Energy 129 (2017) 122–137

Contents lists available at ScienceDirect

Energy

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

Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine

Xiaohui Yuan [a], Qingxiong Tan [a], Xiaohui Lei [b,*], Yanbin Yuan [c], Xiaotao Wu [a]

[a] School of Hydropower and Information Engineering, Huazhong University of Science and Technology, 430074, Wuhan, China
[b] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, 100038, Beijing, China
[c] School of Resource and Environmental Engineering, Wuhan University of Technology, 430070, Wuhan, China

ARTICLE INFO

Article history:
Received 29 September 2015
Received in revised form
1 April 2017
Accepted 16 April 2017
Available online 19 April 2017

Keywords:
Wind power prediction
Autoregressive fractionally integrated moving average
Least square support vector machine
Autocorrelation function
Long memory characteristics

ABSTRACT

Precise prediction of wind power can not only conduct wind turbine's operation, but also reduce the impact on power systems when wind energy is injected into the grid. A hybrid autoregressive fractionally integrated moving average and least square support vector machine model is proposed to forecast short-term wind power. The proposed hybrid model takes advantage of the respective superiority of autoregressive fractionally integrated moving average and least square support vector machine. First, the autocorrelation function analysis is used to detect the long memory characteristics of wind power series, and the autoregressive fractionally integrated moving average model is applied to forecast linear component of wind power series. Then the least square support vector machine model is established to forecast nonlinear component of wind power series by making use of wind speed, wind direction and residual error series of the autoregressive fractionally integrated moving average model. Finally, the prediction of wind power is obtained by integrating the prediction results of autoregressive fractionally integrated moving average and least square support vector machine. Compared with other models, the results of two examples demonstrate that the proposed hybrid model has higher accuracy of wind power prediction in terms of three performance indicators.

1. Introduction

Nowadays owing to the intensification of energy crisis and environmental pollution, wind energy is becoming increasingly important in energy industry. The totally-installed world wind power capacity in 2015 increased to 434.9 GW with the newly-installed 63.7 GW. The corresponding global growth rate was 17.2%. Meanwhile the newly-installed wind power capacity was 33.0 GW and the total capacity reached 114.8 GW with the growth rate of 29.0% for China [1]. Electric energy transformed from the renewable and clean wind power will be the priority in the choice of power systems. However, wind power has the characteristics of randomness, intermittence and instability. If the electricity generated by unstable wind power, especially in a large quantity, is injected into the power grid, it will threat the safe operation of the grid. Fortunately, accurate prediction of wind power is a good way to solve this problem [2]. With the help of accurate wind power prediction, operators of power systems will know the quantity of electric power generated by wind energy timely and make a reasonable dispatching plan for other energies to produce a proper amount of electricity. Therefore, the accurate wind power prediction is very important for the safe and economic operation of power systems.

There are mainly two types of wind power prediction technologies, namely physical model and statistical model. Physical model utilizes physical information (pressure, local terrain and temperature et al.) to establish the formula of wind power prediction [3], which is good at long-term prediction. However, this approach has a high computation complexity in solving the model. The statistical model intends to find out the relationship between historical wind power through time series models, artificial neural network (ANN) and support vector machine

* Corresponding author.
E-mail address: leixiaohui_sky@163.com (X. Lei).

http://dx.doi.org/10.1016/j.energy.2017.04.094
0360-5442/

(SVM). Time series model is widely used for short-term wind power prediction, such as auto regressive (AR), auto regressive moving average (ARMA) and auto regressive integrated moving average (ARIMA), which has higher prediction accuracy when the wind power data shows linearity and stationarity [4]. Although time series model establishes the mapping relationship between wind speed and wind power by extracting information in the historical wind speed signals, it has the disadvantages of nonlinear fitting capability weakness, requirement of large historical records and difficulty in modeling nonlinear problems. To overcome the disadvantages, some improved time series models are put forward for wind power prediction. An improved time series method (ITSM) based ARIMA and the wavelet decomposition (WD) is proposed for hourly wind power forecasting [5]. The chaotic time series model has a better prediction performance after the complex wind power series are de-noised by empirical mode decomposition (EMD) [6]. The multi-scale decomposition of the original wind speed, which is indispensable in improving the prediction accuracy, is widely used. Wavelet transform (WT) and EMD are used to decompose the wind speed into multiple series to eliminate irregular fluctuation and improve the prediction accuracy. It illustrates that the selection of a particular base wavelet and scale as well as intrinsic mode function (IMF) may cause a false wave in WT and EMD. A time series approach based Taylor Kriging (TK) model outperforms ARIMA in prediction of wind speed series in terms of mean absolute error (*MAE*) [7]. This is because TK was used to modify the distance formula between points in covariance function, which promotes wind speed series forecasting. The autoregressive fractionally integrated moving average (ARFIMA) model which can well describe time series with long memory characteristics is adopted to perform day-ahead wind speed prediction [8], and the results indicate that the ARFIMA has higher prediction accuracy than ARIMA because of its capability of correcting local trends in wind power. By making use of linear regression equation, linear models are easy to be formulated and can capture the linear relationship of wind power series. Despite the application of linear models in wind power prediction, such models suffer from some weakness, notably their incapability of capturing nonlinear patterns because the linear correlation structures are assumed among time series. Some nonlinear models are applied to forecast wind power so as to overcome the limitations of linear models. A ridgelet neural network trained by differential evolution outperforms persistence method and multivariate ARIMA for wind power forecasting [9]. Compared with other techniques, wavelet neural network (WNN) trained by clonal selection algorithm has better performance for hourly wind forecasting [10]. Intelligent algorithms (differential evolution and Clonal selection algorithm et al.) are employed to optimize the parameters of artificial neural network model for enhancing forecasting precision of wind power. Compared with ARMA and ANN, single multiplicative neuron model based nonlinear filter performs accurate forecasting for hourly wind speed, which utilizes the recursive information of wind speed without requirement of storage of all historical data. An advanced statistical method based artificial intelligence techniques makes use of the previous power measurements and meteorological information to predict hourly wind power [11]. The simulation results show that it can improve prediction accuracy by taking full advantage of meteorological factors and considering the impact of dependent variables on wind power. ANN has been proven as an efficient forecasting method due to its capabilities of self-learning, establishing complex nonlinear relationships and extracting the dependence between input and output data sets, and not requiring explicit mathematical expressions as used in the physical and statistical approaches. Therefore, ANN is a good choice for forecasting wind power. The model based ANN takes wind speed, direction and previous time interval of wind power as input variables for prediction of short-term wind power [12]. Despite good performance of ANN in wind power prediction, there is a tendency to trap into local minima. ANN is both sensitive to parameter selection and time-consuming. Moreover, it is probable that different structures and parameters of ANN can yield different forecast accuracies for the same wind data in terms of various evaluation criteria. The other statistical machine learning tool, namely SVM, which is used effectively for nonlinear regression problems, has been applied in wind forecasting [13]. SVM has higher generalization ability than ANN and overcomes the disadvantage of ANN in wind speed and power prediction. But SVM has complex optimization process and longer training time. To handle this problem, the least square support vector machine (LSSVM) model is proposed by transforming the complicated quadratic programming into linear equations problem. In real-world experiments, LSSVM has been found with good generalization performance and low computational cost. LSSVM optimized by gravitational search algorithm (GSA) shows stronger ability for hourly wind power forecasting than that of back propagation (BP) and SVM [14]. The simulation results demonstrate that the forecasting accuracy of the LSSVM based GSA is improved by selecting the best kernel function and optimizing the model parameters with the GSA. The forecasting performance of the LSSVM together with the WD for wind energy at different time horizons is compared with that of hybrid ANN and WD [15]. A hybrid LSSVM with empirical wavelet transform (EWT) and coupled simulated annealing (CSA) model is put forward for half-hour wind speed prediction in which the model performance outperforms AR, persistence model and CSA-LSSVM [16]. Since the WD's decomposition preprocesses or eliminates the stochastic volatility of wind power series and the LSSVM has the learning generalization ability, combination of LSSVM and WD outperforms other methods for wind power prediction. As the above mentioned models have flexible nonlinear modeling capability based on the feature representation of the data, they are popular for wind power prediction. However, most of the methods only make use of the single prediction model, such as ARIMA, ANN and SVM, and are incapable of accurately forecasting complicated wind power patterns. Therefore, in recent years, some combined forecasting models, whose basic idea is to combine different models and retain the advantages of each model, have been proposed for wind power prediction. An online time-varying wind power prediction system based several alternative models, whose number of predictors and the corresponding weights change over time depending on their actual performances, can improve the forecasting accuracy [17]. A robust combination approach, which utilizes EWT to extract information of wind speed and makes use of Gaussian process regression (GPR) to combine independent predictions generated by various models (ARIMA, extreme learning machine, SVM and LSSVM), produced more accurate prediction of short-term wind speed comparison of a single model [18]. Compared with Gaussian mixture copula model (GMCM) based ARIMA and support vector regression (SVR), hybrid GMCM and localized GPR model with posterior probability as weight value shows stronger capacity in wind speed prediction [19]. This is because the GPR algorithm is adopted to treat the non-stationary wind speed series and the Bayesian inference strategy is applied to evaluate