figure 1.

Figure 1. Flowchart of data selection in historical data.

3. Sliding mode control

Sliding Mode Control (SMC) is essentially a kind of nonlinear control. It can be continuously changed according to the current state in the dynamic process, so that the system can be controlled at a certain level. The lower edge of the characteristic moves up and down with small amplitude and high frequency along the specified track. The SMC can be designed to work with the system disturbances, which can make the sliding mode motion system be robust.

The properties of SMC are that in the state space of the system $\dot{x} = f(x), x \in \mathbb{R}^n$. It is a tangent plane, $s(x) = s(x_1, x_2, \cdots, x_n) = 0$, which divided the state space into upper part $s > 0$ and lower part. Three situations for the moving point on the tangent plane were defined as shown in Figure 2.

1. Normal point: When the system movement point moves to the vicinity of the tangent plane $s = 0$, it passes through this point A.
2. Starting point: When the system movement point reaches near the tangent plane $s = 0$, it leaves point B from both sides of the tangent plane.
3. Ending point: When the system movement point reaches near the $s=0$, it tends to point C from both sides of the tangent plane.
In the sliding mode structure, normal point and starting point do not have much meaning, but the ending point does have a special meaning. Because all moving points in a certain area are ending points, they will be attracted to move in this area. At this time, the area where all the moving points are the ending points on the tangent plane \( s = 0 \) is called the sliding mode area, and the system moving in this area is called the sliding mode action. According to the requirement that all moving points on the sliding mode area must be ending points, when the moving point reaches the tangent plane \( s(x) = 0 \), they must be formulated in Eqs (4) and (5).

\[
\lim_{s \to 0^+} s \leq 0 \quad \text{or} \quad \lim_{s \to 0^-} s \geq 0 \quad (4)
\]

\[
\lim_{s \to 0^+} s \leq \lim_{s \to 0^-} s \quad (5)
\]

The basic problem of SMC assumes that there is a control system as shown in Eq (6).

\[
\dot{x} = f(x, u, t) \quad \& \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, t \in \mathbb{R} \quad (6)
\]

The control parameters \( u \) is defined in Eq (7).

\[
u = \begin{cases} u^+(x) & s(x) > 0 \\ u^-(x) & s(x) < 0 \end{cases}, \quad u^+(x) \neq u^-(x) \quad (7)
\]

The performance of SMC must satisfy the following conditions.

1. Satisfying the accessibility condition, all moving points outside the tangent plane \( s(x) = 0 \) will reach the tangent plane within a limited time.
2. Ensure the stability of sliding mode action.
3. Meet the dynamic quality requirements of the control system.

4. SMRBFN

SMRBFN consists of the input, hidden and output layers. In the input layer, four variables are connected to the hidden layer. The output layer is defined the “sunshine radiation”. The weights \( w_k \) connecting the \( k \)-th hidden node with the output node. Since the parameter selection of the RBFN has a great influence on the searching procedure, the parameters, the weight ( \( w \) ), the center ( \( C \) ) and the
smoothing parameter ($\sigma$), which were often adjusted by trial and error. In order to adjust three parameters, the adaptability of the SMC is adopted to make the RBFN faster reaches the termination point. The SMRBFN structure is shown in Figure 3, the SMC process is performed in the SMRBFN.

**Figure 3.** The construct of SMRBFN.

4.1. **Input layer**

$X_i$ is the i-th variable of the input layer. For each training data, set input matrix $X = [x_{ji}]_{M \times N}, j = 1, 2, ..., M, i = 1, 2, ..., N$. In this paper, four historical data, including temperature, relative humidity, the hour of sunshine and sunshine radiation, were collected and the data clusters were embedded in the Excel Database according to the season, week and day. Using data selection, those historical data are classified and identified into the similarity of the data. The similarity data are then integrated into a data set for the input layer.

4.2. **Hidden layer**

In the hidden layer, $C_j = [c_{j1}, ..., c_{jk}, ..., c_{jK}]$ is called the j-th center of SMRBFN. $\| x_{ji} - c_{jk} \|$ is the Euclidean distance between the i-th node of the input layer and the k-th node of hidden layer. The Euclidean distance is determined by Eq (8). The k-th hidden layer output is defined as Eq (9).

\[\| x_{ji} - c_{jk} \|^2 = \sqrt{N}(x_{ji} - c_{jk})^2 \]  
\[H_{jk} = \varphi_{jk}\sqrt{\frac{1}{N}(x_{ji} - c_{jk})^2} \]  
\[\varphi(x) = e^{-x^2/\sigma^2} \]
In the Eq (10), the function $\phi(x)$ is Gaussian distribution function and $\sigma$ is smoothing parameter.

### 4.3. Output layer

In the output layer, let $w_{jk}$ be the weight between hidden node $H_{jk}$ and output node $y_j$ and the $j$-th output of output layer be as Eq (11).

$$y_j = \sum_{k=1}^{K} w_{jk} H_{jk}$$

(11)

Calculate the error between simulation output $y_j$ and its expected value $T_j$ by error function. Error function is defined as Eq (12).

$$e_j(n) = \left(T_j - y_j(n)\right)^2 = \left(T_j - \sum_{k=1}^{K} w_{jk}(n) \exp \left(-\frac{||X_j(n) - C_{jk}(n)||^2}{\sigma_{jk}^2(n)}\right)\right)^2$$

(12)

e_j(n) and $y_j(n)$ are the $j$-th error and the $j$-th simulation output of $n$-th epoch, respectively.

SMRBFN is used to optimize the three parameters, weights $w_{jk}$, the center of $C$ and the smoothing parameters $\sigma_{jk}$, and the parameters would refine the accuracy in the dynamic environment and can yield a minimum forecast error. The output layer is defined as “sunshine radiation”. To evaluate the accuracy for SMRBFN, the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are all used in this paper. The MAPE and RMSE are defined as

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{Rad_{\text{true}} - Rad_{\text{predict}}}{Rad_{\text{true}}} \right| \times 100\%$$

(13)

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left| Rad_{\text{true}} - Rad_{\text{predict}} \right|^2}$$

(14)

$Rad_{\text{true}}$ is the actual value of global radiation at hour $t$ and $Rad_{\text{predict}}$ is the forecasting value of sunshine radiation at hour $t$. $T$ is the number of testing data. Figure 4 is the flowchart of proposed algorithm.
5. Results and analysis

The data from the historical meteorological database of Kaohsiung area were used to train and test the proposed algorithm. Many tests, including daily/weekly in summer/winter, were conducted. For comparison purposes, SMRBFN and RBFN were also built for tests. The simulation was implemented with Matlab on an Intel(R) Core (TM) i7-2320 computer with 8 GB RAM.

5.1. Data selection

After the data selection procedure, the number of selected data in the database is set to 20, 40 and 60 at the same period as shown in Table 1. From the Table 1, it can be seen that when the number of selected data is set to 40 in summer, the error value is the lowest. The MAPE and RMSE are 9.3481% and 0.2993, respectively. In winter, the error value of the test is the lowest when the number of selected data is set to 40. The MAPE and RMSE are 11.456% and 0.2536, respectively. The executed time of all tests is under 0.04 s.
Table 1. The results for the number of selected data.

<table>
<thead>
<tr>
<th></th>
<th>The number of selected data</th>
<th>The number of training data</th>
<th>The number of testing data</th>
<th>MAPE (%)</th>
<th>RMSE (kW)</th>
<th>Executed time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td>20</td>
<td>18</td>
<td>2</td>
<td>10.2438</td>
<td>0.3045</td>
<td>0.01707</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>36</td>
<td>4</td>
<td>9.3481</td>
<td>0.2993</td>
<td>0.01926</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>54</td>
<td>6</td>
<td>12.2497</td>
<td>0.3336</td>
<td>0.03665</td>
</tr>
<tr>
<td>Winter</td>
<td>20</td>
<td>18</td>
<td>2</td>
<td>14.8421</td>
<td>0.3637</td>
<td>0.01779</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>36</td>
<td>4</td>
<td>11.456</td>
<td>0.2536</td>
<td>0.02153</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>54</td>
<td>6</td>
<td>16.1487</td>
<td>0.3829</td>
<td>0.04168</td>
</tr>
</tbody>
</table>

Figure 5 shows the sunshine radiation forecast of the various data selection in summer. In Figure 5, the deviation from the actual value is more if the number of data selected is 20. It can track very closely the actual value if the number of data selected is 40 and 60, respectively. The forecasting curve is close to the actual curve when the number of data selected is 40. It is proved that the data selection technique could track the actual data with the less data. This characteristic can be utilized to greatly reduce the training time, and the data storage can be reduced without losing originalities. We can minimize the data storage, shorten the preprocessing needs and reduce the network size.

5.2. Daily sunshine radiation forecast

Figure 6 shows the daily sunshine radiation forecast in summer. The forecasting results of SMRBFN can track very closely the actual value with SMC applied. It can be seen that SMRBFN has the capability to follow the spikes and it is not easily attainable by RBFN. Similarly, Figure 7 is the daily sunshine radiation forecast in winter. The forecasting results of SMRBFN can also track very closely the actual value with SMC applied.
Figure 6. The daily sunshine radiation forecast in summer.

Figure 7. The daily sunshine radiation forecast in winter.

Table 2 shows the daily forecasting errors of GRNN, SMRBFN and RBFN for comparison in summer and winter. In winter, the MAPE and RMSE of GRNN are 6.8594% and 0.1896, respectively, the MAPE and RMSE of SMRBFN are 3.9854% and 0.1066, respectively, and the MAPE and RMSE of RBFN are 8.6719 and 0.2404, respectively. It is obvious that SMRBFN has the ability to find better solutions. Although the errors are larger in summer, the prediction effect of SMRBFN is better than the other algorithms, and the error value is smaller.

Table 2. The daily forecasting errors of GRNN, SMRBFN and RBFN.

<table>
<thead>
<tr>
<th>Method</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAPE (%)</td>
<td>RMSE (kW)</td>
</tr>
<tr>
<td>GRNN</td>
<td>18.3106</td>
<td>0.6954</td>
</tr>
<tr>
<td>RBFN</td>
<td>19.3417</td>
<td>0.6549</td>
</tr>
<tr>
<td>SMRBFN</td>
<td>16.0651</td>
<td>0.5588</td>
</tr>
<tr>
<td>GRNN</td>
<td>6.8594</td>
<td>0.1896</td>
</tr>
<tr>
<td>RBFN</td>
<td>8.6719</td>
<td>0.2404</td>
</tr>
<tr>
<td>SMRBFN</td>
<td>3.9854</td>
<td>0.1066</td>
</tr>
</tbody>
</table>

GRNN: General regression neutral network.
5.3. Weekly sunshine radiation forecast

Figure 8 shows the weekly sunshine radiation forecast, which is a typical summer week. From Figure 8, the forecasting value of SMRBFN is close to the actual values. It can be seen that SMRBFN has the capability to follow the spikes as shown in the 80th ~ 120th hour of this week. Similarly, Figure 9 shows the weekly sunshine radiation forecast, which is a typical winter week. From Figures 8 and 9, it can also be seen that the strength and toughness of SMRBFN is better than that of RBFN, and it is easier to catch up with the peculiarity of sunshine radiation.

Table 3 shows the weekly forecasting errors of SMRBFN and RBFN for comparison in summer and winter. Although the executed time is more than that of RBFN, MAPE and RMSE of SMRBFN are smaller. It is also shown that SMRBFN has better accuracy than RBFN.
Table 3. The weekly forecasting errors of SMRBFN and RBFN.

<table>
<thead>
<tr>
<th></th>
<th>Method</th>
<th>MAPE (%)</th>
<th>RMSE (kW)</th>
<th>Executed time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td>RBFN</td>
<td>11.5419</td>
<td>0.4218</td>
<td>0.017069</td>
</tr>
<tr>
<td></td>
<td>SMRBFN</td>
<td>9.3159</td>
<td>0.3504</td>
<td>0.018953</td>
</tr>
<tr>
<td>Winter</td>
<td>RBFN</td>
<td>13.4771</td>
<td>0.2754</td>
<td>0.015339</td>
</tr>
<tr>
<td></td>
<td>SMRBFN</td>
<td>9.0019</td>
<td>0.2230</td>
<td>0.019794</td>
</tr>
</tbody>
</table>

6. Conclusions

In this paper, the concept of data selection is used to pre-process the historical data of the meteorological database and to select more meaningful data from the database at the same time period, which can reduce the number of modeling data and shorten the calculation time of modeling and prediction. Integration of RBFN and SMC, a SMRBFN is used to forecast the sunshine radiation for finding the output of solar power. SMC is a new toughness algorithm, which is used to adjust the parameters in RBFN training stage to improve the forecasting ability, and a good performance with a close spike tracking capability can be seen. The results of the cases demonstrate that SMRBFN is robust, efficient and accurate. It proved that SMRBFN has the capability to produce better results for forecasting sunshine radiation.

SMRBFN integrated the RBFN and SMC to forecast load based on summer and winter. Data selection algorithm was used to choose meaningful data. It is very important that the proposed algorithm can avoid problems of data insufficiency, bad data or data volatility and can reduce sizes of the training data set. From this study, we can see that more data needs more execution time and could be less efficient. This is opposed to the idea that “more data means more accuracy.” Meaningful data is the key to a better prediction. With the proposed algorithm, we can minimize the data storage, shorten the preprocessing needs, reduce the network size and get better results.

Acknowledgments

We would like to thank the Ministry of Science and Technology, Taiwan for financial support. (Grant Ns. MOST 111-2221-E-230-002).

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

Ming-Tang Tsai assisted the performance of project, model the theory, and prepared the manuscript as the corresponding author. Chih-Jung Huang contributed materials tools, the experiments, and conducted simulations. All authors discussed the simulation results and approved the publication.
References


© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)