Journal of Cleaner Production 134 (2016) 456–462

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression

Kuo-Ping Lin [a], Ping-Feng Pai [b,*]

[a] Department of Information Management, Lunghwa University of Science and Technology, No. 300, Section 1, Wanshou Road, Guishan, Taoyuan 333, Taiwan, ROC

[b] Department of Information Management, National Chi Nan University, No. 1, University Road, Puli, Nantou 545, Taiwan, ROC

ARTICLE INFO

Article history:
Received 30 October 2014
Received in revised form
30 June 2015
Accepted 24 August 2015
Available online 1 September 2015

Keywords:
Solar power output
Forecasting
Seasonal decomposition
Least-squares support vector regression
Genetic algorithms

ABSTRACT

Renewable power output is an important factor in scheduling and for improving balanced area control performance. This investigation develops an evolutionary seasonal decomposition least-square support vector regression (ESDLS-SVR) to forecast monthly solar power output. The construction of the ESDLS-SVR uses seasonal decomposition and least-square support vector regression (LS-SVR). Genetic algorithms (GA) are used simultaneously to select the parameters of the LS-SVR. Monthly solar power output data from Taiwan Power Company are used. Empirical results indicate that the proposed forecasting system demonstrates a superior performance in terms of forecasting accuracy. A comparative study has been introduced showing that the ESDLS-SVR model performance is better than autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA), generalized regression neural network (GRNN) and LS-SVR models.

1. Introduction

Transition to an energy-based economy that relies on renewable resources involves complex issues, such as variability, capacity and reliability of non-dispatchable energy resources (e.g., solar, wind, or tidal). As many countries are actively promoting cleaner sources of energy, the global demand for renewable energy integration to the power grids highlights the importance of economic and technological issues associated with growing levels of flat-panel PhotoVoltaic (PV), Concentrated Solar Power (CSP) and Concentrated PV (CPV) penetration into the power grid. These concerns arise from the variable nature of the solar resource, seasonal deviations in production and load profiles, the high cost of energy storage, and the attempt to obtain a balance between grid flexibility and reliability (Denholm and Margolis, 2007a, 2007b; Inman et al., 2013). In order to improve the integration stability of the output of solar systems into electric grids and optimize decision making at the management level of local storage systems and bidding into markets, the forecasting of solar systems is vital to reliably integrate them into the electricity grid.

Inman et al. (2013) reviewed the solar forecasting methods from 2008 to 2013, and found that high accuracy forecast systems are required for multiple time horizons. Solar power forecasting can be divided into the type of forecasting technique: numerical weather prediction (NWP)-based forecasting (Perez et al., 2010; Mathiesen and Kleissl, 2011), stochastic forecasting (Bacher et al., 2009; Yang et al., 2012), artificial intelligence (AI) forecasting model (Gordon, 2009; Paoli et al., 2010; Martin et al., 2010; Mandal et al., 2012; Pedro and Coimbra, 2012), and the hybrid forecasting model (Marquez et al., 2013). Recently, Bouzerdoum et al. (2013) proposed a hybrid forecasting model that combines Seasonal autoregressive integrated moving average (SARIMA) with support vector machine (SVM), for short-term power forecasting. This hybrid model can obtain better performance than the SARIMA and SVM models. Li et al. (2014) used Auto-Regression Moving Average-eXogenous (ARMAX), which improves the forecast accuracy of power output over the autoregressive integrated moving average (ARIMA) model. The review of past literature on solar power forecasting suggests the followings: (1) the solar power forecasting model for the multiple time horizons and variables should be developed; (2) the AI forecasting model has been successfully applied in solar power forecasting model; (3) Marquez et al. (2013) and Li et al. (2014) proposed hybrid forecasting methods which can successfully improve performance.

* Corresponding author. Tel.: +886 49 2910960.
E-mail address: paipf@ncnu.edu.tw (P.-F. Pai).

http://dx.doi.org/10.1016/j.jclepro.2015.08.099
0959-6526/

Recently, SVMs have proven one of the most powerful tools in dealing with classification problems (Cortes and Vapnik, 1995; Vapnik, 1995). SVMs employ the structural risk minimization principle that aims to minimize an upper bound of the generalization error instead of minimizing the training error. Based on this principle, SVMs achieve an optimum network structure. The SVM is equivalent to solving a linear constrained quadratic programming problem so that the solution of SVM is always unique and globally optimal. Furthermore, with some modifications, the SVM technique has been employed in regression and successfully solved time series problems in many fields (Cao, 2003; Cao and Gu, 2002; Tay and Cao, 2001; Hong and Pai, 2007; Lin and Pai, 2010; Lin et al., 2011). An alternative method, the least square support vector regression (LS-SVR) has been proposed to minimize sum square errors (SSEs) of training data sets (Van Gestel et al., 2001), while simultaneously minimizing the margin error (Peng and Wang, 2009; Suykens et al., 2002). Van Gestel et al. (2001) combined the Bayesian evidence framework with LS-SVR for financial time series forecasting. In predicting the DAX30 index, LS-SVR achieved a better performance than the ARIMA and nonparametric models. The LS-SVR has been successfully used to solve time series problems, and proven to be a superior forecasting model (Li et al., 2007, 2008; Jayadeva et al., 2008; Cui and Yan, 2009; Goodarzi et al., 2010; Lin et al., 2010; Qin et al., 2010; Quan et al., 2010; Yang et al., 2010; Deng and Yeh, 2011; Lin, 2013; Lin et al., 2013; Hung and Lin, 2013; Pai et al., 2014; Xie et al., 2014).

The aim of the present work is to develop an accurate evolutionary seasonal decomposition least-square support vector regression (ESDLS-SVR) model for forecasting the power output by using historical records of the produced power. Inman et al. (2013) and Bouzerdoum et al. (2013) mentioned that solar power output has seasonal deviations that are affected by solar radiation time of day. Moreover, the LS-SVR can achieve a better performance in time series problems by observing previous researches. ESDLS-SVR consists of seasonal decomposition, LS-SVR and genetic algorithm (GA) for improving the accuracy rate of the solar power forecasting model. The seasonal decomposition can reduce complexity of seasonal time series data. Based on the decomposition matrix, the LS-SVR predicts seasonal time series matrix. Furthermore, GA is employed to search optimal parameters. The results are compared with those from the ARIMA, SARIMA, GRNN and LS- SVR models, and the effectiveness of the ESDLS-SVR in long-term power output forecasting is evident.

The remainder of this paper is organized as follows: Section 2 presents the main construct of the ESDLS-SVR; Section 3 shows the solar power output dataset in Taiwan, discusses the experimental results, and compares them with the performances of various other models; Section 4 presents the conclusions.

2. Evolutionary seasonal decomposition least-square support vector regression

By enhancing the characteristics of the SVM model, the LS-SVR technique successfully integrated the least square method into the SVM model, and it has been applied in regression problems. In this study, the ESDLS-SVR is designed to effectively cope with seasonal time series data. The seasonal time series data are often complicated and thus difficult to predict. Effective seasonal time series forecasting is one of the topical subject in the forecasting system; hence, the ESDLS-SVR is combined seasonal decomposition with LS-SVR. Firstly, the seasonal decomposition procedure is conducted. This procedure can effectively produce detrended and seasonally adjusted series data, thereby reducing the complexity of seasonal time series data. Secondly, based on the decomposition data D_t matrix, the SDLS-SVR employs the forecasting capability of LS-SVR to predict the seasonal time series matrix. In addition, the GA technique is utilized to determine the proper parameters of the SDLS-SVR model. Finally, forecasting values and performance criteria are derived. Fig. 1 shows the framework of ESDLS-SVR.

This study uses a decomposition method, which has been successfully applied in time series forecasting models (Wang et al., 2012; Lin et al., 2012; Cheng and Wei, 2014; Xie et al., 2014). In order to capture seasonal characteristics of observations over past years, this study uses a popular multifactor model. The multifactor model can be defined as follows:

$$Y_t = f(T_t) \times S_t \times \varepsilon \tag{1}$$

where Y_t is the forecast value of time series at time t, $f(T_t)$ is the estimative value of trend T_t with SVR method at time t, S_t is the seasonal influence at time t and ε is the model noise. Therefore, the multifactor model can derive decomposition data D_t, which are respectively detrended data DT_t and seasonally adjusted data DS_t as:

$$D_t = [DT_t, DS_t] = [(S_t \times \varepsilon)/Y_t, (f(T_t) \times \varepsilon)/Y_t] \tag{2}$$

The DLS-SVR approach is employed to approximate an unknown function using an adjusted training data set $\{(x_t, D_t)\}$.

The DLS-SVR regression function can be formulated as Eq. (3). EDLS-SVR introduces decomposition and least squares to the SVR method by formulating the regression problem as:

$$\begin{aligned} &Min\ J_1(w,b,e) = \frac{1}{2}\|w\|^2 + \frac{1}{2}C\sum_{t=1}^{N} e_t^2 \\ &subjective\ \ to \\ &\begin{cases} DT_t = w^T\varphi(x_t) + b + e_t, & t = 1,2,\cdots,N.\forall D_t \in DT_t \\ DS_t = w^T\varphi(x_t) + b + e_t, & t = 1,2,\cdots,N.\forall D_t \in DS_t \end{cases} \end{aligned} \tag{3}$$

As the decomposition method can divide, detrend and seasonally adjust time series, the decomposition Lagrangian can be introduced as:

$$L_1(w,b,e,\gamma) = J_1(w,b,e) + \sum_{t=1}^{N} \gamma_t \left(D_t - w^T\varphi(x_t) - b - e_t\right), \tag{4}$$

where γ is the Lagrangian multiplier vector. The conditions for optimality are:

$$\begin{aligned} &\frac{\partial L_1}{\partial w} = 0 \Rightarrow w = \sum_{t=1}^{N} \gamma_t \varphi(x_t), \\ &\frac{\partial L_1}{\partial b} = 0 \Rightarrow \sum_{t=1}^{N} \gamma_i = 0, \\ &\frac{\partial L_1}{\partial e_i} = 0 \Rightarrow e_t = \frac{1}{C}\gamma_t, \quad t = 1,\ldots,N, \\ &\begin{cases} \dfrac{\partial L_1}{\partial \gamma_i} = 0 \Rightarrow DT_t = w^T\varphi(x_t) + b + e_t, & t = 1,\ldots,N.\forall D_t \in DT_t \\ \dfrac{\partial L_1}{\partial \gamma_i} = 0 \Rightarrow DS_t = w^T\varphi(x_t) + b + e_t, & t = 1,\ldots,N.\forall D_t \in DS_t \end{cases} \end{aligned} \tag{5}$$

Here, the value of the kernel function is equal to the inner product of two vectors. The relations of $K = (k_{tj})_{N \times N}$, $k_{tj} = k(x_t, x_j)$ and $V = \mathrm{diag}\{1/C, 1/C, \ldots, 1/C\}$ can be denoted. Thus,