G Model
RECYCL-3200; No. of Pages 8

ARTICLE IN PRESS

Resources, Conservation and Recycling xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Resources, Conservation and Recycling

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

Short-term forecasting of electricity demand for the residential sector using weather and social variables

Hyojoo Son, Changwan Kim *

Department of Architectural Engineering, Chung-Aug University, 84 Heukseok-Ro Dongjak-Gu, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 31 August 2015
Accepted 21 January 2016
Available online xxx

Keywords:
Short-term electricity demand
Forecasting
Residential sector
Feature selection
Support vector regression

ABSTRACT

The aim of this study is to provide a precise model for the one-month-ahead forecast of electricity demand in the residential sector. In this study, a total of 20 influential variables are considered including monthly electricity consumption, 14 weather variables, and five social variables. Based on support vector regression and fuzzy-rough feature selection with particle swarm optimization algorithms, the proposed method established a model with variables that relate to the forecast by ignoring variables that inevitably lead to forecasting errors. The proposed forecasting model was validated using historical data from South Korea between January 1991 and December 2012. The first 240 months were used for training and the remaining 24 for testing. The performance was evaluated using MAPE, MAE, RMSE, MBE, and UPA values. Furthermore, it was compared with that obtained from the artificial neural network, autoregressive integrated moving average, multiple linear regression models, and the methods proposed in the previous studies. It was found to be superior for every performance measure considered in this study. The proposed method has an advantage over the previous methods because it automatically determines appropriate and necessary variables for a reliable forecast. It is expected that it can contribute to more accurate forecasting of short-term electricity demand in the residential sector. The ability to accurately forecast short-term electricity demand can assist power system operators and market participants in ensuring sustainable electricity planning decisions and securing electricity supply for the consumers.

1. Introduction

Short-term electricity demand forecasting plays a significant role in power system planning, including economic scheduling of generating capacity, scheduling of fuel purchases, and power system management (Thatcher, 2007; Shrivastava and Misra, 2008; Kavaklioglu et al., 2009; Mao et al., 2009; Wang et al., 2009; Azadeh et al., 2011; Apadula et al., 2012; Wang, 2012). It is especially obvious that accurate electricity forecasting has great importance to the residential sector, a major contributor to the peak loads in most electricity systems. Overestimating electricity demands misleads planners and wastes resources with expensive expansion plans. Such overestimation also increases operating costs since electricity, unlike other energy sources, cannot be stored on a large scale (Keyno et al., 2009; Qiu, 2013). But underestimation of electricity demands will result in failures and shortages (Kavaklioglu et al., 2009).

Short-term electricity demand forecasting in the residential sector is a complex problem because its rise and fluctuation are caused by the differences in demand from month to month. In addition, consumption is influenced by many nonlinear variables, such as weather conditions, economics, and demographics (Raman and Sunilkumar, 1995; Valor et al., 2001; Altinay and Karagol, 2005; Kavaklioglu et al., 2009; Swan and Ugursal, 2009; Aroonruengsawat and Auffhammer, 2011; Apadula et al., 2012).

Several research studies have been conducted over the last several decades to explore the complex problem of monthly electricity demand forecasting by means of multivariate time series analysis (Quayle and Diaz, 1980; Tserkezos, 1992; Islam et al., 1995; Al-Alawi and Islam, 1996; Al-Zayer and Al-Ibrahim, 1996; Abdel-Aal et al., 1997; Islam and Al-Alawi, 1997; Sailor and Munoz, 1997; Lam, 1998; Yan, 1998; Sailor, 2001; Amato et al., 2005; Hor et al., 2005; Mirasgedis et al., 2006; Pao, 2006; Ruth and Lin, 2006; Lam et al., 2008; Wangpattarapong et al., 2008; Abosedra et al., 2009; Bunnoon et al., 2010; Lam et al., 2010; Pilli-Sihvola et al., 2010; Adam et al., 2011; Chang et al., 2011; Apadula et al., 2012; Bunnoon, 2013; Bunnoon et al., 2013; Chen et al., 2013). Most of the previous studies (Al-Zayer and Al-Ibrahim, 1996; Sailor and Munoz, 1997; Lam, 1998; Hor et al., 2005; Mirasgedis et al., 2006; Lam et al., 2008;

* Corresponding author.
E-mail addresses: hjson0908@live.cau.ac.kr (H. Son), changwan@cau.ac.kr (C. Kim).

http://dx.doi.org/10.1016/j.resconrec.2016.01.016

Please cite this article in press as: Son, H., Kim, C., Short-term forecasting of electricity demand for the residential sector using weather and social variables. Resour Conserv Recy (2016), http://dx.doi.org/10.1016/j.resconrec.2016.01.016

Wangpattarapong et al., 2008; Apadula et al., 2012) have assumed that the input and output series are stationary and applied statistical models. However, real monthly electricity demand series, as well as variables that may influence the electricity demand series, have been found to be nonstationary (Tserkezos, 1992).

When one or more of the series are nonstationary, it is necessary to consider sophisticated models that are capable of describing the nonlinear input and output series. In order to solve nonlinear time series problems, several research studies applied auto-regressive integrated moving average model (Tserkezos, 1992; Islam and Al-Alawi, 1997) and neural network model-based approaches (Islam et al., 1995; Al-Alawi and Islam, 1996; Pao, 2006; Chang et al., 2011; Bunnoon, 2013; Bunnoon et al., 2013). The accuracy of the resulting models has ranged from 1.42% (Bunnoon et al., 2013) to 10.98% (Islam and Al-Alawi, 1997) in terms of mean absolute percentage error. However, it is difficult to say whether these models have been sufficiently validated, because the evaluation periods of most research studies were within about 10 years (Islam et al., 1995; Abdel-Aal et al., 1997; Islam and Al-Alawi, 1997; Adam et al., 2011; Chang et al., 2011; Bunnoon, 2013; Bunnoon et al., 2013; Chen et al., 2013). To our knowledge, there have been no research studies examining the method that can develop a model by determining appropriate and necessary variables as well as handling nonlinear time series problems for a reliable forecast of monthly electricity demand in the residential sector.

The aim of this study is to provide a precise model for one-month-ahead forecast of electricity demand in the residential sector. Based on support vector regression and fuzzy-rough feature selection with particle swarm optimization algorithms, the proposed method automatically develops a forecasting model with variables that relate to the electricity demand series by ignoring variables that inevitably lead to forecasting errors. To evaluate the forecasting performance of the proposed method, we performed a comprehensive comparison of the prediction performance of the proposed method versus that of the artificial neural network, auto-regressive integrated moving average, multiple linear regression models, and the methods proposed in the previous studies. A data set covering the period from January 1991 to December 2012 was collected from South Korea and used for training and testing experiments.

In Section 2, a comprehensive review of the literature is presented. In Section 3, the data set and pre-processing and some materials on the proposed methodology are described. In Section 4, a discussion and analysis of the experimental results are presented. Section 5 contains conclusions and suggestions for future research.

2. Related works

According to the Annual Energy Outlook 2014 by the U.S. Energy Information Administration, although the growth in electricity use has slowed down, it is still forecasted to increase by 29% from 2012 to 2040, with the residential sector accounting for the largest portion of energy consumption (US Energy Information Administration, 2014). Depending on the country, the residential sector accounted for anywhere between 16% and 50% of energy consumption, averaging approximately 30% worldwide (Swan and Ugursal, 2009). This considerable amount of the energy consumption in the residential sector shows that a thorough understanding is required to prepare for and help guide consumption in an increasingly energy-conscious world, particularly with regard to supply, efficient use, and effects of consumption (Swan and Ugursal, 2009).

As part of this point of view, several research studies have explored this complex problem of monthly electricity demand forecasting by means of multivariate time series analysis (Quayle and Diaz, 1980; Tserkezos, 1992; Islam et al., 1995; Al-Alawi and Islam, 1996; Al-Zayer and Al-Ibrahim, 1996; Abdel-Aal et al., 1997; Islam and Al-Alawi, 1997; Sailor and Munoz, 1997; Lam, 1998; Yan, 1998; Sailor, 2001; Amato et al., 2005; Hor et al., 2005; Mirasgedis et al., 2006; Pao, 2006; Ruth and Lin, 2006; Lam et al., 2008; Wangpattarapong et al., 2008; Abosedra et al., 2009; Bunnoon et al., 2010; Lam et al., 2010; Pilli-Sihvola et al., 2010; Adam et al., 2011; Chang et al., 2011; Apadula et al., 2012; Bunnoon, 2013; Bunnoon et al., 2013; Chen et al., 2013). This chapter comprehensively reviews the variables which appear to have been considered in the literature. To date, many nonlinear variables, such as weather conditions, economics, and demographics, were found to have influence on electricity demand series (Raman and Sunilkumar, 1995; Valor et al., 2001; Altinay and Karagol, 2005; Kavaklioglu et al., 2009; Swan and Ugursal, 2009; Aroonruengsawat and Auffhammer, 2011; Apadula et al., 2012). Table 1 summarizes some of the many representative variables related to the weather conditions, economics, and demographics. The comprehensive list in Table 1 was developed from 28 studies.

The variables that were considered no more than twice in the literature were not included in Table 1, so research studies by Mirasgedis et al. (2006) and Apadula et al. (2012) were not included the list. For weather variables, the variables excluded from Table 1 were *absolute maximum temperature* and *average maximum relative humidity* (Islam et al., 1995; Islam and Al-Alawi, 1997), *maximum relative humidity* (Al-Alawi and Islam, 1996), *comfort index* (Al-Alawi and Islam, 1996; Islam and Al-Alawi, 1997), *temperature at peak load* and *relative humidity at peak load* (Islam and Al-Alawi, 1997), *cooling radiation-days* (Lam, 1998), *clo* (Yan, 1998), *cloud cover* (Yan, 1998; Apadula et al., 2012), *annual trend heating degree days* and *annual trend cooling degree days* (Amato et al., 2005), *total number of monthly heating degree days* and *total number of monthly cooling degree days* (Mirasgedis et al., 2006), *dry-bulb temperature* and *wet-bulb temperature* (Lam et al., 2008; Bunnoon et al., 2010), *clearness index* (Lam et al., 2008, 2010), *rainy time* and *average air pressure* (Chang et al., 2011), and *heating degree months* and *cooling degree months* (Apadula et al., 2012).

For social variables, the variables excluded from Table 1 were *number of consumers connected* (Al-Alawi and Islam, 1996; Islam and Al-Alawi, 1997), *household size* (Lam, 1998), *household income* (Lam, 1998; Wangpattarapong et al., 2008), *number of non-working days during the month* (Mirasgedis et al., 2006), national income (Pao, 2006), *number of houses* and *monthly number of air conditioners sold* (Wangpattarapong et al., 2008), *real import* (Abosedra et al., 2009), *weighted calendar variable* (Apadula et al., 2012), and *gross value of export-import*, *industry increasing value*, and *social retail sales of consumer goods* (Chen et al., 2013). Although some of the selected variables may inevitably lead to forecasting errors, most of the previous studies have selected such variables based on human intuition and expertise in this subject area or based on validation results of one-to-one correlations between the variables and electricity demand series to develop forecasting models (Islam et al., 1995; Al-Zayer and Al-Ibrahim, 1996; Islam and Al-Alawi, 1997; Sailor and Munoz, 1997; Lam, 1998; Hor et al., 2005; Lam et al., 2008; Wangpattarapong et al., 2008; Apadula et al., 2012).

Most of the previous studies have assumed that certain inputs among various weather variables and social variables have an impact on the electricity demand series, and therefore they fed these inputs to develop their models. In some cases, although some selected variables may inevitably lead to forecasting errors, the variables were selected based on human intuition and expertise in this subject area or based on validation results of one-to-one correlations between the variables and electricity demand series (Islam et al., 1995; Al-Zayer and Al-Ibrahim, 1996; Islam and Al-Alawi, 1997; Sailor and Munoz, 1997; Lam, 1998; Hor et al., 2005; Lam et al., 2008; Wangpattarapong et al., 2008; Apadula et al., 2012). None have examined the method that can develop a model

Please cite this article in press as: Son, H., Kim, C., Short-term forecasting of electricity demand for the residential sector using weather and social variables. Resour Conserv Recy (2016), http://dx.doi.org/10.1016/j.resconrec.2016.01.016