Applied Soft Computing 36 (2015) 422–441

Contents lists available at ScienceDirect

Applied Soft Computing

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

Pattern similarity-based methods for short-term load forecasting – Part 2: Models

Grzegorz Dudek*

Department of Electrical Engineering, Czestochowa University of Technology, Al. Armii Krajowej 17, 42-200 Czestochowa, Poland

ARTICLE INFO

Article history:
Received 9 December 2014
Received in revised form 16 June 2015
Accepted 27 July 2015
Available online 5 August 2015

Keywords:
Artificial immune systems
Nonparametric regression
Short-term load forecasting
Similarity-based methods

ABSTRACT

Models for the short-term load forecasting based on the similarity of patterns of seasonal cycles are presented. They include: kernel estimation-based model, nearest neighbor estimation-based models and pattern clustering-based models such as classical clustering methods and new artificial immune systems. The problem of construction of the pattern similarity-based forecasting models and the elements and procedures of the model space are characterized. Details of the model learning and optimization using deterministic and stochastic methods such as evolutionary algorithms and tournament searching are described. Sensitivities of the models to changes in parameter values and their robustness to noisy and missing data are examined. The comparative studies with other popular forecasting methods such as ARIMA, exponential smoothing and neural networks are performed. The advantages of the proposed models are their simplicity and a small number of parameters to be estimated, which implies simple optimization procedures. The models can successfully deal with missing data. The increased number of the model outputs does not complicate their structure. The local nature of the models leads to their simplification and accuracy improvement. The proposed models are strong competitors for other popular univariate methods, which was confirmed in the simulation studies.

1. Introduction

The importance of the short-term load forecasting (STLF) in the power system control, scheduling and security translates into a large number of forecasting models. In the last few decades various forecasting methods have been proposed. They can be generally divided into conventional and unconventional methods. Conventional STLF models use regression methods, smoothing techniques and statistical analysis. Regression methods, linear and nonlinear, parametric or nonparametric, are usually applied to model the relationship between load consumption and other factors (weather, day type, customer class). Examples of semi-parametric additive models were recently presented in [1], whilst the nonparametric model using kernel estimators was presented in [2].

Gross and Galiana in their review paper [3] consider two basic conventional STLF models: time-of-day models and dynamic models. The former defines the load as a linear combination of a finite number of explicit time function, usually sinusoids with a period of 24 or 168 h. The latter take into account the most recent behavior of the time series and also exogenous variables and random component. Dynamic models are of two basic types: autoregressive moving average (ARMA) and state-space models. These approaches are used successfully up to today. Some examples such as ARIMA, exponential smoothing and the structural time series models are presented in the first part of this work [4]. Nowadays the conventional methods are often hybridized with new computational intelligence methods. As an example a new self-organizing model of fuzzy autoregressive moving average with exogenous input variables proposed in [5] can be given. In this approach a combined use of heuristics and evolutionary programming scheme is relied on to solve the problem of determining optimal number of input variables, best partition of fuzzy spaces and associated fuzzy membership functions. Good overview of the autoregressive moving average and other statistical approaches to modeling and forecasting electricity loads and prices can be found in [6]. Some conventional approaches to load forecasting such as static and dynamic state estimation are described in book [7].

The rapid development of computational intelligence observed in recent years has brought new methods of STLF. They are based on artificial neural networks (ANNs), fuzzy logic and expert systems. Also intelligent searching methods, such as evolutionary algorithms and swarm intelligence, are often applied to optimize the STLF models.

The multilayer perceptron (MLP), ANN which is most often applied in load forecasting, is an attractive tool to modeling of nonlinear problems due to their universal approximation property.

* Tel.: +48 343250896; fax: +48 343250803.
E-mail address: dudek@el.pcz.czest.pl

http://dx.doi.org/10.1016/j.asoc.2015.07.035
1568-4946/

Its other useful properties are: massive parallelism among a large number of simple units, learning capabilities, robustness in the presence of noise, and fault tolerance. Many forecasting models based on the MLP are used in practice by electric companies. An example would be ANNSTF system, which uses more than 40 companies from the U.S. and Canada [8]. Examples of some new publications on the use of MLP in STLF are: [9], where the complexity of MLP applied to STLF problems has been controlled by the Bayesian approach [10], where a new hybrid forecasting method composed of wavelet transform, MLP and evolutionary algorithm is proposed [11], where a generic framework that combines similar day selection, wavelet decomposition, and MLP is presented [12], where MLP is combined with wavelet transform and particle swarm optimization [13], where an approach of MLP with rough sets for complicated STLF with dynamic and non-linear factors is proposed, and [14], where the neural model generates the prediction intervals.

A radial basis function (RBF) network is an alternative to MLP in STLF. The RBF network approximates the target function by a linear combination of radial functions (usually Gaussian), which nonlinearly transform the input data. The learning algorithms for RBF are simpler than for MLP. The RBF network has a property of universal approximation. Some new publications concerning the STLF models based on the RBF network are: [15], where RBF is combined with fuzzy inference system and genetic algorithm [16], where a model to STLF is established by combining the RBF network with the adaptive neural fuzzy inference system and [17], where RBF is combined with the wavelet transform.

A self-organizing feature map (SOFM) is an another ANN used in STLF. This network is trained using unsupervised competitive learning to produce a low-dimensional representation of the input pattern space. The input patterns are grouped and represented by neurons. Some examples of STLF models using SOFM are: [18], where a hierarchical model composed of two SOFM is presented [19], where an adaptive two-stage hybrid network with SOFM and support vector machine is proposed [20], where SOFM is combined with MLP and a flexible smooth transition autoregressive model, and [21], where nonlinear model based on SOFM and predictors determined using curvilinear component analysis is described.

Many other types of neural networks have been used for STLF including: recurrent networks, generalized regression ANN [22], probabilistic ANN, adaptive resonance theory ANN, functional link network and counter propagation ANN. The survey of ANN applications to STLF can be found in [23] and [24].

Fuzzy logic allows to take into account imprecise, incomplete and ambiguous information in the STLF models. Fuzzy models are often simpler and more accurate than standard statistical models and allow to enter input information by rules formulated verbally by experts. The advantage of fuzzy inference systems is that they describe the behavior of complex systems by using linguistic expressions, mimicking the action of man. The fuzzy rule base consists of if-then statements that are almost natural language. To obtain a set of if-then rules two approaches are used. First, transforming human expert knowledge and experience, and second, automatically generating the rules from examples. The fusion of neural networks and fuzzy logic in neuro-fuzzy models achieves readability and learning ability (extracting rules from data) at once. The fuzzy inference mechanism leads to a nonlinear global model, which is an interpolation of local models implemented in the individual rules. The fuzzy STLF models are based on: fuzzy interpolation [25], fuzzy linear regression [26], fuzzy C-regression [27], Takagi–Sugeno–Kang model [28], fuzzy inductive reasoning [29] and neuro-fuzzy networks. The main advantages of the latter hybrid approach are: the ability to respond accurately to unexpected changes in the input variables, the ability to learn from experience, and the ability to synthesize new relationships between the load demand and the input variables. Examples of such STLF models are: [16], where the neuro-fuzzy system is used to adjust the results of load forecasting obtained by RBF network [30], where two neuro-fuzzy networks are proposed: a wavelet fuzzy neural network using the fuzzified wavelet features as the inputs and fuzzy neural network employing the Choquet integral as the outputs [31], where an efficient adaptive fuzzy neural network is proposed which can reduce its complexity removing the unneeded hidden units [32], where an integrated approach which combines a self-organizing fuzzy neural network learning method with a bilevel optimization method [33] where a neuro-fuzzy system working on the seasonal cycle patterns is proposed, and [34], where fuzzy logic is combined with wavelet transform and neural network.

Another useful tools for STLF are: support vector machines [35–37], clustering methods [19,20,38] and ensembles of models [39–41]. An interesting approach which can be classified as the similarity-based one is presented in [42]. It uses the clustering of the normalized daily curves for definition new inputs: sequences of the group labels for the successive days. The sequences are paired with load curves of the next days. The forecast is composed from the daily curves paired with the sequences from the history which are the same as the current sequence. Another interesting similarity-based model for STLF is described in [43]. The forecast is calculated as the weighted average of past daily load segments, the shape of which is similar to the expected shape of the load segment to be predicted.

It is noteworthy that many of the models developed in recent years are hybrid solutions (most papers concerning STLF published in IEEE Transaction on Power Systems in the last 10 years relate to just such models). These approaches combine data preprocessing methods (e.g. wavelet transform) with approximation models (such as neural and neuro-fuzzy networks) and methods of optimization and learning of these models (e.g. evolutionary and swarm algorithms). Sometimes forecast is adjusted depending on additional factors affecting the load demand and not included in the basic model.

This paper presents the univariate STLF models based on similarities between patterns of the daily cycles of the load time series. The principles of the models were described in the first part of this work [4]. The main advantage of the pattern similarity-based forecasting models (PSBFMs) is their simplicity: they have a clear structure and comprehensible principles of operation. The number of parameters is low and the optimization and learning procedures are fast.

The remainder of this paper is divided into seven sections. The problems of construction of the PSBFMs in Section 2 are presented. In Sections 3–5 the STLF models based on pattern similarity including non-parametric regression methods and clustering methods are presented. In Section 6 we analyze the proposed forecasting methods and we compare the results to other STLF methods: ARIMA, exponential smoothing and MLP. An overview of the work is given in Section 7.

2. Construction of PSBFMs

The forecasting models considered in this work are memory-based inductive approximation models which learn the relevant relationships between variables on the basis of observed instances. Instances (examples or samples) are pairs of x- and y-patterns extracted from the load time series using the functions f_x and f_y (see Section 3 in [4]). Instances form the series $S = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_N, \mathbf{y}_N)\}$, $i = 1, 2, \ldots, N$, $\mathbf{x}_i \in X = \mathbb{R}^n$, $\mathbf{y}_i \in Y = \mathbb{R}^n$, where X and Y are domain and codomain, respectively. The goal