Short-term Photovoltaic Power Prediction based on Sparse Representation Method

DanYang Zheng, Jun Li *

Big Data School, School of Automation & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou Gansu, China

*675189222@qq.com

Abstract. In the research of solar power prediction, providing accurate prediction data in real time is one of the most effective means to enhance the capacity of wind power acceptance and improve the power reliability and economy. The existing prediction models based on statistical methods are often unavoidable in data preprocessing and model training stage, and their adaptive ability needs to be improved. Considering that the sparse coding method does not require model training, and has the characteristics of high solving efficiency and strong self-adaptability, an online solar energy prediction model using sparse coding is proposed. Firstly, the historical time series data is composed of input-output pairs with delay, and the dictionary is respectively constructed in atomic form. Then, the sparse weight is calculated for the delay input data vector to be predicted, and the corresponding predicted output is obtained by borrowing the dictionary. Taking the actual solar power data of Alberta, Canada as sample, the simulation was carried out in MATLAB. The simulation results show that the model can accurately predict the solar power and improve the effectiveness and practicability of the prediction.

Keywords: The sparse coding, Output power prediction, Feature extraction

1. Introduction

In recent years, renewable energy, especially solar energy, has attracted more and more attention. However, due to the randomness and intermittency of solar energy resources and the uncertainty influence caused by the high penetration of solar energy, it is easy to produce "adverse effect" in the operation of power system. Solar power prediction can avoid the influence of power fluctuation caused by the increase of unit installed capacity, which is one of the most important research directions in the energy field.[1-3].

Solar power prediction methods are usually divided into physical methods, statistical methods and so on. Traditional statistical methods are the prediction methods of linear regression models such as Persistence model and auto-regressive integral moving average (ARIMA) model. Such models are low in complexity and have certain applications in engineering, but their prediction accuracy is low. More advanced statistical methods, such as neural network, support vector machine and various combination prediction methods, etc. For example, literature proposed ultra-short-term wind power prediction based on wavelet threshold noise reduction and BP neural network. Literature proposed an online prediction...
algorithm based on OS-ELM, which can meet the online application requirements in terms of computing time by taking advantage of the ELM modeling learning speed. Literature [4] established a prediction model by using ADA-Boost algorithm to integrate support vector machine. These methods all need to do the corresponding data preprocessing, and the model needs to go through a relatively complex training stage.

Sparse Representation (SR) [5], as an advanced machine learning method, has been widely used in image denoising, pattern recognition and other fields. The SR method is composed of dictionary learning and sparse coding. In 2015, literature [6] proposed a dictionary-based sparse coding prediction method, which has been successfully applied in benchmark time series.

In view of the advantages of sparse representation algorithm in feature extraction and modeling, this paper proposes a new method for short-term prediction of photovoltaic power based on dictionary learning K-SVD-OMP, combined with the KELM method to form a global prediction model. The proposed method is applied to the photovoltaic benchmark power prediction example provided by the 2014 Global Energy Forecasting Competition (GEFCOM2014). Under the same conditions, the prediction results are compared with prediction methods such as SVM and KELM and sparse representation feature extraction modeling methods based on non-dictionary learning to evaluate the effectiveness of the K-SVD-OMP method.

2. Principles of photovoltaic power online prediction

In actual conditions, the parameters are often affected by many factors such as location, weather, and unit aging, and the power value cannot be accurately calculated. According to the time series forecasting theory, many uncertain information of the system are implicit in the historical solar power data. Through historical power data to learn an effective prediction model, the prediction of future power can be realized.

Suppose a wind power sequence \(\{y(t)\}\) of length \(N\), first construct a data set \(\{x(t); y(t + h)\}\) \(i=1\), then the prediction model to be established is as follows

\[
y(t + h) = f(x(t)) = f[y(t), y(t - 1), \ldots, y(t - m + 1)]
\]

(1)

After the model is established, at time \(t\), the input vector \(x'(t)\) is constructed from the current power \(y(t)\) and the previous \(m - 1\) historical power values, and the corresponding predicted output \(^y\) can be obtained by substituting it into the prediction model \(f(t)\), meanwhile, dynamically update the prediction model \(f(\cdot)\) to ensure real-time follow-up of the model.

3. Online prediction of sparse coding based on adaptive dictionary

3.1. Sparse coding prediction model

In sparse representation, dictionary learning converts samples into appropriate sparse representations, simplifying the learning task and reducing the complexity of the model. Therefore, different dictionaries can be selected for different types of data. Such as: Discrete Cosine Transform (DCT) dictionary, data dictionary, structured dictionary, etc. In order to better address the sparse representation of different signals, this article directly selects a better K-means singular value decomposition dictionary.

The essence of the Sparse Coding method is to solve the sparse solution of the vector based on an over-complete dictionary. The commonly used initial dictionaries include Fourier dictionary, wavelet dictionary, etc. This article directly selects historical training input-output data pairs constructed from time series as dictionaries, without the need for model training.

Establish dictionary \(\Psi \in \mathbb{R}^{m \times N}\) as the whole input vector of training data set, where \(N\) is the number of training data set, \(\Psi\) is: \(x(t) = [y(t), y(t - 1), \ldots, y(t - m + 1)]^T\), \(\Psi\) once determined, it does not change. The other dictionary \(\Phi \in \mathbb{R}^{m \times N}\) consists of the corresponding target \(y(t)(t + h)\) as the atom.

3.2. Sparse weight solution algorithm

Sparse representation is a linear combination of coefficients of an over-complete basis function to represent the original signal. The sparse representation of the signal vector \(x \in \mathbb{R}^m\) is:
For formula (1), it can be obtained by solving the l0-norm optimization problem of coefficient vector $\alpha \in \mathbb{R}$, namely:

$$\min_{\alpha} \| \alpha \|_0 \quad \text{s.t.} \quad \| \Psi \alpha - x \|_2^2 \leq \varepsilon$$

Limiting the number of non-zero coefficients in $\alpha$ and formulating (2) as an M-sparse optimization problem, we can get:

$$\min_{\alpha} \| \Psi \alpha - x \|_2^2 \quad \text{s.t.} \quad \| \alpha \|_0 \leq M$$

Using Lagrangian multipliers, formula (2) can be optimized as:

$$\min_{\alpha} \frac{1}{2} \| \Psi \alpha - x \|_2^2 + \lambda \| \alpha \|_0$$

For the non-stationary solar power time series, once the prediction model is fixed, it will be difficult to track the drastic changes of signals and produce a large prediction error. In this paper, the sparse coding method based on dictionary directly imports data into dictionary without model training, which greatly reduces the complexity of online updating. Based on this consideration, with the continuous entry of test set data, the existing dictionaries can be updated adaptically, and the dictionary capacity can be constrained to avoid memory overflow, so as to carry out online prediction of time series. This paper presents the following dictionary adaptive updating strategy:

1) Replace Old (RO) strategy;
2) Replace Neighborhood (RNr) strategy;
3) Keep nearby atoms (KR) strategy.

4. Experiment and example analysis

The experiment used an actual solar data set from a solar power plant in Canada in 2011, with a sampling interval of 10min. The forecast period is the data between 20 o’clock on December 21 and 24 o’clock on December 25, 2010. Samples were collected every 30min. Embedding dimension $m=6$, Step length $h=1$, Delay constant $\tau=1$.

Figure 1 shows the prediction results of EN-SPARSE -2 at an out-of-synchronization length of half an hour in advance. Each graph contains two subgraphs, which are the comparison of predicted value and true value respectively, and the relative error comparison graph of each point. From Figure 1, it is clear that the sparse coding method in this paper has a very good prediction effect.
Fig. 1 Graph and error graph of EN-Sparse -2 method for power prediction half an hour in advance

Table 1 lists the numerical comparison results of 10 representative methods and SVM methods. Table 1 shows that THE MAPE and RMSE values of EN-SPARSE method are the lowest under the premise of a certain average sparse degree, with high prediction accuracy.

| Prediction method   | MAPE    | RMSE    | Mean Sparsity |
|---------------------|---------|---------|---------------|
| SVM                 | 9.1613  | 42.982  | 0%            |
| Basic sparse-1      | 8.4446  | 42.3227 | 95.3%         |
| Basic sparse-2      | 8.4433  | 42.3184 | 95.3%         |
| Basic sparse-3      | 13.243  | 64.8179 | 98.8%         |
| EN-sparse-1         | 7.9304  | 40.947  | 89.9%         |
| EN-sparse-2         | 7.9289  | 40.9327 | 89.9%         |
| Basic sparse-1(RO)  | 8.2926  | 40.2987 | 95.3%         |
| Basic sparse-2(KR)  | 7.7674  | 39.9872 | 95.3%         |
| EN-sparse-1(KR)     | 7.4511  | 38.9949 | 89.7%         |
| EN-sparse-2(KR)     | 7.4515  | 38.9996 | 89.7%         |

5. Conclusion
Sparse coding based on dictionary is proposed for short-term on-line photovoltaic power prediction. Basic sparse coding method is similar to THAT of SVM. In the premise of a certain mean sparse degree, en-Sparse method is better than the above method. With the addition of online dictionary updating strategy, the prediction accuracy of sparse coding method is improved to some extent on the original basis, and the prediction effect is obvious, which can meet the engineering requirements at the present stage. The model presented in this paper is suitable for online photovoltaic power prediction and has good application potential. Further research includes a prediction method that considers the combination of dictionary learning algorithm and sparse coding.

References
[1] Tuohy A, Zack J and Haupt S E et al 2015 Solar forecasting: methods, challenges, and performance IEEE Power and Energy Magazine 13(6) pp50-59
[2] Tao Hong, Pierre Pinson and Shu Fan et al 2016 Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond J International Journal of Forecasting 32(3) pp896-913
[3] Ammar R B, Oualha 2017 A Photovoltaic power prediction using recurrent neural networks J
Engineering Applications of Artificial Intelligence (63) pp54-68

[4] Liu Xindong, Chen Huanyuan, She Caiqi 2011 Ultra-short term wind electrical power rate prediction based on wavelet threshold noise reduction and BP neural network J World science and technology research and development 33(6) pp1006-1010

[5] Wang Yan, et al. 2014 Ultra-short-term Wind power prediction based on OS-ELM and Bootstrap methods J Power system automation 38 (6) pp14-19

[6] Hu Qian, Chen Hongkun, SUN Zhida 2016 Integrated Learning Prediction Model of Wind Farm Short-term Wind Speed J Journal of power systems and automation 28 (1) p1-4

[7] Xu Zhengyuan, Zhang Cheng, Wei Sui 2013 Research and Improvement of Sparse Representation Face Recognition Algorithm J Computer simulation 30 (6) pp405-408

[8] erzijbergh V R A, Heijnen P. W and de Roode S R 2015 Alexander Los, Jonker H J J Improved model output statistics of numerical weather prediction based irradiance forecasts for solar power applications Solar Energy 118 pp634-645

[9] oyant V C, Notton G and Kalogirou S et al 2017 Machine learning methods for solar radiation forecasting: A review", Renewable Energy 105 pp569-582

[10] Al-Dahidi S, O. Ayadi, M. Alrbai and J. Adeeb 2019 Ensemble Approach of Optimized Artificial Neural Networks for Solar Photovoltaic Power Prediction in IEEE Access 7 pp81741-81758

[11] Louazni M, N Omran 2020 A Local Training Strategy-Based Artificial Neural Network for Predicting the Power Production of Solar Photovoltaic Systems in IEEE Access 8 pp150262-150281

[12] Huang C and Kuo P 2019 Multiple-Input Deep Convolutional Neural Network Model for Short-Term Photovoltaic Power Forecasting in IEEE Access 7 pp74822-74834

[13] Sun Y et al 2015 Short-term prediction model of module temperature for photovoltaic power forecasting based on support vector machine International Conference on Renewable Power Generation (RPG 2015) Beijing pp1-6

[14] A. Karabiber and Ö. F. Alçın 2019 Short Term PV Power Estimation by means of Extreme Learning Machine and Support Vector Machine 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG) Istanbul Turkey pp 41-44

[15] Natarajan, Y J, Subramaniamb Nachimuthu D 2016 A new PCA pretreatment method Journal of Process Control 20 pp181-194