Short-term Global Horizontal Irradiance Prediction Based on Deep Echo State Network

Gaohong Ma, Jun Li

1Automation & Electrical Engineering School, Lanzhou Jiaotong University, China
e-mail: 201509410@stu.lzjtu.edu.cn

Abstract. The prediction of global horizontal irradiance have a great impact on the stability and economic benefits of photovoltaic (PV) power generation. In this paper, we adopt the method of Deep Echo State Network (DESN) to predict the global horizontal irradiance in different areas one hour in advance. Under the same conditions, the results show that DESN are better than BP, SVM, ESN methods in the prediction accuracy. Experiments show that the proposed models show superior ability in predicting solar irradiance and have great application potential in power grid integration.

1. Introduction
In recent years, solar energy and other renewable energy are rising rapidly, the popularization rate is higher and higher, and the energy storage system is expanding. The prediction of solar irradiance is becoming more and more important. At present, Autoregressive integrated moving average neural network(ARIMA)[1], support vector machine and k-nearest neighbors[2] have been successfully applied to solar irradiance prediction. Through these examples, it is proved that these prediction methods can model solar irradiance, so as to overcome some nonlinear problems in the prediction process. However, for the atmospheric data with multiple variable characteristics, the prediction accuracy will be significantly reduced in the short-term prediction of irradiance.

Echo state network(ESN)[3] is a method based on reservoir. Although it is simple but effective. Similar to the structure of ordinary neural network, echo state network is also composed of input, hidden and output layers. The only difference from traditional neural networks is the hidden layer structure. The hidden layer structure of echo state network is composed of sparse connected neuron matrix, which needs to meet the echo state attribute, which mapping the input information to the high-dimensional state space. After initialization and iterative update, the hidden layer neurons are no longer updated. Only the weight matrix of output needs to be trained. Therefore, the ESN avoids a large number of calculations and local minimums problems of the traditional RNN based on the gradient descent algorithm. In addition, the mapping capabilities and nonlinear processing capabilities of ESN were stronger than classic RNNs and has been used in many fields, including control[4-5] and time series forecasting[6-7]. To further improve the performance of ESN, Claudio Gallicchio[8] combined deep learning theory and proposed the Deep Echo State Network (DESN).

2. The ESN algorithm
The single layer of ESN is shown in Figure 1. The state equation of reserve pool neurons of ESN network is given:
\[ x(t) = (1-a) x(t-1) + a f(W_{in} t(t) + \hat{W} x(t-1)) \]  

(1)

Where \( f(*) \) is the activation function of state neurons, Tanh function is generally used in echo state network. \( \hat{W} \in \mathbb{R}^{n \times n} \) is the weight state matrix obtained by random initialization, \( W_{in} \) is also the weight input matrix obtained by random initialization, and the dimension is depends on input datasets. While \( a \in [0,1] \) represents the leaky parameter, which select the input to better represent the feature. The larger the value of \( a \) is, the faster the response speed will be. The network weight must meet the Echo State Property (ESP) during initialization[9]:

\[ \rho(1-a) I + a \hat{W} < 1 \]  

(2)

**Fig. 1** The diagram of Echo State Network structure

### 3. DESN

The schematic diagram of DESN is shown in the figure below. The status transfer function of DESN is shown in Eq. (3), to facilitate the representation, the bias terms in the formula are omitted.

\[ x^{(l)}(t) = (1-a^{(l)}) x^{(l)}(t-1) + a^{(l)} f(W_{in}^{(l)} t^{(l)}(t) + \hat{W}^{(l)} x^{(l)}(t-1)) \]  

(3)

\[ t^{(l)}(t) = \begin{cases} u(t) & \text{if } l = 1 \\ X^{(l-1)}(t) & \text{if } l > 1 \end{cases} \]  

(4)

It is noted that the update formula of DESN is different from ESN is the number of hidden layers deeper. Where \( l \) represents the layers, \( t^{(l)}(t) \) is the input of each layer, which is given by Eq. (4). \( x^{(l)}(t) \)
represents the state of the reservoir at time \( t \) of the L-th layer. \( W^{(l)} \) is the weight input matrix obtained by random initialization of the L-th layer. \( \hat{W} \in R^{N_s \times N_s} \) is the weight state matrix obtained by random initialization of the L-th layer. And \( d^{(l)} \in [0,1] \) represents the leaky parameter, which select the input to better represent the feature. The larger the value of \( a \) is, the faster the response speed will be.

After initializing the state of the reservoir according to Eq. (3), the state of each layer can be obtained, as shown in Eq. (5):

\[
X^{(N_s)} = \begin{bmatrix}
    x_{11}, & x_{12}, & \cdots & x_{1T} \\
    x_{21}, & x_{22}, & \cdots & x_{2T} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{N_s1}, & x_{N_s2}, & \cdots & x_{N_sT}
\end{bmatrix}_{N_s \times T}
\]  

(5)

Where, \( N_s \) is the size of the reservoir, \( T \) is the training samples, and \( N_L \) is the total number of layers. It can be seen from Fig.1 that the hidden layer of DESN can be formed by stacking the reservoir together and adding the input sample data, as shown in Eq. (6):

\[
X = \left[ (X^{(1)})^T, (X^{(2)})^T, \cdots, (X^{(N_s)})^T, u(t) \right]_{L \times (N_s \times N_s + N_p)}
\]  

(6)

After training, the prediction result of the DESN is shown in Eq. (7):

\[
\hat{y} = W_{out}X
\]  

(7)

Generally, the output weights of the DESN are trained by linear regression with regularized least squares, as shown in Eq. (8):

\[
W_{out} = \left( X^TX + \lambda I \right)^{-1}X^Ty
\]  

(8)

4. Experiments and results

This study adopts the data from three sites collected by the US National Ocean and Atmospheric Administration, which are Bondville, Boulder, and Desert Rock, respectively. For each dataset, the data from 2010 and 2011 are hired as the training part and the data of 2012 as a test set. Since the radiation at night is negligible, it is removed from the datasets. For missing data, use adjacent data to impute, and finally, normalize the data. The external meteorological variables are used as input variables, as shown in Table 1. And the training goal is global horizontal irradiance.

| Tab. 1 The exogeneous meteorological variable | Unit |
|---------------------------------------------|------|
| DGS                                        | W/m² |
| Upwelling GS                               | W/m² |
| DNS                                        | W/m² |
| DDS                                        | W/m² |
| DTI                                        | W/m² |
| DICT                                       | K    |
| DIDT                                       | K    |
| UTI                                        | W/m² |
| UICT                                       | K    |
| UIDT                                       | K    |
| GUVB                                       | mW/m²|
| PAR                                        | W/m² |
| NS                                         | W/m² |
The model evaluation indexes selected in this paper are common root mean square error (RMSE), mean absolute error (MAE) and standard deviation (Std) of MAE. The parameters of the network are obtained through experiment and correction, and select the initial parameters as follows: the deep of layer is determined as 5, each reservoir size is determined as 100, $\text{scale}_e$ is set as 0.1, $M$ is set as 100, the spectral radius about echo status attribute is 0.83, and the randomly selected state matrix $\hat{W}$ obey the uniform distribution on [-1, 1].

Table 2-4 show the comparison of the performance index of 1 hour ahead solar irradiance prediction in Bondville, Boulder, and Desert rock regions based on different methods. In order to verify the effectiveness of the DESN method proposed in this article, persistence is selected as the reference method. It can get the result from the table that, no matter in which region, the DESN method are obviously better than the benchmark methods and other training methods in predicting one hour ahead solar irradiance, which reflects the generalization ability of the methods.

### Tab. 2 Performance comparison of 1 hour ahead of GHI models in Bondville in 2012

| METHOD | RMSE  | MAE  | Std  |
|--------|-------|------|------|
| Persistence | 102.3952 | 5.3547 | 34.4366 |
| BP     | 29.0179 | 19.9477 | 36.7707 |
| SVM    | 33.7995 | 22.8483 | 38.1587 |
| ESN    | 26.3907 | 16.8752 | 30.7234 |
| DESN   | 21.9442 | 12.7196 | 24.5907 |

### Tab. 3 Performance comparison of 1 hour ahead of GHI models in Boulder in 2012

| METHOD | RMSE  | MAE  | Std  |
|--------|-------|------|------|
| Persistence | 110.3893 | 5.7836 | 37.1555 |
| BP     | 35.0808 | 22.3475 | 41.8951 |
| SVM    | 45.0067 | 28.9371 | 54.1835 |
| ESN    | 31.3414 | 15.8824 | 34.8436 |
| DESN   | 26.9642 | 12.1328 | 30.1319 |

### Tab. 4 Performance comparison of 1 hour ahead of GHI models in Desert rock in 2012

| METHOD | RMSE  | MAE  | Std  |
|--------|-------|------|------|
| Persistence | 116.0186 | 6.0743 | 38.9665 |
| BP     | 41.3173 | 27.7139 | 50.7221 |
| SVM    | 47.3797 | 31.1605 | 56.0262 |
| ESN    | 31.7516 | 16.9544 | 35.462 |
| DESN   | 27.6212 | 13.2868 | 31.0704 |

To show the prediction effect more intuitively, the prediction chart of 50 data samples of solar irradiance in the Bondville area is shown in Fig. 3, and the chart of 1000 samples points error is shown in Fig. 4. From the chart, it can be concluded that the accuracy of DESN is better, and the fluctuation range of error is smaller and more stable, which can more accurately predict the solar irradiance.
5. Conclusion
Aiming at the problem of short-term GHI prediction, this paper proposes the algorithms of DESN, which is used to map the data to high-dimensional states, linear regression is used to train the output of DESN. Compared with the training of recurrent neural networks such as RNN, echo state network does not need to train the neurons of the whole neural network, just train the output structure, that is the weights from the reservoirs to the output, which can greatly reduce the training difficulty.

To verify the accuracy of the method, DESN is applied to GHI prediction in different regions and time scales, and RMSE, MAE, and Std are used to assessment the accuracy of the DESN methods. From the table and chart above, the results show that DESN are better than BP, SVM, ESN methods in the prediction accuracy. Experiments show that the proposed models show superior ability in predicting solar irradiance and have great application potential in power grid integration.

References
[1] Conejo A, Plazas M, Espinola R, and Molina A 2005 Day-ahead electricity price forecasting using the wavelet transform and ARIMA models IEEE Trans. Power Syst. vol. 20 pp 1035–1042
[2] Mellit A, Pavan A M A 2010 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid connected PV plant at Trieste Italy Solar Energy 84(5) p 807e21
[3] Jaeger H, Haas H 2004 Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication Science 304(5667) pp 78-80
[4] Shi G, Liu D, Wei Q 2017 Echo state network-based Q-learning method for optimal battery control of offices combined with renewable energy Iet Control Theory & Applications 11(99) pp 915-922
[5] Bo, Ying-Chun, Xin Z 2018 Online Adaptive Dynamic Programming Based on Echo State Networks for Dissolved Oxygen Control Applied Soft Computing 62 pp 830-839
[6] Chouikhi N, Ammar B, Rokbani N, et al. 2017 PSO-based analysis of Echo State Network parameters for time series forecasting Applied Soft Computing 55 pp 211-225
[7] Shen L, Chen J, Zeng Z, et al. 2017 A novel echo state network for multivariate and nonlinear time series prediction Applied Soft Computing 62 pp 524-535
[8] Gallicchio C, Micheli A, Pedrelli L 2017 Deep Reservoir Computing: A Critical Experimental Analysis Neurocomputing 268(11) pp 87-99
[9] Claudio G, Alessio M, Luca P 2018 Design of deep echo state networks Neural Networks 108 pp 33-44