Electrical Power and Energy Systems 63 (2014) 353–362

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework

Tao Xiong [a], Yukun Bao [b,*], Zhongyi Hu [b]

[a] College of Economics and Management, Huazhong Agricultural University, Wuhan 630027, PR China
[b] School of Management, Huazhong University of Science and Technology, Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 17 October 2013
Received in revised form 12 May 2014
Accepted 5 June 2014
Available online 3 July 2014

Keywords:
Interval-valued data
Electricity demand forecasting
Bivariate empirical mode decomposition (BEMD)
Support vector regression (SVR)

ABSTRACT

Highly accurate interval forecasting of electricity demand is fundamental to the success of reducing the risk when making power system planning and operational decisions by providing a range rather than point estimation. In this study, a novel modeling framework integrating bivariate empirical mode decomposition (BEMD) and support vector regression (SVR), extended from the well-established empirical mode decomposition (EMD) based time series modeling framework in the energy demand forecasting literature, is proposed for interval forecasting of electricity demand. The novelty of this study arises from the employment of BEMD, a new extension of classical empirical model decomposition (EMD) destined to handle bivariate time series treated as complex-valued time series, as decomposition method instead of classical EMD only capable of decomposing one-dimensional single-valued time series. This proposed modeling framework is endowed with BEMD to decompose simultaneously both the lower and upper bounds time series, constructed in forms of complex-valued time series, of electricity demand on a monthly per hour basis, resulting in capturing the potential interrelationship between lower and upper bounds. The proposed modeling framework is justified with monthly interval-valued electricity demand data per hour in Pennsylvania–New Jersey–Maryland Interconnection, indicating it as a promising method for interval-valued electricity demand forecasting.

Introduction

To supply high quality electric energy to the customer in a secure and economic manner, an electricity operator faces many technical and economical problems in operation, planning, control and reliability management of an electric power system. In achieving this goal, accurate forecast of electricity demand is the first prerequisite.

According to extent literature investigation, it is not hard to find that a wide variety of methodologies and techniques have been used for electricity demand forecasting with many degrees of success, such as exponential smoothing models [1], regression models [2], fuzzy logic approach [3], fuzzy inference system [4], grey-based approaches [5], wavelet transforms and adaptive models [6], kernel-based method [7], artificial neural networks [8,9], support vector machines [10,11], semi-parametric method [12] and hybrid model [13,14]. The reader is referred to [15] for a recent survey of the presented methodologies and techniques employed for electricity demand forecasting. However, an important point to note from past studies mentioned above is their preoccupation with point forecasting rather than interval one.

An interval forecasting of electricity demand has the advantage of taking into account the variability and/or uncertainty so as to reduce the amount of random variation relative to that found in classic single-valued load time series. As García-Ascanio and Maté [16] pointed out, the interval-valued time series forecasting (ITS) methods as a potential tool that will lead to a reduction in risk when making power system planning and operational decisions. To date, several suitable tools for managing ITS have been developed (see [17] for a recently review), such as interval Holt's exponential smoothing methods [18], interval multi-layer perceptrons (iMLP) [19], vector autoregressive (VAR) model [16], and vector error correction (VEC) model [20]. Most of the conventional methodologies available for ITS in the literature propose the use of computational methods or modeling schema, accounting for the capability of dealing with interval-valued data. For instance, neural networks are an area that has recently witnessed substantial improvements in interval analysis. Notable earlier work on interval analysis using neural networks includes that of Simoff [21], while Beheshti et al. [22] developed a multi-layer perceptron (MLP) in which inputs, weights, biases, and outputs are intervals, and

* Corresponding author. Tel.: +86 27 87558579; fax: +86 27 87556437.
E-mail addresses: yukunbao@hust.edu.cn, y.bao@ieee.org (Y. Bao).

http://dx.doi.org/10.1016/j.ijepes.2014.06.010
0142-0615/

proposed an algorithm so as to obtain the optimal weights and biases. Recently, Roque et al. [19] proposed and analyzed a new model of MLP (named iMLP in [19]) based on interval arithmetic that facilitates handling input and output interval-valued data, but where weights and biases are single-valued and not interval-valued.

Our study focuses on extending the EMD-based time series modeling framework to adapt to the scenario of interval forecasting of electricity demand. Following the philosophy of 'divide and conquer', EMD-based time series modeling framework has been recently well-established and well-justified in energy market [23–25], tourism management [26], hydrology [27], transportation research [28], and forecasting competition [29]. Generally speaking, there are three main steps involved in the classical EMD-based time series modeling framework, i.e., decomposition, single forecast, and ensemble forecast. First of all, EMD is used to divide the original single-valued time series into a finite number of intrinsic mode function (IMFs) components and a residue component. Secondly, some powerful modeling techniques such as artificial neural network (ANN) and support vector regression (SVR), are applied to model and predict all components (IMFs and residue) respectively. Finally, these prediction results of all extracted IMFs components and the residue in the previous step are combined to generate an aggregated output using one modeling technique, which can be seen as the final prediction results for the original time series. However, the classical EMD applied in the aforementioned studies [23–29] is only applicable to one-dimensional (univariate) single-value time series decomposition. One straightforward solution to extend the classical EMD-based modeling framework for interval forecasting of electricity demand is to decompose and forecast the lower and upper bound series of interval-valued electricity demand respectively as several studies [30,31] do, without considering the possible interrelations that are presented amongst themselves, which has been criticized in [17]. To enhance the capability of classical EMD, fortunately, BEMD [32] has been recently proposed to extend this decomposition method to treat complex-valued signals. Although BEMD was not specially developed for interval-valued time series analysis purposes, in view of the BEMD's advantages in decomposing complex-valued signals, this study proposes to construct a interval-valued electricity demand time series consisting of both lower and upper bounds of monthly electricity demand per hour in forms of complex-valued time series. This is to say, the lower and upper bounds series are taken as the real and imaginary parts of the complex-valued signal respectively. After this generic construction, the interval-valued time series of electricity demand can fully be decomposed through BEMD and then be fed into the proposed BEMD-based time series modeling framework for forecasting task after the selection of a specific modeling technique such as SVR in this study.

Our contributions could be outlined as follows. First, as was mentioned above, although extensive mounts of approaches [1–14] have been developed for electricity demand modeling and forecasting, most of them rely only on single-valued electricity demand series. The interval forecasting of electricity demand has not been widely explored (in fact, we have only come across interval-valued electricity demand forecasting in one published work [16]). Second, by introducing BEMD, the well-established EMD-based time series modeling framework [23–29] can be extended to deal with interval-valued time series forecasting, which can fully take the advantages of its simplification of modeling process in nature of 'divide and conquer' as well as avoiding to increase computational cost while employing complex-valued modeling techniques to deal with interval-valued time series. Third, since the work of Rilling et al. [32], BEMD has attracted particular attention in engineering technology filed [33–36]. However, there have been very few, if any, studies for interval-valued time series forecasting using the BEMD-based modeling framework. So, we hope this study would fill this gap. The fourth contribution is straightforward to provide the empirical evidence on the interval-valued electricity demand forecasting with real-world data from Pennsylvania–New Jersey–Maryland (PJM) market. Given the hourly single-valued electricity demand series from PJM market, we calculate the maximum and minimum value of the demand per hour and month from 2000 to 2011. This produces an interval-valued time series where each observation is formed by an interval that collects, as its lower bound, the minimum value of the electricity demand and, as its upper bound, the maximum value of the electricity demand for a specific hour, month and year.

This paper is structured as follows. In Section 'BEMD with interval-valued electricity demand time series', we provide brief introduction of BEMD and illustrate the data representation of interval-valued electricity demand time series analysis. Afterwards, the proposed BEMD-based SVR modeling framework are discussed in detail in Section 'The proposed BEMD-SVR modeling framework'. Section 'Research design' details the research design on data process and preliminary analysis, accuracy measure, methodologies implementation, and experimental procedure. Following that, in Section 'Results', the experimental results are discussed. Section 'Conclusions' finally concludes this work.

BEMD with interval-valued electricity demand time series

In this section, the overall formulation process of the BEMD for interval-valued electricity demand time series forecasting is presented. First, the data representation of interval-valued electricity demand time series is illustrated. Then the BEMD for the obtained ITS is formulated in details.

Constructing interval-valued electricity demand time series

Classical statistics and data analysis deal with individuals who can be described by a classic variable that takes as its value either a real value (for a quantitative variable) or a category (for a nominal variable). However, observations and estimations in the real world are usually incomplete to represent classic data exactly. In the electricity market, for instance, electricity demand has its daily (or weekly, or monthly) bounds and varies in each period-day, week, or month. Representing the variations with snap shot points, say the highest demand, only reflects a particular number at a particular time; it does not properly reflect its variability during the period. This problem can be reduced if the higher and lower demand per period is considered, giving rise to an interval-valued time series.

Interval-valued data is a particular case of symbolic data in the field of symbolic data analysis (SDA) [37]. SDA states that symbolic variables (lists, intervals, frequency distributions, etc.) are better suited than single-valued variables for faithfully describing complex real-life situations [31]. It should be noted that interval-valued data in the field of SDA do not come from noise assumptions, but rather from the expression of variation or aggregation of huge databases into a reduced number of groups [18].

In the context of SDA, an interval-valued variable, $[Y]$, is a variable defined for all the elements i of a set E, where $[Y_i] = \{[Y_i^L,\ Y_i^U] : Y_i^L,\ Y_i^U \in \mathbb{R},\ Y_i^L \le Y_i^U\},\ \forall i \in E$. The particular value of $[Y]$ for the ith element can be either denoted by the interval lower and upper bounds $[Y_i] = [Y_i^L,\ Y_i^U]$ or the center (mid-point) and radius (half-range) $[Y_i] = [Y_i^C,\ Y_i^R]$, where $Y_i^C = (Y_i^L + Y_i^U)/2$ and $Y_i^R = (Y_i^U - Y_i^L)/2$. In Table 1, the interval-valued in every month of the hourly spot electricity demand in Pennsylvania–New Jersey–Maryland (PJM) Interconnection in MW h per day and per hour in 2011 is showed.