Energy 59 (2013) 126–132

Contents lists available at ScienceDirect

Energy

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

Online training algorithms based single multiplicative neuron model for energy consumption forecasting

Xuedong Wu [a,b,*], Jianxu Mao [c], Zhaoping Du [a], Yanchao Chang [a]

[a] School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212003, China
[b] Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, USA
[c] College of Electrical and Information Engineering, Hunan University, Changsha, Hunan 410082, China

ARTICLE INFO

Article history:
Received 8 October 2012
Received in revised form
12 May 2013
Accepted 30 June 2013
Available online 26 July 2013

Keywords:
Energy consumption forecasting
Small dataset and nonlinearity
Single multiplicative neuron model
Iterated nonlinear filters

ABSTRACT

Although traditional approaches can yield accurate forecasting results of energy consumption, they may suffer from several limitations such as the need for large dataset and the linear assumption. Two novel hybrid dynamic approaches, which are based on the SMN (single multiplicative neuron) model and the iterated nonlinear filters, have been proposed for forecasting energy consumption with small dataset and nonlinearity in our study. The forecasting models are established by using the weights and the biases of SMN model to present the state vector and the output of SMN model to present the observation equation, and the input vector to the SMN model is composed of the known energy consumption values with a rolling mechanism. The SMN model has advantages of better approximation capabilities, simpler network structures and faster learning algorithms. The nonlinear filters can deal with additive noises and can update model parameters when a new observation data arrives due to their iterative algorithm structure. Two case studies of energy consumption have been used to demonstrate the reliability of the proposed models, and the experimental results have indicated that the proposed approaches outperform existing models in forecasting energy consumption.

1. Introduction

Energy consumption has been increased remarkably over the past decades all over the world due to the increasing population and the economic development [1,2]. Energy is considered as an important factor in the economic and social development of a country and consequently in the people's wealth. Therefore, many countries are concerned with energy–related issues. For instance, China is the country that products and consumes the most energy in Asia, the China statistical yearbook [3] states that the rate of energy production increases 4.5% annually while the rate of energy consumption increases 5.2% annually. In addition, the energy information administration of the United States has forecasted that global energy consumption will increase by 49% from 2007 to 2035 [4]. Given this fact, a highly precise model for forecasting energy consumption must be developed in order that energy policy makers can either implement an energy conservation policy or allocate a certain amount of energy based on such a model.

The ARIMA (autoregressive integrated moving average) model is extensively used to forecast time series data [5]. However, the forecasting accuracy of the ARIMA model is poor when data are few or nonlinear [6]. Forecasting models that are based on conventional statistical methods are limited because real world data are commonly few, nonlinear and fail to satisfy statistical assumptions.

To solve the forecasting problem of time series with few and nonlinear observations, the grey system theory based first–order one–variable grey model (GM(1,1)) had been adopted for energy consumption forecasting in Refs. [7,8]. Lee and Tong [7] combined residual modification with residual GP (genetic programming) sign estimation to increase the effectiveness of ANN (artificial neural network) in estimating the residual signs of GM(1,1). To increase the accuracy of GM(1,1) applied to original energy consumption data and to prevent inaccurate forecasting using conventional linear time series models when residual series were complex patterns, Lee and Tong [8] developed a novel hybrid dynamic approach that combined a dynamic grey model with GP to forecast energy consumption.

Forecasting models can also be developed using conventional ANNs [1,9–16]. In Ekonomou [1], a developed ANN by testing several possible architectures and selecting the one with the best

* Corresponding author. School of Electronics and Information, Jiangsu University of Science and Technology, No. 2, Mengxi Road, Zhenjiang, Jiangsu 212003, China.
E-mail address: woolcn@163.com (X. Wu).

0360-5442/$ – see front matter
http://dx.doi.org/10.1016/j.energy.2013.06.068

generalizing ability was suggested for Greek long–term energy consumption prediction. Azadeh et al. [9] presented an approach by combining GA (genetic algorithm) with ANN to estimate and predict electricity demand using stochastic procedures. Three modeling techniques, which included traditional regression analysis, decision tree and ANN, were considered for the prediction of electricity energy consumption in Ref. [10]. An ANN model was used to forecast building energy consumption in Ref. [11]. González and Zamarreño [12] suggested an approach for short–term load prediction using a feedback ANN. Dombaycı [13] used a three–layer ANN to predict hourly energy consumption of a modeled house in Denizli–Turkey. A comparative study between a hybrid genetic algorithm–adaptive network–based fuzzy inference system (GA–ANFIS) method and an ANN approach for building energy consumption forecasting was presented in Ref. [14]. Azadeh et al. [15] developed an ANN approach for annual electricity consumption in high energy consumption industrial sectors. In Refs. [16], a back–propagation three–layered ANN was used for the prediction of the heating energy requirements of different building samples.

However, the development of conventional ANNs for time series prediction involves many difficulties such as the selection of inputs to network, the selection of the network structure and the calculation of model parameters [17]. Recently, the single multiplicative neuron (SMN) model has been proposed as an alternative to conventional ANNs. The SMN model derives its inspiration from the single neuron computation in neuroscience [18,19], and only the input connections need to be determined during the learning process. Therefore, the SMN model is much simpler in structure and lower in computational complexity than the conventional ANNs and can offer better performances if properly trained [20,21], and it has been successfully applied to different time–series prediction [22–24]. Compared with conventional ANNs, although the SMN model has advantages of better approximation capabilities, simpler network structures and faster learning algorithms, the development of SMN model may suffer from the basic limitation on estimation of the model parameters in the training stage [24], which is similar to conventional ANNs.

Nonlinear filtering algorithms can deal with nonlinear systems, and the EKF (extended Kalman filter) is a well–known approach in the integration of nonlinear systems. Although the EKF has the advantage in real–time estimation, the linearization of a nonlinear system by Taylor series expansion may be a biased estimator [25,26]. The UKF (unscented Kalman filter) and the IEKF (iterated extended Kalman filter) are the representative filters in overcoming the flaw of the EKF. The UKF essentially provides derivative–free higher order approximations by approximating a Gaussian distribution [25], and the IEKF can effectively reduce the bias and the estimation error by increasing iterative operations [27,28]. Enlightened by the development of IEKF as well as the superiority of UKF, a recently developed filtering technique called IUKF (iterated unscented Kalman filter) is proposed. This algorithm is developed from UKF but it can obtain more accurate state and covariance estimation and has been used in different fields such as [29,30].

In the existing literature, ANNs are normally constructed using an entire dataset in the training stage. This study has developed the IEKF–based–SMN model and the IUKF–based–SMN model with a rolling mechanism in which only a minimal amount of recent data are used for energy consumption forecasting. The system dynamic model is established by using the weights and the biases of SMN model to present the state vector and the output of SMN model to present the observation equation, and the forecasting observation values of the nonlinear filters are the energy consumption forecasting values. The input vector to the SMN model is composed of the continuous known energy consumption data and its dimension is decided according to their minimum MSE (mean squared error) values because the nonlinear filters are based on the minimum MSE principle.

The rest of this paper is organized as follows. Section 2 demonstrates the proposed approaches including the iterated nonlinear filters in Section 2.1, the SMN model in Section 2.2 and the iterated nonlinear filters based SMN model for energy consumption forecasting in Section 2.3. Section 3 evaluates the forecasting accuracy of the proposed models and compares them with the existing energy consumption models. The reasons for obtaining the experimental results are discussed in detail as well. Section 4 concludes the paper.

2. The proposed approaches

2.1. The iterated nonlinear filters

Suppose that a state vector $\mathbf{x}_k$ at instant k is propagated through the following nonlinear state transition equation,

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k) + \mathbf{v}_k, \tag{1}$$

where $f(\bullet)$ is the state transition function and $\mathbf{v}_k$ is the process noise with covariance matrices $\mathbf{Q}_k$. The observation equation is given as,

$$\mathbf{y}_k = h(\boldsymbol{x}_k, \boldsymbol{u}_k) + \mathbf{w}_k, \tag{2}$$

where $h(\bullet,\bullet)$ is the observation equation, $\mathbf{u}_k$ is the known input vector and $\mathbf{w}_k$ denotes the observation noise with covariance matrices $\mathbf{R}_k$.

The following Section 2.1.1 and 2.1.2 briefly describe the IEKF and the IUKF, and the reader can refer to literature [27,29] for more detailed descriptions on them.

2.1.1. The iterated extended Kalman filter

The main difference between the IEKF and the EKF lies in the step of observation update [27]. The update equations of IEKF for mean and covariance of the Gaussian approximation to the posterior distribution of the state are as follows.

(1) Initialized with $\widehat{\mathbf{x}}_0 = E[\overline{\mathbf{x}}_0]$, $\widehat{\mathbf{P}}_0 = E[(\mathbf{x}_0 - \overline{\mathbf{x}}_0)(\mathbf{x}_0 - \overline{\mathbf{x}}_0)^T]$.
(2) For $k = 1,2,\ldots$
  (a) Time update.

$$\overline{\boldsymbol{x}}_k = f\left(\widehat{\boldsymbol{x}}_{k-1}\right), \quad \overline{\boldsymbol{P}}_k = \boldsymbol{F}_k \widehat{\boldsymbol{P}}_{k-1} \boldsymbol{F}_k^{\mathrm{T}} + \mathbf{Q}_k, \tag{3}$$

where $\mathbf{F}_k = \frac{\partial f}{\partial \mathbf{x}}|.\mathbf{x} = \widehat{\mathbf{x}}_{k-1}$ is the Jacobian matrix of state transition equation.
  (b) Observation update.

After the state prediction value $\overline{\mathbf{x}}_k$ and the corresponding covariance $\overline{\mathbf{P}}_k$ are obtained, the following iteration will be recursively carried out, for $j = 0,1,2,\ldots,N$

$$\overline{\boldsymbol{x}}_{k,0} = \overline{\boldsymbol{x}}_k, \quad \overline{\boldsymbol{P}}_{k,0} = \overline{P}_k, \tag{4}$$

$$\boldsymbol{H}_{k,j} = \frac{\partial h}{\partial x}\Big|\boldsymbol{x} = \overline{\boldsymbol{x}}_{k,j}, \quad \boldsymbol{K}_{k,j} = \overline{\boldsymbol{P}}_k \boldsymbol{H}_{k,j}^T \left[\boldsymbol{H}_{k,j} \overline{\boldsymbol{P}}_k \boldsymbol{H}_{k,j}^T + \boldsymbol{R}_k\right]^1, \tag{5}$$

$$\widehat{\boldsymbol{x}}_{k,j} = \overline{\boldsymbol{x}}_k + \boldsymbol{K}_{k,j}\left[\boldsymbol{y}_k - h\left(\overline{\boldsymbol{x}}_{k,j}, \boldsymbol{u}_k\right)\right], \widehat{\boldsymbol{P}}_{k,j} = \overline{\boldsymbol{P}}_k - \boldsymbol{K}_{k,j} \boldsymbol{H}_{k,j} \overline{P}_k. \tag{6}$$

where j is the iterated number for the same observation value and $\mathbf{K}_{k,j}$ is the Kalman gain.