Renewable Energy 76 (2015) 637–647

Contents lists available at ScienceDirect

Renewable Energy

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

Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA

Osamah Basheer Shukur, Muhammad Hisyam Lee*

Department of Mathematical Sciences, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

ARTICLE INFO

Article history:
Received 20 August 2014
Accepted 25 November 2014
Available online

Keywords:
Wind speed forecasting
Forecasting accuracy
Hybrid KF-ANN
ARIMA
Kalman filter
ANN

ABSTRACT

The accuracy of wind speed forecasting is important to control, and optimize renewable wind power generation. The nonlinearity in the patterns of wind speed data is the reason of inaccurate wind speed forecasting using a linear autoregressive integrated moving average (ARIMA) model. The inaccurate forecasting of ARIMA model reflects the uncertainty of modelling process. The aim of this study is to improve the accuracy of wind speed forecasting by suggesting a more appropriate approach. An artificial neural network (ANN) and Kalman filter (KF) will be used to handle nonlinearity and uncertainty problems. Based on the ARIMA model, a hybrid KF-ANN model will improve the accuracy of wind speed forecasting. First, the effectiveness of ARIMA will be helped to determine the inputs structure for KF, ANN and their hybrid model. A case study will be carried out using daily wind speed data from Iraq and Malaysia. The hybrid KF-ANN model was the most adequate and provided the most accurate forecasts. In conclusion, the hybrid KF-ANN model will result in better wind speed forecasting accuracy than its separate components, while the KF model and ANN separately will be provide acceptable forecasts compared to ARIMA model that will provide ineffectual wind speed forecasts.

1. Introduction

Frequent, extreme wind speeds and the nonlinear nature of wind speed data makes forecasting a complex process. Some authors have proposed using ARIMA models, a classical statistical approach, to forecast wind speeds. Finding the appropriate wind speed ARIMA model can be accomplished by following the methodology proposed by Box–Jenkins. Benth and Benth [1] proposed an ARIMA model for estimating and forecasting wind speeds for three different wind farms in New York State. Shi, Qu and Zeng [2] adopted a simplified ARIMA model for direct and indirect short-term forecasting methods then compared the performances of both approaches using the wind speed and power production data from an offshore 2-MW wind turbine. Zhu and Genton [3] reviewed statistical short-term wind speed forecasting models, including autoregressive models and traditional time series approaches, used in wind power developments to determine which model provided the most accurate forecasts. AR, ARIMA, or seasonal ARIMA models have been used for comparing and for determining KF state equation structures and ANN inputs structure such as those proposed by Refs. [4–9].

The nonlinear pattern of wind speed data may be one reason for the inaccuracy of ARIMA forecasting, which is a linear time series model [5]. An ANN was used to handle the nonlinear nature of wind speed data. Cadenas and Rivera [4] presented a comparison of ARIMA and ANN approaches for wind speed forecasting using seven years of wind speed data. Six years of data was used for training and one year of data was reserved for validation. Li and Shi [10] compared one hour ahead forecasts for hourly wind speeds in North Dakota using three different types of artificial neural networks. They used an autocorrelation function (ACF) and partial autocorrelation function (PACF) to determine the ANN input variables. Pourmousavi-Kani and Ardehali [11] used ANN to develop very short-term wind speed forecasts after combining the ANN with a Markov chain to create an ANN–MC hybrid model. Assareh, Behrang, Ghalambaz, Noghrehabadi and Ghanbarzadeh [12] forecasted wind speeds using twelve years of data from the Manjil station. The final year of data was used for testing and the first eleven years were used for training. ANN was proposed as a way to represent the relationship among wind speeds, as an output, and other meteorological time series data, and to accurately forecast

* Corresponding author. Tel.: +60 7 5534236; fax: +60 7 556 6162.
E-mail addresses: assomeeh@yahoo.com (O.B. Shukur), mhl@utm.my (M.H. Lee).

http://dx.doi.org/10.1016/j.renene.2014.11.084
0960-1481/

wind speed. Bilgili and Sahin [13] used ANN to forecast daily, weekly, and monthly wind speeds using data from four different measuring stations in the Aegean and Marmara regions of Turkey. They obtained successful forecasting results. Peng, Liu and Yang [14] suggested an individual ANN and hybrid strategy based on physical and statistical approaches for short term wind power forecasting. The individual ANN approach resulted in highly accurate forecasts.

In this paper, an ANN was constructed based on an autoregressive order to simulate the ANN structure. An autoregressive order can be determined by observing PACF. Khashei and Bijari [15] and Khashei and Bijari [16] explained the use of an autoregressive order (p) for determining the inputs structure of ANN. Guo, Zhao, Haiyan and Wang [9] proposed many methods for wind speed forecasting. One of these methods was a feed-forward neural network whose input variables were determined using a partial autocorrelation function that was dependent on the autoregressive order. Liu, Tian and Li [6] confirmed that the performance of the hybrid method in terms of its wind speed predictions was consistently better than that of its components. They proposed two types of hybrid methods. One method has been called a new ARIMA-ANN hybrid method, while it contains using ANN based on AR(p) model and it performed well.

Although an ARIMA model is the perfect statistical model for forecasting, it leads to inaccurate results for wind speed forecasting. The KF model can be used for meteorological purposes, such as wind speed forecasting. To obtain the best initial parameters for the KF, an ARIMA model will be used to create the structure of the KF model that will be regarded as the best model for handling the stochastic uncertainty and improve wind speed forecasting. An ARIMA model will be used with the KF model to construct the structure of the state-space equation. This model is also called a hybrid ARIMA-KF.

Malmberg, Holst and Holst [17] used the KF model based on an AR to model and forecast the large scale component of bounded areas of near-surface ocean wind speeds. Galanis, Louka, Katsafados, Pytharoulis and Kallos [18] proposed implementing nonlinear polynomial functions in classical linear Kalman filter algorithms as a new methodology that would improve regional weather forecasts. Louka, Galanis, Siebert, Kariniotakis, Katsafados, Pytharoulis and Kallos [19] applied the KF model as a post-processing method for numerical wind speed forecasting and employed two limited area atmospheric models with different horizontal resolution to improve wind speed forecasts. Cassola and Burlando [20] proposed a mixed approach based on the use of a NWP model coupled with a statistical model based on the KF model to generate a way to forecast wind speed and wind power data sets collected from two anemometric stations located in the eastern Liguria. Liu, Tian and Li [6] proposed two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed forecasting, and compared their performance. They combined an ARIMA with ANN and the KF model in order to create the structure of an ANN model and initialize the state equation for a KF. Zhu and Genton [3] suggested using traditional time series statistical models of wind speed forecasting, including a KF method, to handle uncertainty in the system, noise, and observation noise. Tatinati and Veluvolu [7] proposed many approaches for short term wind speed forecasting. One of these approaches was a hybrid model that combined the KF model with an AR model to improve forecasting accuracy.

In this paper, a hybrid KF-ANN model was proposed based on an ARIMA model to further improve the forecasting accuracy of wind speed. Artificial neural network (ANN) and Kalman filter were useful for handling nonlinearity and stochastic uncertainty problems associated with wind speed data. Many recent studies have also combined the KF model to handle stochastic uncertainty, with another approach, such as support vector machines which handle the nonlinearity of wind speed. Tatinati and Veluvolu [7] proposed several approaches for short term wind speed forecasting such as a the KF model based on AR, a least squares version of support vector machine (SVM), an empirical mode of decomposition (EMD), and their hybrid model for average wind speed and the direction in Beloit for the period 2003–2004. Chen and Yu [8] integrated unscented Kalman filter (UKF) with support vector regression (SVR) based on a state-space model. The hybrid SVR–UKF approach was employed firstly to handle a nonlinear state-space model via studying support vector regression and then stochastic uncertainty via studying an unscented KF. In the proposed KF-ANN approach, first the KF state (system) and observation (measurement) equations were created based on an ARIMA model. In a second step, the inputs variables of the ANN approaches were generated from the new state series that was the output of the state equation, while the target was the original wind speed series. As a result, the output of the ANN represents the final fitted or forecasting series.

This paper is organized as follows: Section 2 states the framework of this study and presents a hybrid KF-ANN based on ARIMA through ARIMA, KF, hybrid ARIMA-KF, and ANN theoretically. Section 3 displays and discusses the forecasting results and the computational steps of the methods in Section 2. Section 4 provides the conclusion of this study.

2. Material and method

2.1. Data and framework used in the study

In this study, daily wind speed data from two meteorological stations was collected. The first data set was collected from the Mosul Dam Meteorological Station in Mosul, Iraq. It covered four hydrological years (1 October 2000–30 September 2004) which was used for training. Another four months' of hydrological data (1 October 2004–31 January 2005) was reserved for testing. The other data set was collected from the Muar Meteorological Station in Johor, Malaysia. It covered four hydrological years (1 October 2006–30 September 2010) which was used for training. An additional three months' of hydrological data (1 October 2010–31 December 2010) was used for testing.

The framework of this study includes the following:

a. Determining the most appropriate ARIMA model following Box–Jenkins methodology.
b. Constructing the most appropriate ANN based on AR.
c. Constructing the most appropriate hybrid ARIMA-KF model.
d. Combining the KF model and ANN based on ARIMA to create a hybrid KF-ANN model.
e. Comparing the studied approaches to determine what model would provide the best forecasts. Fig. 1 demonstrates the framework of this study.

2.2. Mathematical models

2.2.1. Autoregressive integrated moving average (ARIMA) model

An ARIMA model was used to forecast wind speed. The modelling strategy followed Box–Jenkins methodology through the identification, estimation, diagnostic checking, and forecasting stages. A general expression of the seasonal ARIMA$(p,d,q)(P,D,Q)_s$ model is shown below: