Energy Conversion and Management 171 (2018) 787–806

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Variational mode decomposition based low rank robust kernel extreme learning machine for solar irradiation forecasting

Irani Majumder, P.K. Dash*, Ranjeeta Bisoi

Siksha O Anusandhan Deemed to be University, Bhubaneswar, India

A R T I C L E I N F O

Keywords:
Solar irradiation prediction
Variational mode decomposition
Empirical mode decomposition
Low rank robust kernel extreme learning machine
Morlet Wavelet Kernel

A B S T R A C T

In this paper a new hybrid method has been implemented by combining Variational Mode Decomposition (VMD) and a new low rank robust kernel based Extreme Learning Machine (RKELM) for solar irradiation forecasting. This hybrid model presents an efficient and effective short term solar irradiation prediction approach using the historical solar irradiation data. The original non-stationary time series data is decomposed into various modes using VMD approach. The proposed VMD-RMWK (VMD based reduced Morlet Wavelet Kernel extreme learning machine) method is used to predict the solar irradiation of an experimental 1 MW solar power plant in Odisha, India. Different time intervals of 15 min, 1 h and 1 day ahead in different weather conditions are considered for forecasting purpose. The VMD technique decomposes the original nonlinear irradiation into a set of Variational Mode Functions (VMFs), and the extracted VMFs are used to train the kernel based robust ELM. Comparison with empirical mode decomposition (EMD) based low rank kernel is also presented in this paper. As a new contribution to the previously performed literature survey this paper presents a more accurate solar irradiation prediction paradigm for distinctive weather conditions, and different time intervals varying from very short duration of 15 min to one day ahead. Also to improve the reliability of the KELM and to make it robust under noisy conditions and the presence of outliers in the data, a weight loss matrix has been derived using a non-parametric kernel density estimation method and incorporated in the new formulation (RKELM). A typical solar power experimental solar power station in India has been taken for detailed study showing clearly the accuracy and robustness of the proposed approach. For cross validation of the proposed model, solar irradiation data from a solar power plant located in the state of Florida has been implemented.

1. Introduction

Due to increased demand in renewable energy, photovoltaic generation systems have gained more importance. PV power penetration into the grid led to precious solar irradiation prediction for stable functioning of the power plant. Effective solar power prediction is necessary mostly for high energy integration [1,2]. The intermittent nature of the solar irradiation gives way to variable problems like stability, power quality issues and voltage fluctuations [3]. The solar output is not easily predicted and differs based on variable weather conditions and location of the solar plant as this affects the irradiation.

Previously many researchers have applied different forecasting methods for solar irradiation or power forecasting. These methods can be broadly categorised into two different types namely linear and nonlinear forecasting techniques, predominantly implemented linear methods are ARMA, ARIMA and physical model like (NWP) [4–10]. The numerical weather prediction method (NWP) previously used requires more detailed information of the input data accompanied with greater computational complexity. The ARMA and ARIMA models are very popular for solar irradiation forecasting but suffer from inaccuracies due to inability to handle highly nonlinear and fluctuating data. On the other hand the latter method includes different artificial intelligence and machine learning techniques. This category includes fuzzy logic system [11–13], artificial neural network (ANN) [14–17], support vector machine (SVM) [18–22], etc. Although the ANN based methods for irradiation forecasting are known for their widespread implementations, but they suffer from problems like generalization, over fitting, local minima, and lower convergence speed.

The ANN techniques proved to be unstable in case of small variation in time series data and selection of different attributes, contributing to inaccurate solar irradiation prediction. A number of hybrid methods were introduced to predict the solar irradiation time series data which is a combination of two or more methods. The hybrid models imparted better prediction accuracy and hence are considered to be superior when compared to other methods. The different hybrid models initially used for solar irradiation prediction include hybrid ANN like ARMA-

* Corresponding author.
E-mail address: pkdash.india@gmail.com (P.K. Dash).

https://doi.org/10.1016/j.enconman.2018.06.021
Received 1 March 2018; Received in revised form 20 May 2018; Accepted 6 June 2018
0196-8904/

ANN [10,18], etc. In order to overcome the limitations of ANN based models, many researchers have proposed a new learning machine algorithm known as the Extreme Learning Machine (ELM) [23–28]. This is a single layer feed forward neural network (SLFNs) which structurally shows a resemblance with random vector functional link (RVFL) network and where the input weights are randomly selected. The randomly chosen weights cause variations in the output for different trial runs and hence making the system non-robust. A new incremental ELM (I-ELM) was proposed by Huang et al. [26], and in this case the accuracy of the training samples was improved. But the major drawback of ELM and its variants are their dependence on the correct choice of neurons in the hidden layer and the right activation function, which are still problems for larger input data set. In order to mitigate the problem of hidden layer selection, kernel functions can be used. The kernel function improves the stability of the prediction system when applied in ELM technique and is termed as Kernel based Extreme Learning Machine (KELM) [29,30]. Different types of kernel functions namely Gaussian kernel, Polynomial Kernel, Sigmoid kernel and Wavelet Kernel are implemented for prediction purpose. The kernel based formulation provides a consolidated framework and generalized model for forecasting time series data. In spite of the afore mentioned advantages, the kernel technique suffers from a disadvantage of larger training time while dealing with large data set. A recent literature [31] shows that the training time and computational overhead can be substantially decreased by reducing the size of the kernel matrix. In this paper the row rank kernel method is described for reducing the execution time without affecting the accuracy of the system to a greater extent. Further the KELM is made robust against the presence of noise and outliers in the fluctuating solar irradiation data during weather changes by incorporating a variable weight for the residual error using a non-parametric kernel density technique [35] and this model will be known as RKELM.

The nonlinearity of the time series data can be handled by decomposition techniques. Empirical mode decomposition (EMD) technique is used to decompose the data into intrinsic mode functions and a residue [32]. The basic disadvantage of this technique is that it lacks mathematical foundation. In recent time, a decomposition technique known as the Variational mode decomposition (VMD) [33,34] has been implemented as a better alternative to the previously used EMD technique. Unlike the EMD model VMD processes an exact mathematical model and sensitive to both noise and sampling. VMD is a non-recursive variational model that explores for a number of modes and there central frequencies in a band limited manner which reconstructs the original signal in the least square sense. As compared to EMD VMD has a superior denoising property and ability to separate tones of similar frequencies. In this paper the VMD is used to decompose the original solar irradiation data in order to obtain the fundamental variational mode functions (VMF) which combine together to form the original data set. The VMFs act as the predictor for KELM to forecast the future data. Although various VMD based hybrid models have been implemented for wind power and price forecasting, they are still not implemented for solar irradiation or power forecasting. The main contribution of this paper is that it presents a hybrid technique implementing VMD and different low rank robust kernel extreme learning machines (RKELM) namely VMD-RMWK, VMD-RMHWK, VMD-RGK, VMD-RPK and VMD-RSK) for solar irradiation prediction under different weather conditions like sunny, rainy and foggy at various time horizon (15 min, 1 h and 1 day). The RKELM reduces the effects of the outliers and hence increasing the reliability of the samples, and this process decreases the execution time without affecting the accuracy of prediction to a greater extent. The decomposition technique (VMD) divides the main nonlinear data into different modes. This reconstructs the original signal with the help of least square technique which removes the nonlinearity of the original signal. Further to validate the superiority of the proposed prediction model VMD-RMWK, its performance is compared with the basic model (VMD-MWK) and decomposition based RKELM (EMD-RMWK), and also other RKELM variants.

The rest of the paper is organised as follows: Section 2 describes the RKELM (robust kernel extreme learning machine) model along with its reduced or low rank version. In Section 3 VMD and EMD techniques are described for the decomposition of solar irradiation time series data into different modes which are used as inputs to the forecasting kernel models. In Section 4 the detailed numerical experimentation results are outlined for both VMD based RKELM and EMD based RKELM prediction models in different atmospheric conditions followed by conclusion in Section 5.

2. Robust kernel based extreme learning (RKELM)

The output function of the basic ELM with L hidden nodes can be represented by

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i(x) = h(x)\beta \tag{1}$$

where the output vector $\beta = [\beta_1, \beta_2, \ldots\ldots\ldots\ldots, \beta_L]$ between the L hidden neurons and the output neuron and the ELM feature mapping function$h(x) = [h_1(x), h_2(x), \ldots\ldots\ldots\ldots, h_L(x)]$; the number of input samples$x = [x_1, x_2, \ldots\ldots\ldots\ldots, x_N]$, and N represents the number of patterns. The initial weights between the input layer consisting m inputs for each pattern and hidden layer with L neurons need not be tuned and the activation function of the hidden neurons could comprise almost all nonlinear piecewise continuous functions. Thus using the tanh function as the activation function

$$h_i(x) = \tanh(w_{i0} + w_{i1}x_{i1} + w_{i2}x_{i2} + \ldots\ldots\ldots\ldots + w_{im}x_{im}) \tag{2}$$

Thus the hidden layer randomized matrix is written as

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h(x_1)\cdots h_L(x_1) \\ \vdots \;\vdots\; \vdots \\ h(x_N)\cdots h_L(x_N) \end{bmatrix} \tag{3}$$

and the target vector is given by

$$T = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix} \tag{4}$$

Eq. (1) is expressed in a matrix form as

$$H\beta = T \tag{5}$$

To solve for β, a constrained optimization problem needs to be solved that overcomes the over fitting problem and provides better generalization ability in comparison to the original ELM. This is similar to the structural risk minimization of the statistical learning theory and is expressed as

$$L_{C_{ELM}} = \frac{1}{2}\,\|\beta\|^2 + \frac{1}{2}C\sum_{i=1}^{N} w_i \xi_i^2 \tag{6}$$

Subject to $h(x_i)\beta = t_i - \xi_i,\ \ i = 1, 2, \ldots\ldots\ldots\ldots, N$ where the error vector ξ is obtained as: $\xi = [\xi_1, \xi_2, \ldots\ldots\ldots\ldots, \xi_N]$, C is the regularization parameter. However, if the data samples are corrupted by noise or outliers, the reliability of the regularized ELM is not robust. Therefore, in Eq. (6) a weighting factor $w_i$on a sample to sample basis is introduced using the nonparametric kernel density estimation [35] to the output error. This provides better reliability of the samples by unevenly weighting each sample, so that the effect of noise or outliers in the data is eliminated. This makes the ELM robust. Using KKT theorem the constrained optimization problem in Eq. (6) is converted to a dual optimization problem as

$$L_{D_{ELM}} = \frac{1}{2}\,\|\beta\|^2 + \frac{1}{2}C\sum_{i=1}^{N} w_i \xi_i^2 - \sum_{i=1}^{N} \alpha_i(h(x_i)\beta - t_i + \xi_i) \tag{7}$$