Available online at www.sciencedirect.com
SciVerse ScienceDirect
Energy Procedia 12 (2011) 488 – 494
Energy Procedia

ICSGCE 2011: 27–30 September 2011, Chengdu, China

Solar Irradiance Short-Term Prediction Model Based on BP Neural Network

Zhe Wang[a], Fei Wang[a*], Shi Su[b]

[a] *School of Electrical and Elec tronic Engineering, North China Electric Power University, Baoding, China*
[b] *Yunnan Electric Power Test & Research Institute (Group) co., Ltd Electric Power Research Institute, Kunming, China*

Abstract

A short-term solar irradiance prediction model is established based on BP neural network and time series. Firstly, several different network structures of the solar irradiance prediction model are established based on BP neural network. Secondly, the most suitable neural network is gained by the comparison of different network structures and cross-validation. Lastly, the chosen model is trained through setting the suitable network parameters. This model can avoid over-fitting and help to gain a more accurate solar irradiance prediction model. The result of simulation indicates that the model can be effectively used for short-term solar irradiance predicting

Keywords: Solar Irradiance, BP Neural Network, Time Series, Short-term Prediction

1. Introduction

With high-speed development of global economy and the increase of environment consciousness of human beings, clean energy has caught worldwide attention. As a clean renewable energy, solar energy has been used in many fields such as constructions and photovoltaic systems. The lack of solar radiation data has limited the widespread using of solar energy [1]. People have been researching in modeling solar radiation to make the best use of solar energy [2].

There are two main methods to model solar radiation: One is indirect prediction method, which uses sunshine duration, diffused radiation, temperature, and relative humidity as the input data of the model. Japan has been researching in solar radiation modeling since 1970s [3]. The diverging separate model proposed by Udagawa gives the relation of solar altitude and irradiance with less error and simple form. Although the model adopts less observation data and lacks of time representation, it lays the foundation of

* Corresponding author. Tel.: +86-312-7522257.
E-mail address: wangfei.ncepu@gmail.com.

doi:10.1016/j.egypro.2011.10.065

future research. In recent years many scholars put forward kinds of solar radiation prediction models based on neural network [4]-[10]. Another method is direct prediction model, which predicts the value of the next moment form the former solar irradiance data. A growing number of facilities are working on collecting data from solar radiation [11]-[12]. For example, the solar radiation monitors FY-3A onboard the Fengyun-3 satellite can collect the solar irradiance data. These facilities provide guarantee to the realization of the direct prediction method. Presently the short-term prediction methods at home and abroad include the traditional prediction method, such as regression method, trend extrapolation, time series, the artificial neural network prediction method based on artificial intelligence, etc. The traditional prediction method has limitations when dealing with nonlinear problems [13]-[14]. Neural network provides a new approach for the nonlinear problems. It is valuable to apply neural network on short-term prediction.

In this paper, measured solar radiation data from National Renewable Energy Laboratory (NREL) are employed for the solar irradiance modeling. Back propagation (BP) neural network and time series are used to establish the short-term solar irradiance prediction model. Network structures affect the solar irradiance prediction precision. The most suitable network structure is selected by comparison and cross-validation, then the chosen model is trained through setting the suitable network parameters, which can avoid over fitting and help to gain a more accurate solar irradiance prediction model. Simulation shows that the model can predict short-term solar irradiance accurately.

2. Data Pretreatment

The data are obtained from NREL Solar Radiation Research Laboratory (SRRL) at 39.74° N, 105.18° W, and 1828.8 meters above mean sea level. The interval of samples is 1 hour. The data pretreatment includes the exclusion of abnormal data and normalization.

2.1. Abnormal data processing

1) Exclude abnormal data, such as the negative data, and the data of which the radiation value is greater than the theoretical extraterrestrial radiation value.

2) Use the linear interpolation method to fix the missing data.

3) If the relative humidity data is greater than 100, make it to be 100.

4) If the calculated value of the sunshine duration is longer than the theoretical value, replace it with the theoretical value.

2.2. Normalization

When neural network applies to solar irradiance prediction, the orders of magnitude differ because of the different variable units. So the data should be normalized before they are trained. In this paper the normalization function “mapminmax” is adopted. The normalization interval is [0.1, 0.9], not the [0, 1] as usual, this ensures convergence of the network. While the random classification function “dividevec” is adopted to draw a certain proportion sample data at random to form the training input, cross-validation is used to set the network topology, which ensures that the sample is representative, reliable, and uniformly distributed. And this method can effectively avoid over fitting.

3. Solar Irradiance Prediction Model

BP neural network is a backward propagation of errors and multilayered feed-forward neural network [15]. It contains one input layer, one or more hidden layers, and one output layer. The topology of BP neural network with double hidden layers is shown in Fig. 1.

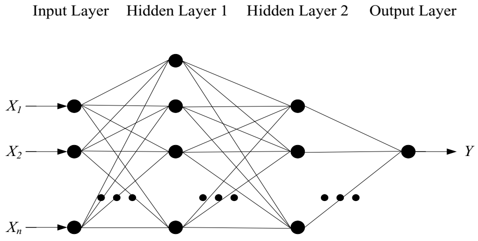

Fig. 1. BP neural network topology.

BP neural network design should consider input and out parameter, network layer, node number, transfer function, and training function, etc.

3.1. Input and output

Traditional solar irradiance modeling usually chooses meteorology factors such as diffused radiation, temperature, relative humidity and time as the input parameters. The models always ignore some factors to simplify the calculation. However, it makes the models less accurate, and the traditional models can not deal with the place which lacks of meteorology factors, so the direct prediction model based on its own solar irradiance data is called to solve the problem. The model is designed to predict irradiance value with the previous and current irradiance data. Both the input and output data are the solar irradiance data with sampling interval of 1hour, and the number of neuron in output layer is 1. In order to get a precise model, the number of neurons in input layer and the number of hidden layers should be confirmed, and then through setting the network parameters, a most suitable BP neuron network is obtained.

3.2. Network layers

It has been proved that if a BP network with one hidden layer has enough number of neurons, it can realize any nonlinear mapping. But if the sample number is large, the network with one hidden layer cannot reach an accurate function and the calculation efficiency declines intensively [16]. In order to save training time and get an accurate model, the network with two hidden layers is considered.

3.3. Node number

In this paper, the former 24 hours solar irradiance data are used to predict the next value with sampling interval of 1 hour. So the number of neurons in input and output layer is 24 and 1 respectively. The number of neurons in the single hidden layer can be calculated by (1).

$$m = \sqrt{n+l} + a \tag{1}$$

where n is the number of neurons in the input layer, l is the number of neurons in the output layer, α is the constant, and $1< \alpha<10$.

The number of neurons in the single hidden layer network is 6-15. The number of neurons in a double hidden layers network can be designed by (1) primarily, then the network is debugged over. The most suitable model is chosen according to statistic indicators, it include coefficient of determination (R^2), root mean square error (RMSE), mean percentage error (MPE), and mean average bias error (MABE).

Cross-validation are employed to choose the stable model, the samples are divided into five groups. Any four groups are chosen as training sample, and the remaining group as validation sample.

The comparison is shown in Table 1.

Table 1. Statistic indicators of different hidden layers and number of neurons

neuron number	R^2	RMSE	MPE	MABE
7	0.9755	0.0362	1.4203	5.2390
11	0.9704	0.0366	1.5162	0.0163
14	0.9775	0.0344	0.8794	0.0155
17	0.9796	0.0339	0.8469	0.0155
18	0.9810	0.0332	1.0695	0.0153
20	0.9782	0.0338	1.1064	0.0158
25	0.9772	0.0343	0.6341	0.0159
27	0.9776	0.0350	1.4533	0.0163
15,12	0.9812	0.0396	1.0858	0.0151
17,16	0.9799	0.0324	0.9030	0.0144
18,6	0.9807	0.0342	0.5772	0.0142
18,13	0.9912	0.0331	0.7720	0.0138
19,15	0.9878	0.0355	0.7613	0.0142
19,22	0.9764	0.0360	1.1967	0.0158
20,17	0.9796	0.0338	1.2178	0.0145

From Table 1 some conclusions can be made as follows:

1) As the number of neurons in the single hidden layer increases, the four indicators R^2, RMSE, MPE, and MABE perform best when the number of neurons is 18. The number is a little larger than that of the experience formula, which is 6-15. As the number of neurons increases more, the indicators become worse.

2) The double hidden layers model is more accurate than the single hidden layer model.

According to the comparison, the most suitable neural network is the double hidden layers model with number of neurons of 18 and 13.

3.4. Transfer function and training function

The transfer functions are “tansig” and “logsig” respectively. The output layer adopts the linear transfer function “purelin”.

Levenberg-Marquardt (LM) algorithm, which is a combination of Grade method and Gauss-Newton method, is chosen to avoid the time-consuming one-dimension searching. It has the local convergence of Gauss-Newton method as well as the whole characteristic of grade method [17].

4. Simulation and Result Analysis

A four-year solar radiation data measured by NREL form 2006 to 2010 is selected to establish the prediction models.

Two direct prediction models are established based on BP neural network and time series with double hidden layers. The input and output data of the two direct models are the solar irradiance data with sampling interval of 1 hour.

The difference of the two models is that, the input data of Model 1 are collected from 6:00 am to 8:00 pm. The input data of Model 2 are the whole 24 hours data, that is, the input data of two models are 15 and 24 respectively.

The statistic indicators of the two models are shown in Table 2.

Table 2. Statistic indicators of two models

Model	R^2	RMSE	MPE	MABE

Model 1	0.9645	0.0450	0.5904	0.0257
Model 2	0.9912	0.0331	0.7220	0.0138

By comparison, the statistic indicators of Model 2 are better than Model 1. R^2 of Model 2 reaches 0.9912. So the former 24 solar irradiance data are used to predict the next value, and the number of neurons in the input layer is 24.

After the construction and parameters are fixed, the sample data should be classified reasonably, they are divided into two groups at random: training and validation. The group proportions are 70% and 30% respectively. Go along by turns, by this way, the over-fitting problem can be avoided efficiently. The simulation of the training error is shown in Fig. 2.

Mean square error (MSE) of the training sample becomes small gradually as iterative steps increase. But the MSE curve climbs a little after the 24th iterative step, which indicates that from the 25th step on, the network occurs over fitting, so stop training and output the validation error at the 24th iterative step. The statistic indicators of the training error are shown in Table 3.

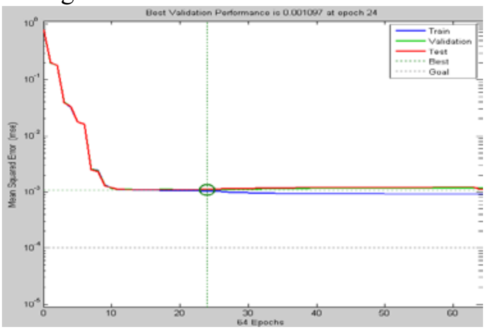

Fig. 2. Training curve.

Table 3. Statistic indicators of training error

step	MSE	RMSE	MPE	MABE	R^2
5	0.005837	0.0764	6.2576	0.0601	0.9008
10	0.001163	0.0341	1.0122	0.0142	0.9886
15	0.001471	0.0338	0.9626	0.0140	0.9887
20	0.001142	0.0337	0.8820	0.0144	0.9905
24	0.001097	0.0331	0.7220	0.0138	0.9912
30	0.001176	0.0343	1.5849	0.0150	0.9881
35	0.001239	0.0352	0.9286	0.0148	0.9867

The indicators of the 24th iterative step are the best. R^2 reaches 0.9912. It is an efficient way to get the suitable network.

The first three-day-data of the validation sample are chosen to test the accuracy and reliability of the network. The scatter diagram of the measured value and predicted value is shown in Fig. 3.

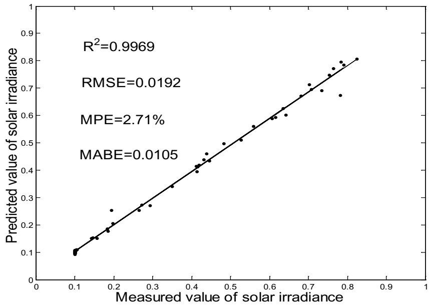

Fig. 3. Scatter diagram of solar irradiance predicted value and measured value

X is the measured value, and y is the predicted value. The expression is y=0.0048+0.9660x , R^2 reaches 0.9969, which is much closer to 1, and other statistic indicators are shown in Fig. 1. By comparing the predicted value and the measured data, a relatively ideal model is obtained.

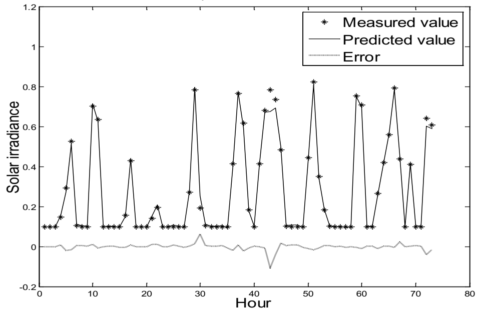

Fig. 4. The curve of solar irradiance predicted value and measured value.

The curve of irradiance predicted value and measured data is shown in Fig. 4. It can be seen that the BP neural network model can reflect the solar irradiance change trend precisely with little error, and the extreme points are in good agreement. The accuracy and applicability are fully validated.

5. Conclusion

In this paper, several different network structures for the short-term solar irradiance model are established based on BP neural network and time series. A relatively ideal model is obtained though comparison of different network structures by cross-validation and adjusting the network parameters.

This paper collects solar irradiance data from NREL. According to the simulation results, the number of neurons in the double hidden layers is 18 and 13, the transfer functions are “tansig” and “logsig” respectively. R^2 is 0.9912, MPE is 0.7720 and MABE is 0.0138. The predicted value and the measured value are in good agreement. The proposed model can be used in solar irradiance short-term prediction, and more research is needed for other place or situation.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (11ZG09) and the Key Science and Technology Project of China Southern Power Grid (2011ZD24).

References

[1] R. Sun, L. Liang, and L. Yang, "Estimation of daily solar radiation in China based on weather station data," *Meteorological and Environmental Sciences,* vol. 30, no. 1, pp. 24-27 ,Feb. 2007.

[2] J. H. Yang, L. Ge, and X. L. Jiang, "Establishment of solar radiation data-base on variously tilted surface in China," Renewable Enercy, vol. 5, pp. 14-16, May. 2004.

[3] A. G. Song, "General situation and analysis of model of solar radiation in Japan," *Journal of Capital Normal University (Natural Science Edition)*, vol. 18(sup), pp. 68-72, Sept. 1997.

[4] J. Zhou, Y. Z. Wu, and Z. J. Ma, "Solar radiation estimation using artificial neural networks," *Acta Energiae Solaris Sinica,* vol. 26, no. 4, pp. 509-512, Aug. 2005.

[5] R. Shafiqur, M. Mohamed, "Artificial neural network estimation of global solar radiation using air temperature and relative humidity," *Energy Policy,* vol. 36, pp. 571-576, Feb. 2008.

[6] C. L. Joseph and Y. Liu, "Solar radiation modelling using ANNs for different climates in China," *Energy Conversion and Management,* Vol. 49, pp. 1080-1090, May. 2008.

[7] S. H. Cao and J. C. CAO, "Study of chaos optimization neural networks for the forecast of hourly total solar irradiation," *Acta Energiae* Solaris *Sinica*, vol. 27, no. 2, pp. 164-169, Feb. 2006.

[8] B. Zhou, C. Z. Chen, and N. Song, "Application on forecast of day-by-day solar irradiance based on blind source separation neural network," *Acta Energiae Solaris Sinica,* vol. 28, no. 9, pp. 1008-1011, Sept. 2007.

[9] Y. N. Jiang, X. J. Cheng, and J. Liu, "Models for estimating monthly mean daily diffuse radiation," *Renewable Energy Resources,* vol. 26, no. 5, pp. 8-12, Oct. 2008.

[10] S. Adnan, A. Erol, andO. Mehmet, "Forecasting based on neural network approach of solar potential in Turkey," *Renewable Energy,* vol. 30, pp. 1075-1090, Jun. 2005.

[11] D. J. Yang, W. Fang, H. Qiu, and J. Qi, "On-orbit data calibration of FY-3A solar irradiance monitor," *Chinese Journal of Luminescence,* vol. 31, no. 5, pp. 676-681, Oct. 2010.

[12] W. Fang, B. X. Yu, and H. H. Yao, etc, "Solar irradiance absolute radiometer and international comparison," *Acta Optica Sinica*, vol. 23, no. 1, pp. 112-116, Jan. 2003.

[13] J. F. Lou, J. L. Zheng, and S. Y. Sun, "Prediction algorithm based on the combination of chaos and neural network and it' s applications," *Chinese Journal of Radio Science*, vol. 14, no. 2, pp. 172-177, Jun. 1999.

[14] G. Y. Zhang, "Short-term forecasting method of chaotic time series based on NNT," *Computer Engineering,* vol. 28, no. 11, pp. 197-198, Nov. 2002.

[15] L. Q. Han, Artificial *neural network tutorial.* Beijing: Beijing University of Posts and Telecommunications Press, 2006, pp. 22-30.

[16] Z. Q. Chen, W. H. Liu, and Y. S. Ou, "Cross-talk resistant adaptive neural network algorithm for speech enhancement," *Chinese Journal of Scientific Instrument,* vol. 29, no. 4, pp. 617-620, Apr. 2008.

[17] H. Y. Yang, H. Ye, and G. Z. Wang, "Applications of chaos theory to load forecasting in power system," *Relay,* vol. 33, no. 16, pp. 26-30, Apr. 2005.