A hybrid model with dual channel feature processing for short-term photovoltaic power prediction

Jing Yao¹, Yuanhao Chen¹, Siyuan Yang¹, Yuhui Wang¹, Tao Li¹, Bo Zhu¹, Guanyin Cheng¹ and Xueqin Liu¹*

¹ School of Science, Chongqing University of Technology, Chongqing, 400054, China
*
Corresponding author’s e-mail: xqliu@cqut.edu.cn

Abstract. Adverse effects of random fluctuations and intermittent characteristics of solar irradiance usually hamper the proper operation of the photovoltaic power grid. It is therefore desirable to improve the accuracy of photovoltaic (PV) power prediction. In this work, PV forecasting is realized through a Bayesian optimized model which combines the long short-term memory and radial basis function neural network (BOA-LSTM-RBF). The hybrid model presents a dual channel feature processing by extracting the historical data of PV generation via long-short-term memory network (LSTM) and extracting the forecasted weather conditions via radial basis function neural network (RBF). Then the number of hidden layer neurons and the training batch size are simultaneously optimized by the Bayesian optimization algorithm (BOA). The testing results of three stations demonstrate that, compared with other available models, the RMSE values of BOA-LSTM-RBF model decreased by 2% ~ 17%, which has striking advantages in prediction precision and generalizability. More interestingly, high-precision PV power forecasting can be achieved even under dramatic weather changes.

1. Introduction

With accelerating global energy resources depletion, it is essential and even imperative to develop alternative renewable energy sources [1]. Over the past decades, photovoltaic (PV) power generation that can convert solar irradiation into electric energy has received the attention of the whole world and thus yielded large-scale grid-connected PV power plants [2]. However, the variability and uncertainty of PV power generation arising from the negative influence of weather and environment pose great challenges to the reliable operation of the power grid. It is therefore of practical significance to promote the short-term prediction accuracy of PV power for economical dispatch of network power system and decision-making on the energy market [3].

Going through the literature, one can find that various approaches have been developed and recorded for PV solar power forecasting. Generally speaking, these approaches can be divided mainly into three classes: (i) physical method; (ii) persistence method; and (iii) statistical method. Machine learning approach based on artificial neural network (ANN) has strong ability of nonlinear transformation to achieve PV power output data. When compared with the traditional prediction models, ANN forecasting model exhibits obvious advantages [4], whereas, the features contained in the data may not be completely extracted by shallow networks. With the rapid development of artificial intelligence (AI) technology, various deep neural network models are incorporated into time series prediction of PV output. Gao et al. proposed a day-ahead PV power forecasting model based on long short-term memory (LSTM), which improves the accuracy of power prediction in comparison with back propagation (BP) learning algorithm [5]. Also, different RBF models were established for
PV power prediction under different weather conditions and further tested in actual PV stations [6]. However, forecast accuracy of single model often has limited success. A lot of hybrid models are developed for PV prediction due to their outstanding performance. Wang et al. proposed a PV prediction method that combines (Convolutional Neural Networks) CNN with LSTM and verified the placement order of CNN and LSTM [7]. Wang F et al. forecasted the output power of PV systems by long-short-term memory recurrent neural network (LSTM-RNN), whereas with only historical power as input in this model, the influence of weather features on PV power was ignored [8].

As well known, it is because of the influence of weather features such as wind speed, air temperature, and relative humidity that PV power has the characteristics of instability and intermittence. The weather features at the target time play an important role in the prediction, which can be gained by the weather forecast [9]. In this work, a hybrid dual channel model based on advantages of both LSTM and RBF is proposed for PV forecasting, in which LSTM can capture deep features among time-dependent series, RBF has strong ability to map relationship between PV power and discrete forecasted weather information at target time, predicted power of this model can respond quickly to the weather changes.

2. Methodology

2.1 Long short-term memory
LSTM is more suitable for long-term dependent tasks, which is an improved RNN [10]. LSTM is composed of a forget gate, an input gate, and an output gate. The forget gate decides whether the information should be discarded or retained. The input gate determines how many input parameters at the current moment are saved into the unit state. The output gate is used to determine which value is delivered to the next hidden state and output.

2.2 Radial basis functions neural network
The RBF is a three-layer feed forward neural network, which includes input layer, hidden layer, and output layer [11]. The input layer only serves to transmit information of raw data. The hidden layer realizes to map non-linearly of the input information, whose neuron kernel function is the Gaussian kernel function. Information from the hidden layer is entered into the output layer for linear weighting to obtain the output result of the entire neural network. The general expressions are as follows in equation (1), equation (2):

\[ f(x) = w_0 + \sum_{i=1}^{n} w_i \phi(x) \]  

(1)

\[ \phi(x) = \exp(-\|x - c_i\|^2 / \beta_i^2) \]  

(2)

here the symbol \(\|\|\) denotes the Euclidean norm, \(w_i(i = 1,2,\ldots,n)\) is the \(i\)th weight parameter, \(c_i(i = 1,2,\ldots,n)\) is the \(i\)th center of radial basis function, \(\beta_i\) is the \(i\)th “width” parameter.

2.3 Proposed neural network model
In this article, PV forecasting is achieved through the Bayesian optimization algorithm (BOA) model which combines LSTM and RBF (BOA-LSTM-RBF). As shown in Figure 1, in our proposed model, data are divided into the historical data and forecasted weather data at target time, which are input respectively into the LSTM and RBF channels of the dual channel feature processing model. LSTM is good at processing time series, the hidden layer of RBF network can map the input of low-dimensional space to a high-dimensional space, and find a surface through Gaussian function that can fit the training data best to achieve effectively nonlinear feature extraction Then the different features from the two channels are incorporated into the feature merge layer in the way of one-dimensional splicing. The hybrid model adopts BOA to optimize the number of neurons of hidden layers, training batch size until the optimal predicted value is obtained. The flow chart of BOA is shown in Figure 2.
channel model can take the advantages of each other, capture hidden correlations, and enhance the accuracy of prediction significantly.

![Figure 1: Schematic illustration of LSTM-RBF model](image)

**2.4 Evaluation metrics**

Below, some metrics are considered to evaluate the performance of the proposed model, which includes mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). These quantities of interest are defined in equation (3)-(5):

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \\
MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \\
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}
\]

Where \(y_i\) and \(\hat{y}_i\) stands respectively for the \(i\)th real and predicted PV power, and \(N\) is the number of \(y_i\). The smaller the values of MAE, MSE, and RMSE, the better the prediction accuracy becomes.

**3. Results and discussion**

**3.1 Data description**

The data used in the present work are available from DC competition [12], which include three stations. The data are relative values, which involve historical PV power output and forecasted weather features: solar irradiance, air temperature, relative humidity, wind speed, air pressure, and wind direction. Data were divided into training set and testing set. The first 80% of the samples are used for training the models and the rest of 20% are used for testing.
3.2 Comparative analysis

For giving prominence to the excellent performance of the proposed BOA-LSTM-RBF model, we compare its errors with those of available LSTM model, as well as LSTM optimized by BOA (BOA-LSTM) model and LSTM-RBF hybrid model. There is no optimization algorithm in LSTM-RBF model. Table 1 summarizes the metrics of different forecasting models mentioned above. The forecasting performances of all of hybrid models such as BOA-LSTM, LSTM-RBF, and BOA-LSTM-RBF are superior to the isolated LSTM model. In more detail, it can be found that the BOA-LSTM-RBF model achieves the minimal error values among all models, the MAE, MSE, and RMSE of which for station1 decrease down to 0.0237, 0.0024, and 0.0494, respectively. This is in sharp contrast to the single LSTM model in which the MAE, MSE, and RMSE reaches 0.0410, 0.0036, and 0.0597, respectively. Similar results for station2 and station3 show the generalizability of the proposed model.

|           | LSTM   | BOA-LSTM | LSTM-RBF | BOA-LSTM-RBF |
|-----------|--------|----------|----------|--------------|
| Station1  |        |          |          |              |
| MAE       | 0.0410 | 0.0331   | 0.0240   | 0.0237       |
| MSE       | 0.0036 | 0.0028   | 0.0026   | 0.0024       |
| RMSE      | 0.0597 | 0.0532   | 0.0507   | 0.0494       |
| Station2  |        |          |          |              |
| MAE       | 0.0399 | 0.0274   | 0.0234   | 0.0208       |
| MSE       | 0.0030 | 0.0024   | 0.0023   | 0.0021       |
| RMSE      | 0.0554 | 0.0493   | 0.0481   | 0.0465       |
| Station3  |        |          |          |              |
| MAE       | 0.0393 | 0.0261   | 0.0240   | 0.0207       |
| MSE       | 0.0027 | 0.002    | 0.002    | 0.0019       |
| RMSE      | 0.0526 | 0.0457   | 0.0456   | 0.0438       |

Yang D regarded RMSE as an excellent accuracy measure [13]. As can be seen from the Table 1, comparing BOA-LSTM-RBF model with LSTM model, the RMSE values for three stations are reduced by 17.252%, 16.06% and 16.7%, respectively. The RMSE values of BOA-LSTM-RBF model for three stations are reduced 7.14%, 5.67% and 4.15% than BOA-LSTM model, respectively. The RMSE values of BOA-LSTM-RBF model for three stations are reduced 2.56%, 3.32% and 3.94% after hyper-parameters were optimized by BOA. The improvement is obvious of the proposed model probably due to the involvement of dual channel feature processing.

The prediction effect of station1 is better than the other two stations; therefore, the prediction results of station1 are analyzed in more depth. Figure 3 shows the comparative results between predicted PV data and the real power data by using different models. The diagonal solid line represents a perfect prediction, while the squares indicate cases of prediction in all testing process. The closer the square is to the line, the more accurate PV power prediction is. The predicted data using all models shows good correlation with the real data, the distribution of predicted values by the proposed model with relative to real values is closer to the line of prefect prediction than the others, which confirm the robustness of BOA-LSTM-RBF model.

Data under different weather conditions are randomly selected to further corroborate the exceptional performance of BOA-LSTM-RBF model. As can be seen from Figure 4 (a), all of four models exhibit satisfactory predictive ability in sunny day. However, Figure 4 (b) shows that forecasting results of the LSTM-RBF and BOA-LSTM-RBF models are obviously superior to those of
the LSTM and BOA-LSTM models when PV power values are located around the peak in wet days. In particular, predicted power of BOA-LSTM-RBF model indicates a remarkable degree of consistency with real power, the proposed model is suitable for PV power prediction under multi weather conditions.

Figure 3: Scatter plots of real and predicted values for all models

Figure 4: Comparison of forecasting results of different models with the real power on a day of steady power change on a (a) sunny and (b) wet weather day

Figure 5: Comparison of forecasting results of different models on a day of variable weather
PV power prediction is a huge challenge when the weather changes dramatically, it is necessary to explore the prediction effect in this case. Figure.5 shows the forecasting results of different models of a day with variable weather. The predicted curve of BOA-LSTM-RBF model is more overlapped with the real power, evidencing its superiority over the other models. The proposed model can perceive the trend and respond to the change of PV power quickly, this is because forecasted weather features as a separate input into the dual channel model.

4. Conclusions
To achieve high-precision PV forecasting, a hybrid dual channel feature processing model is proposed, which is based on LSTM, RBF, and BOA, namely BOA-LSTM-RBF network model. This dual channel feature processing model can not only retain the more original features, but also take into account the influence of forecasted weather features on the PV output power. It can be found that the BOA-LSTM-RBF model achieves the minimal error values among all available models, the MAE, MSE, and RMSE for station1 decreases down to 0.0237, 0.0024, and 0.0494, the MAE, MSE, and RMSE for station2 decreases down to 0.0208, 0.0021, and 0.0465, Similarly, the MAE, MSE, and RMSE for station3 decreases down to 0.0207, 0.0019 and 0.0438, respectively. Compared with LSTM single model, RMSE value of BOA-LSTM-RBF model for three stations was reduced by 17.252%, 16.06% and 16.7%. The result of multiple stations shows that the proposed model has advantages in prediction accuracy and generalizability. We verify that the proposed models have obvious advantages in predicting PV power under multi weather conditions.

The BOA-LSTM-RBF model proposed here can obtain satisfactory results in the short-term PV prediction. This research provides a support for PV power forecasting and has practical significance for the design and operation of actual PV power stations.

Acknowledgement
This research was supported by the Graduate Student Innovation Program of Chongqing University of Technology (Grant No. clgycx 20202107).

References
[1] Li, L.L., Wen, S.Y., Tseng ML, CS, Wang. (2019) Renewable energy prediction: A novel short-term prediction model of photovoltaic output power. J. Clean Prod, 228: 359–375.
[2] Li, P., Zhou, K., Lu, X., S, Yang. (2020) A hybrid deep learning model for short-term PV power forecasting. Appl. Energy, 259: 1-11.
[3] Sobri, S., Kooli-Kamali, S., Rahim, N. A. (2018) Solar photovoltaic generation forecasting methods: A review. Energy Conv. Manag, 156: 459–497.
[4] Pedro, H.T., Coimbra, CF. (2012) Assessment of forecasting techniques for solar power production with no exogenous inputs. Sol. Energy, 86: 2017–2028.
[5] Gao, M., Li, J., Hong, F., Long, D. (2019) Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy, 187: 1-12.
[6] Huang, C.M., Chen, S.J., Yang, S.P., Kuo, C. J. (2015) One-day-ahead hourly forecasting for photovoltaic power generation using an intelligent method with weather-based forecasting models. IET Generation, Transmission & Distribution, 9: 1874–1882.
[7] Wang, K., Qi, X., Liu, H. (2019) Photovoltaic power forecasting based LSTM-Convolutional Network. Energy, 189: 1-11.
[8] Wang, F., Xuan, Z., Zhen, Z. (2020) A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. Energy Conv. Manag, 212: 112766.
[9] Huang, X., Zhang, C., Li, Q. (2020) A Comparison of Hour-Ahead Solar Irradiance Forecasting Models Based on LSTM Network. Math. Probl. Eng, 1-15.
[10] Li, P., Zhou, K., Lu, X., Yang, S. (2020) A hybrid deep learning model for short-term PV power forecasting. Appl. Energy, 259: 1-11.
[11] Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L., Yang, S. (2008) RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. IEEE Trans. Power Syst., 23(3): 853-858.

[12] DC Competition platform (2019). https://challenge.datacastle.cn/v3/cmptDetail.html?id=309

[13] Yang, D. Alessandrini, S., Antonanzas, J. (2020) Verification of deterministic solar forecasts. Sol. Energy, 210: 20-37.