Electrical Power and Energy Systems 44 (2013) 604–614

Contents lists available at SciVerse ScienceDirect

Electrical Power and Energy Systems

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

Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm

Wei-Chiang Hong [a,*], Yucheng Dong [b], Wen Yu Zhang [c], Li-Yueh Chen [d], B. K. Panigrahi [e]

[a] Department of Information Management, Oriental Institute of Technology, 58, Sec. 2, Sichuan Rd., Panchiao, Taipei County 220, Taiwan, ROC
[b] School of Business, Sichuan University, Chengdu 610065, China
[c] School of Information, Zhejiang University of Finance & Economics, Hangzhou 310018, China
[d] Department of Global Marketing and Logistics, MingDao University, 369 Wen-Hua Rd., Peetow, Changhua 52345, Taiwan, ROC
[e] Department of Electrical Engineering, Indian Institutes of Technology (IITs), India

ARTICLE INFO

Article history:
Received 1 July 2011
Received in revised form 3 August 2012
Accepted 8 August 2012
Available online 26 September 2012

Keywords:
Support vector regression (SVR)
Chaotic genetic algorithm (CGA)
Seasonal adjustment mechanism
Cyclic electric load forecasting

ABSTRACT

Application of support vector regression (SVR) with chaotic sequence and evolutionary algorithms not only could improve forecasting accuracy performance, but also could effectively avoid converging prematurely (i.e., trapping into a local optimum). However, the tendency of electric load sometimes reveals cyclic changes (such as hourly peak in a working day, weekly peak in a business week, and monthly peak in a demand planned year) due to cyclic economic activities or climate seasonal nature. The applications of SVR model to deal with cyclic electric load forecasting have not been widely explored. This investigation presents a SVR-based electric load forecasting model which applied a novel hybrid algorithm, namely chaotic genetic algorithm (CGA), to improve the forecasting performance. With the increase of the complexity and the larger problem scale of tourism demands, genetic algorithm (GA) is often faced with the problems of premature convergence, slowly reaching the global optimal solution or trapping into a local optimum. The proposed CGA based on the chaos optimization algorithm and GA, which employs internal randomness of chaos iterations, is used to overcome premature local optimum in determining three parameters of a SVR model. A numerical example from an existed reference is used to elucidate the forecasting performance of the proposed SSVRCGA model. The forecasting results indicate that the proposed model yields more accurate forecasting results than ARIMA and TF-ε-SVR-SA models. Therefore, the SSVRCGA model is a promising alternative for electric load forecasting.

1. Introduction

1.1. Traditional electric load forecasting approaches

Along with the rise of power market privatization and deregulation, the electric power industry is getting into free competitive era, thus, accurate electric load forecasting has become the most important issue in a regional or a national system for the market participants. For inaccurate electric load forecasting, it may increase operating costs [1,2]. For example, over estimation of future electric load results in unnecessary spinning reserve, wastes limited energy resources, even leads to distribution inefficiency, and, furthermore, is not accepted by international energy networks owing to excess supply. In contrast, under estimation of load causes failure in providing sufficient reserve and implies high costs in the peaking unit, which discourage any economic and industrial developments. Thus, the accuracy of future electric demand forecasting have received growing attention, particularly in the areas of electricity load planning, energy expenditure/cost economy and secure operation fields, in regional and/or national systems. However, the electric load forecasting is complicated and the electric load data presents nonlinear data patterns caused by the influencing factors, such as climate factors (temperature and humidity), social activities (human social activities) and seasonal factors (seasonal climate change and load growth).

In the past decades, various approaches have been proposed. The famous approach is, weather insensitive, employing historical load data to forecast future electric load, such as Box–Jenkins' ARIMA models [3–6], exponential smoothing models [7,8], multiplicative autoregressive (AR) model [9], Bayesian estimation model [10], and the state space and Kalman filtering technology [11–14]. The disadvantage of these methods is time consuming and unable to avoid the observation noise in the forecasting process, particularly for the situation when the number of variables is increased. Recently, to avoid a lot of variables selection problem, such as Azadeh et al. [15] employ fuzzy system to provide an ideal rule base to determine which type of ARMA models should be used. Wang et al. [16] propose hybrid ARMAX (auto-regressive and moving average with exogenous variables) model with particle swarm optimization to efficiently solve the problem of trapping into local minimum which is caused by exogenous variable (e.g., weather condition).

* Corresponding author. Tel.: +886 2 7738 8000x5316; fax: +886 2 7738 6310.
E-mail address: samuelsonhong@gmail.com (W.-C. Hong).

0142-0615/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijepes.2012.08.010

The second approach is regression model, which is based on the cause–effect relationships between electric load and relevant independent variables (weather, holiday, temperature, wind conditions, humidity, and so on), such as linear regression, lots of explanatory variables, such as ''weather'' (relative humidity, heating, and cooling degree-days), ''holiday'', ''temperature'', and ''economic and geographic'' (GDP, electricity price, and population), are considered into the regression model [17–21]. These forecasting approaches are difficult to take some significant improvement in terms of forecasting accuracy due to their theoretical limitations, such as unable to capture the rapid variational process changes underlying of electric load from historical data pattern. In addition, these models are based on linear assumption, however, these independent variables are unjustified to be used because of the terms are known to be nonlinear. Recently, Vilar et al. [22] introduce a nonparametric regression techniques with functional explanatory data and a semi-functional partial linear model to forecast 1-day electricity demand, by cutting the observed time series into a sample of functional trajectories and considering the past functional observation (last day trajectory) rather than multiple past time series values to incorporate vector covariates in the forecasting procedure.

To improve the performance of nonlinear electric load forecasting, artificial intelligence techniques are employed. Knowledge based expert system (KBES) approach [23] extracts rules from received relevant information (e.g., daily temperature, day type, load from the previous day, and so on), then, derives training rules and transforms the information into mathematical equations. This approach has the rule-based mechanism to transform new rule from received information, i.e., the expert capability is trained and increased by the existence presuming to increase forecasting accuracy [23–25]. This approach is derivation of the rules from on-the-job training and sometimes transforming the information logic to equations could be impractical. Recently, applications of fuzzy inference system and fuzzy theory in load forecasting to improve the shortcomings of KBES are also received attentions. Ying and Pan [26] introduce adaptive network fuzzy inference system (ANFIS), by looking for the mapping relation between the input and output data to determine the optimal distribution of membership functions, to forecast regional load. Chen and Wang [27] and Chang et al. [28] all employ fuzzy approaches to get superior performance in terms of load forecasting.

Accompanying the mature nonlinear mapping capabilities and data processing characteristics, artificial neural network (ANN) has received wide successful applications in improving load forecasting accuracy [14,29–32]. These experimental results indicate that the ANN model is superior to ARIMA and regression models in terms of forecasting accuracy. ANN-based models seem to obtain improved and acceptable performance in load forecasting issue and to provide the possible nonlinear extensions to the ARIMA models in load time series. However, the training procedure of ANNs models is not only time consuming but also possible to get trapped in local minima and subjectively in selecting the model architecture [33]. Recently, to improve the drawbacks of ANN, many applications of hybrid ANN with statistical methods or other intelligent approaches have received a lot of attentions, such as hybrid with Bayesian inference [34,35], self-organizing map [36,37], wavelet transform [38,39], particle swarm optimization [40], and cooperative co-evolutionary approach [41].

1.2. Applications of support vector regression models

Support vector regression (SVR) [42] has been successfully used to solve forecasting problems in many fields, such as financial time series (stocks index and exchange rate) forecasting [43–47], tourist arrival forecasting [48,49], atmospheric science forecasting [50–53], and traffic flow forecasting [54,55]. Meanwhile, SVR model has also been successfully applied to forecast electric load [56–61]. The empirical results indicate that the selection of the three parameters C (to trade off the training errors and large weights), ε (the width of the insensitive loss function), and σ (the parameter of Gaussian kernel function) in a SVR model influences the forecasting accuracy significantly. Although, numerous publications in the literature had given some recommendations on appropriate setting of SVR parameters [62], however, those approaches do not simultaneously consider the interaction effects among the three parameters. It is feasible to employ optimization solving procedure to obtain suitable parameters combination, such as minimizing the objective function describing the structural risk mentioned above, thus, evolutionary algorithms are employed to determine appropriate parameter values. Authors have conducted a series of relevant researches, by employed different evolutionary algorithms (such as genetic algorithm, simulated annealing algorithms, immune algorithms, particle swarm optimization, and Tabu search) for parameters determination, to identify that which algorithm is suited for specified data patterns. In which, all SVR with different evolutionary algorithms are superior to other competitive forecasting models (ARIMA, regression models, and ANNs etc.), however, these algorithms are almost lack of knowledge memory or storage functions which would lead to neither time consuming nor inefficiency in the searching the suitable parameters (i.e., being trapped in local optimum). Therefore, authors have also conducted several trials by employing hybrid chaotic sequence with evolutionary algorithms for parameters determination to overcome the immature convergence (trapped in local optimum) problem [48,50,57,59]. In order to continue testing the stability and superiority of hybrid chaotic sequence with evolutionary algorithms, this investigation tries to employ the chaotic genetic algorithm (CGA) to determine the values of three parameters in a SVR model.

1.3. Chaotic genetic algorithm in parameter determination of a SVR model

Genetic algorithm (GA) is auto-adaptive stochastic search techniques [63], it generates new individuals with selection, crossover and mutation operators. GA starts with a coding of the parameter set of all types of objective functions, thus, it has the ability in solving problems those traditional algorithms are not easily to solve. In Pai and Hong [49,60], GA is able to reserve a few best fitted members of the whole population. However, the selection operation rules of GA mean only the few best fitting members of the whole population in a generation can survive. The population diversity is significantly reduced after some generations, meaning that GA might lead to a premature convergence to a local optimum in the searching for suitable parameters of a SVR model. To overcome these drawbacks, some effective approaches and improvements on GA need to be discovered to maintain the population diversity and avoid leading to misleading local optimum. One possible approach is to divide the chromosome population into several subgroups, and limit the crossover between the members in different subgroups in order to maintain the population diversity. However, this method requires a very large population size, which is not typical in business forecasting application problem solving.

Another feasible scheme focuses on the chaos approach, due to its easy implementation and unique ability to avoid becoming trapped in local optima [64]. Chaos, defined as highly unstable motion in finite phase space, often occurs in deterministic nonlinear dynamic systems [65,66]. Such motion is very similar to a random process (''*randomicity*''). Therefore, any variable in the chaotic space can travel ergodically over the whole space of interest (''*ergodicity*''). The variation of those chaotic variables has a delicate inherent rule, even though its variation looks like disorder (''*regularity*''). Additionally, it is extremely sensitive to the initial condition, which