Energy 93 (2015) 41–56

Contents lists available at ScienceDirect

Energy

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

A robust combination approach for short-term wind speed forecasting and analysis – Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model

Jianzhou Wang [b], Jianming Hu [a,*]

[a] School of Mathematics & Statistics, Lanzhou University, Lanzhou 730000, China
[b] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

ARTICLE INFO

Article history:
Received 27 March 2015
Received in revised form
17 July 2015
Accepted 11 August 2015
Available online 29 September 2015

Keywords:
Gaussian Process Regression
Wind speed forecasting
Empirical Wavelet Transform
Extreme Learning Machine
Support Vector Machine

ABSTRACT

With the increasing importance of wind power as a component of power systems, the problems induced by the stochastic and intermittent nature of wind speed have compelled system operators and researchers to search for more reliable techniques to forecast wind speed. This paper proposes a combination model for probabilistic short-term wind speed forecasting. In this proposed hybrid approach, EWT (Empirical Wavelet Transform) is employed to extract meaningful information from a wind speed series by designing an appropriate wavelet filter bank. The GPR (Gaussian Process Regression) model is utilized to combine independent forecasts generated by various forecasting engines (ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM)) in a nonlinear way rather than the commonly used linear way. The proposed approach provides more probabilistic information for wind speed predictions besides improving the forecasting accuracy for single-value predictions. The effectiveness of the proposed approach is demonstrated with wind speed data from two wind farms in China. The results indicate that the individual forecasting engines do not consistently forecast short-term wind speed for the two sites, and the proposed combination method can generate a more reliable and accurate forecast.

Definition

Multistep ahead forecast: suppose that we are at the time index $h$and are interested in forecasting $\widehat{r}_{h+l}$, where $l \geq 1$. The time index h is called the forecast origin and the positive integer l is the forecast horizon. Let $\widehat{r}_h(l)$ be the forecast of r_{h+l}, we refer to $\widehat{r}_h(l)$ as the l-step ahead forecast of r_t at the forecast origin h. When $l = 1$, we refer to $\widehat{r}_h(1)$ as the **one-step ahead forecast** of r_t at the forecast origin h.

Hyberparameters: In Bayesian statistics, a hyperparameter is a parameter of a prior distribution; the term is used to distinguish them from parameters of the model for the underlying system under analysis.

1. Introduction

Due to increasing energy demand and environmental concerns, wind power, as an environmentally friendly source of renewable energy, has attracted global attention. The rapid development of wind energy gives wind power the potential to support sustainable economic development and environmental protection [1]. However, the stochastic and intermittent nature of wind power complicates the large-scale penetration of wind power into the grid system because this could decrease system reliability and power quality. Wind speed forecasts can effectively reduce the risk to the power system induced by wind-related uncertainties.

Many methods and models have been proposed in recent decades to obtain accurate wind speed predictions, including physical approaches and statistical models. Physical approaches are generally used for long-term wind forecasts and utilize weather data, while statistical models are more commonly used for short-term

* Corresponding author. Tel.: +86 189 93173569; fax: +86 931 8912481.
E-mail address: hujm11@lzu.edu.cn (J. Hu).

http://dx.doi.org/10.1016/j.energy.2015.08.045
0360-5442/

Symbols and notation

EWT

V_n	the nth segment in $[0,\pi]$
ω_n	the nth maxima of Fourier spectrum
$\forall$	randomly assign
$\widehat{\phi}_n(\omega)$	the empirical scaling function
$\widehat{\varphi}_n(\omega)$	the empirical wavelets
γ	the real number belong to [0,1]
$\langle\rangle$	inner product
$f(t)$	the signal
$\overline{\varphi_n(\tau-t)}$	the conjugate of the empirical wavelets
$w_f^{\varepsilon}(n,t)$	The detail coefficients

ARIMA model

$\varphi(B)$	the AR (autoregressive) model
$\theta(B)$	the moving average function
q	the lag order of moving average terms
p	the lag order of autoregressive terms
d	lag order of non-seasonal differences

LSSVM model

W	the coefficient vector
$\Phi(x)$	the mapping function
γ	regulation constant
e_i	error variable
α_i	the Lagrange multipliers
$K(x_i,x)$	kernel function
b	bias term

SVM model

w	the coefficient vector
$\phi(x)$	the mapping function
ξ_i	the slack variable (feasible case)
ξ_i^*	slack variable (non-feasible case)
α	the Lagrange multipliers for ξ_i
α^*	the Lagrange multipliers for ξ_i^*
$K(x,y)$	kernel function
ε	the approximation precision
C	regulation constant
b	bias term

ELM model

β	the coefficient
$h(x)$	the mapping function
t_i	the indicator vector
H	the mapping function matrix
C	constant
ξ_i	training error vector

GPR model

$p(\bullet)$	the probability function
ϕ	parameters of observation model
$m(x)$	the mean function
f	the latent variables
θ	parameters of covariance function
$k(x,x'\vert\theta)$	the covariance function

prediction and are primarily performed through analysis of historical data. NWP (Numerical Weather Prediction) models are representative of the physical methods most often presented in the literature [2,3]. Statistical models mainly include time series techniques (ARMA (Auto-Regressive Moving Average) [4], ARIMA (Autoregressive Integrated Moving Average) [5], FARIMA (fractional ARIMA) [6], exponential smoothing techniques [7] and gray predictors [8]) and statistical learning methods (ANN (Artificial Neural Networks) [9–14], fuzzy logic [15,16], and SVM (Support Vector Machine) [17–19]).

Recent studies have predominantly focused on short-term wind forecasting ranging from minutes to hours because of the importance of these forecasts for power systems. Various attempts have been made to use hybrid methods for short-term wind forecasting. The combined approaches most commonly seen in the literature are data preprocessing-based approaches, parameter-optimization-based approaches and weighting-based approaches.

When employing data preprocessing-based approaches, techniques such as WT (Wavelet Transform) and EMD (Empirical Mode Decomposition) are used as data preprocessors to decompose wind series or eliminate stochastic volatility. The WT method is widely utilized for wind speed series data processing by researchers [17,20–22] due to its adaptive ability of time–frequency analysis. For example, Liu et al. [17] proposed a hybrid model consisting of WT, GA (Genetic Algorithm) and SVM in which WT decomposed the wind speed series into two components for wind speed forecasting. The EMD method, introduced by Huang et al. [23] for the decomposition of non-stationary signals into a series of IMFs (Intrinsic Mode Functions), has also been recently utilized for the purpose of wind data processing [9,19,24–26]. For example, Hu et al. [19] suggested a forecasting approach associated with ensemble EMD and the SVM, with Ensemble EMD used to decompose the wind speed series for the SVM to enhance wind speed prediction precision. Among the parameter-optimization-based approaches, many studies have concentrated on stochastic heuristic optimization algorithms that have the capability of fast convergence to global optima and relatively simple implementation. These stochastic heuristic optimization algorithms include the GA [27,28], DE (Differential Evolution) algorithm [29], PSO (Particle Swarm Optimization) algorithm [30], evolutionary programming [31], MTS (Memory Tabu Search) algorithm [32], and ICA (Imperialist Competitive Algorithm) [33]. The weighting-based approaches presented in the literature [4,33–36] combine several independent forecasting techniques through a weight that determines the relative effectiveness of each model. Most of the independent forecasting techniques involved include time series techniques and statistical learning methods.

The aforementioned models are useful for point or deterministic wind speed or power forecasts. These combined models are mostly parametric models that offer point estimation of wind speed given appropriate inputs. However, due to the stochastic and variable nature of the wind, the accuracy of such forecasts cannot be guaranteed and tends to be fairly low. In addition, the uncertainty in wind energy gives rise to economic risks for wind farm owners, especially in competitive electricity markets [4,5]. In such circumstances, probabilistic forecasting of wind power becomes highly meaningful for both utilities and system operators who manage wind farms. Probabilistic forecasting can effectively reflect the uncertainties associated with prediction results, and help to assess the risk of relying on a forecast [6]. For these reasons, some researchers have examined the interval wind speed or power prediction with respect to uncertainty