Electric Power Systems Research 95 (2013) 66–76

Contents lists available at SciVerse ScienceDirect

Electric Power Systems Research

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

Day-ahead electricity price forecasting based on panel cointegration and particle filter

X.R. Li[a], C.W. Yu[a,*], S.Y. Ren[b], C.H. Chiu[c], K. Meng[d]

[a] Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong
[b] Institute of Textiles & Clothing, The Hong Kong Polytechnic University, Hong Kong
[c] Sun Yat-Sen Business School, Sun Yat-Sen University, China
[d] Centre for Intelligent Electricity Networks, The University of Newcastle, Australia

ARTICLE INFO

Article history:
Received 3 March 2012
Received in revised form 7 June 2012
Accepted 18 July 2012
Available online 2 October 2012

Keywords:
Price forecasting
Panel data
Cointegration model
Particle filter

ABSTRACT

An accurate forecasting of energy price is important for generation companies (GENCOs) to develop their bidding strategies or to make investment decisions. Nowadays, day-ahead electricity market is closely associated with other commodity markets such as fuel market and emission market. Under such an environment, day-ahead electricity price is volatile and its volatility changes overtime due to the uncertainties from the multi-market. This paper proposes a two-stage hybrid model based on panel cointegration and particle filter (PCPF). Panel cointegration (PC) model, which utilizes information of both the inter-temporal dynamics and the individuality of interconnected regions, provides powerful forecasting tool for electricity price. Particle filter (PF) has achieved significant successes in tracking applications involving non-Gaussian signals and nonlinear systems. This paper has two main focuses: (1) To expand the dimension of electricity price dataset from time series to panel data so that the dynamics of interconnected regions can be analyzed simultaneously and considered as a whole. (2) Regarding the model coefficients as a time-varying process, PF is used to forecast electricity price adaptively. In the case study, the proposed PCPF model is applied to the real electricity market data of PJM in the year 2008. Promising results show clearly the superior predicting behavior of the proposed modeling.

1. Introduction

Nowadays, day-ahead electricity market is closely associated with other commodity markets such as fuel market and emission market. This evolution has changed the roles of market participants and complicated the analysis of electricity price [1]. Day-ahead electricity price is volatile and its volatility changes overtime due to the uncertainties from the multi-market. GENCOs always adjust their bidding strategies to achieve their maximum benefits according to their analysis. Similarly, consumers would derive a plan to optimize their purchased electricity from the pool, or use other financial derivatives to protect themselves against high prices [2,3].

The available forecasting methods can be broadly classified into three categories: system simulation models, market equilibrium analysis, and time series models [4,5]. System simulation models usually concentrate on detailed insight of price formation [6]. Factors such as actual dispatch according to system operating requirements and transmission constraints are considered. Market equilibrium analysis, on the other hand, involves economics and game theory [7]. In addition to the forecasted prices, these two categories always come up with general equilibrium or market strategic behaviors. Time series models are widely used to forecast electricity price. Electricity price is forecasted through statistical methods with little attentions paying to the reasons of the price changing. This type of methods can be divided into three subtypes, namely regression based models, stochastic time series models and intelligent learning models.

Regression based models analyze the assumed relationship between electricity price and a number of independent variables that are known or estimated [8]. These methods overcome serial correlation problems. However, they do not always work well in practice since they assume the variables are stationary or stationary after the application of statistical techniques such as differencing.

Stochastic time series models are proposed to deal with nonstationary time series. Both autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models work by iteratively identifying a parametric model from hypothesized models and estimating the corresponding parameters based on observations. When the series have high volatility and price spikes, GARCH model [5,9] is a good alternative because it considers the conditional variance as time dependent. However, the

* Corresponding author.
E-mail address: eecwyu@inet.polyu.edu.hk (C.W. Yu).

0378-7796/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.epsr.2012.07.021

identification and estimation of these models can be badly distorted by outlier effects.

Intelligent learning models derived from Neural Networks (NN) and data mining have also been studied. Artificial neural networks (ANN) are defined as information processing systems which have common specific characteristics associated to biological networks. ANN is capable to model nonlinear input/output mapping functions. Its families have strong fault tolerance ability though they usually require long training time. A standard ANN is a group of interconnected neural processing units imitating the brain activation. Recent studies on ANN focus on the determination of the best forecasting model by comparing various neural architectures, applying several decomposition techniques or selecting proper transfer functions [10]. Wu and Shahidehpour [11] proposed a hybrid time-series and adaptive wavelet neural network (AWNN) model, composed of linear and nonlinear relationships of prices and explanatory variables, for day-ahead price forecasting. AWNN was used to present the nonlinear, nonstationary impact of load series on electricity prices. Amjady [12] developed a fuzzy neural network (FNN) which combined fuzzy logic and standard ANN to provide more accurate results than ARIMA, wavelet-ARIMA, multilayer perceptron, and radial basis function neural networks. Kernel-based machine learning method such as support vector machine (SVM) [9] has shown good accuracy and efficiency in some real-world problems. Furthermore, relevance vector machine (RVM) [9] has been proved outperforms SVM in both the forecast accuracy and computational efficiency. However, the performance of these machine learning models relies on heuristics, e.g. the choices of kernel and penalty functions.

The integration of different commodity markets leads to uncertainties in both time and space dimensions. Furthermore, electricity price features high volatility, non-stationary and non-linear patterns [12]. Having a robust and accurate prediction model for day-ahead electricity price is important under these circumstances. However, rare literature has studied both inter-temporal dynamics and inter-regional interactions of uniform day-ahead price among different interconnected regions. Furthermore, little attention has been paid to develop methods that can handle both linear and nonlinear problems simultaneously. Therefore, a novel panel cointegration and particle filter (PCPF) model is proposed in this paper to predict day-ahead electricity price.

The rest of this paper is organized as follows: In Section 2, the problem formulation and the evaluation criteria are described. Section 3 presents the forecasting framework of the proposed PCPF model. A case study is conducted in Section 4 to demonstrate the effectiveness of the PCPF model. Section 5 concludes the paper.

2. Problem formulation

A pool-based electricity market (EM) with N interconnected regions is studied in this paper. In this kind of market (e.g. PJM), the operator coordinates the movement of electricity through the interconnected power grid. The uniform price in the day-ahead market is affected significantly by the regional loads because electricity price is calculated based on the consideration of the entire gird [4]. On the other hand, the variability of the uniform price can influence the energy-usage patterns and introduce new trends. Fig. 1 shows a snapshot of seven regional loads of PJM from Jan 1 to Jan 7, 2008.

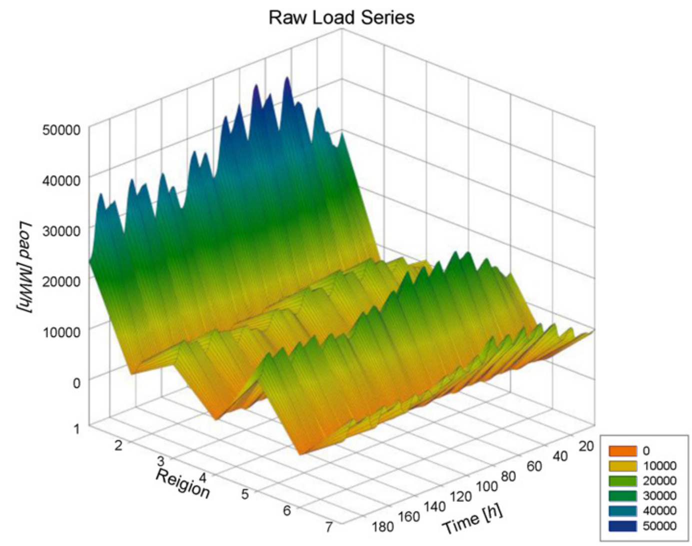

Fig. 1. Loads of 7 regions from Jan 1, 2008 to Jan 7, 2008.

It can be seen that significant loading differences, which change from time to time, exist among regions. Besides, significant individual intra-day variations can be observed. In other words, both the inter-temporal dynamics and the inter-regional interactions exist in the panel. In addition, nonlinear patterns exist in the relationship among the uniform day-ahead price and the loads of different regions. In this case, using time series models such as regression based models and stochastic time series models cannot capture the complex nonlinear behavior. On the other hand, using intelligent method such as NN will yield mixed results and the efficiency will vary from case to case.

The proposed PCPF model is introduced to tackle these difficulties by using a two-stage forecasting framework. It has two features that differentiate it from other existing techniques. First, it makes prediction by using historical loading data of different regions in the pool and constructs the regional loading data together with the uniform day-ahead price as a panel [13]. Using panel data, both the impacts of inter-temporal dynamics and inter-regional loading differences on the uniform day-ahead price can be taken into account. Secondly, PF is applied as a post-processor to effectively handle the nonlinearity and the volatility of electricity price. A case study on PJM day-ahead electricity market is conducted. The main reason of selecting PJM is due to the fact that its market participants' decisions are affected not only by the uncertainties in the day-ahead electricity market, but also by the fuel market and the emission market in the Regional Greenhouse Gas Initiative (RGGI) scheme. GENCOs' bidding strategies and investment plans have been changed after the implementation of RGGI in year 2009 [14]. Through the case study, it can be demonstrated that the proposed method can effectively deal with volatility and nonlinearity in a pool environment with power resources allocated dispersedly in different regions.

The data in the period Jan 1, 2007–Dec 31, 2007 is used for the estimation of the models, while the data in the period Jan 1, 2008–Dec 1, 2008 is used for out-of-sample testing. To assess and compare the performance of the models, weekly mean absolute percentage error (*WMAPE*), *WMAPE* for period j($WMAPE_j$), daily mean absolute percentage error (*DMAPE*), weekly mean absolute error (*WMAE*) and weekly root mean square error (*WRMSE*) indices are adopted in this paper.

$$WMAPE = \frac{1}{168}\sum_{t=1}^{7}\sum_{j=1}^{24}\frac{|X_{j,t}^{A} - X_{j,t}^{F}|}{X_{j,t}^{A}} \tag{1}$$

$$WMAPE_j = \frac{1}{7}\sum_{t=1}^{7}\frac{|X_{j,t}^{A} - X_{j,t}^{F}|}{X_{j,t}^{A}} \tag{2}$$