CrossMark

METHODOLOGIES AND APPLICATION

Optimal design of adaptive neuro-fuzzy inference system using genetic algorithm for electricity demand forecasting in Iranian industry

Shahram Mollaiy-Berneti[1]

Abstract Demand planning for industrial's electricity consumption is an important factor to efficiently plan the generation and distribution of power utilities. However, this can only be possible if the demand is predicted accurately. Recent advancement in adaptive neuro-fuzzy inference system aimed at mapping input to output for highly non-linear processes such as energy management field, provide reliable approach to forecast energy demand. Despite the wide range of applications and flexibility of adaptive neuro-fuzzy inference system, complexity of the rule base is featured with certain limitations associated with combinatorial explosion of rules, parameters and data. This paper proposes a hybrid procedure, subtractive clustering technique coupled with genetic algorithm, to develop adaptive neuro-fuzzy inference system. Genetic algorithm finds the optimum value of cluster radius which guaranteed the minimum number of rules and error. The empirical data regarding the industrial's electricity demand in Iran from 1967 to 2011 are investigated to demonstrate the applicability and merits of the present method. The performance of hybrid approach is found to be better than that of conventional adaptive neuro-fuzzy inference system based on gird partitioning, fuzzy c-means, and subtractive clustering in terms of both accuracy and the number of rules.

Keywords Adaptive neuro-fuzzy inference system · Subtractive clustering · Genetic algorithm · Electricity demand · Iranian industrial sector

✉ Shahram Mollaiy-Berneti
  sh.m.berneti@gmail.com

[1] Young Researchers and Elites Club, Bandar Abbas Branch, Islamic Azad University, Bandar Abbas, Iran

1 Introduction

Electricity is one of the most significant input for industrial sector. It is used both in the building components for cooling, heating, and lighting, and in the operational process for mechanical and electronic processing. Demand for electricity is increasing with increasing industrial activities.

In recent years, industrial's electricity consumption in Iran has been growing very rapidly. In 2011, electricity consumption in industrial sector occupied 34.6 % of total electricity sale by ministry of energy, increased by 3.9 % compared to 2010 (Ministry of Energy 2013). In the same year, the number of customers in industrial sector has increased to 1,74,000 (Tavanir holding company 2012). Some factors including low price of electric energy, low efficiencies and old technologies of equipment cause a large amount of waste and inefficiency in electricity consumption.

Forecasting electricity consumption in industries has always been an essential part of efficient power system planning and operation. There are different techniques that contribute to the prediction of future electricity demand. Time series, regression, support vector machine, and various types of artificial neural network (ANN) are samples of methods that have been applied in this area (Kumar and Jain 2010; Al-Shobaki and Mohsen 2008; Zhang et al. 2009; Limanond et al. 2011; Hong 2009; Ekonomou 2010; Fadare 2009; Geem and Roper 2009; Xia et al. 2010).

One popular and robust soft computing method is adaptive neuro-fuzzy inference system (ANFIS) which has advantages of the simple learning procedures and computational power of neural networks and the high-level, human-like thinking, and reasoning of fuzzy systems. ANFIS relates input space to output space based both on human knowledge and on datasets through a hybrid learning scheme combining the gradient descent technique and least-squares method.

Ying and Pan (2008) applied ANFIS model to forecast the regional electricity loads in Taiwan. Zhang and Mao (2009) investigated the interrelationship between energy consumption and economic growth in China based on ANFIS model. Cheng and Wei (2010) used ANFIS for electricity load forecasting and their experimental results indicated that the ANFIS model is superior to ANN modeling. Long-term electric load forecasting has been predicted using ANFIS model (Akdemir and Çetinkaya 2012). Using ANFIS, Al-Ghandoor et al. (2012) modeled and forecasted the transport energy demand of Jordan. Cárdenas et al. (2012) presented a Load forecasting framework of electricity consumptions based on ANFIS. Zahedi et al. (2013) developed ANFIS method to predict electricity demand in Ontario province of Canada up to 2015.

Generating a fuzzy inference system (FIS) with the minimum number of rules is an important issue in developing an ANFIS, because the total number of ANFIS parameters that need to be adjusted dramatically increases with an increase in the number of rules. For FIS generation, several methods like grid partitioning, fuzzy c-means (Bezdek 1981), mountain (Yager and Filev 1994), and subtractive clustering (Chiu 1994) have been proposed.

In grid partitioning, all the possible rules are generated based on the number of membership functions (MFs) for each input. The wide application of this strategy is blocked by the curse of dimensionality, which means that the number of fuzzy rules increases exponentially with the number of input variables. For example, considering a problem with five input variables each of which has four MFs, the total number of possible rules in grid partitioning will be 1024. This algorithm is only suitable for cases with small number of input variables.

The fuzzy c-means (FCM) algorithm is a method of clustering which allows one piece of data to belong to two or more clusters. This algorithm requires a priori knowledge of the number of clusters. Thereupon, the number of rules must be known in advance, because each cluster is assigned to a specific rule. Another major drawback of FCM is the initial guess positions for each cluster center. The random initialization may give sub-optimal results, especially when complex data structures are present or the data distribution has a large variance.

The automatic generation capability for estimating the number and initial position of cluster centers was introduced in the mountain method. This method operates based on gridding the data space and computing a potential value for each grid point which is a function of the distance measure. In this method, a grid point with highest potential is considered as a cluster center and gird points close to new cluster center are penalized to control the emergence of new cluster centers. This method is simple and effective. However, its computation grows exponentially with the dimension of the problem.

Subtractive clustering (SC) is an extension of the mountain method that eliminates the restriction imposed by this method. It clusters data points through calculating the potential of data points in the feature space. Then, the data point which has less distance than the cluster radius from the cluster center is subtracted and the potential values are iteratively updated. The method continues until all the datasets are tested. Since in this method each data point (not a grid point) is considered as a potential cluster center, it is independent of the dimension of the problem and no longer necessary to specify a grid resolution.

The selection of an appropriate cluster radius is an important initial consideration in SC technique, because it has a significant impact on the complexity and generalization capabilities of this method. A large radius generally leads to fewer clusters (and hence fewer rules), while a small radius produces many smaller clusters (hence excessive number of rules). Unfortunately, there is no standard procedure for choosing this parameter. The suitable influence radius is difficult to be found and requires a large number of trials to obtain the appropriate value.

Due to the capability of irregular searching, evolutionary algorithms such as genetic algorithm (GA) can be used for determination of optimum value of cluster radius. These algorithms perform a global exploration of the search space and use several heuristic methods to avoid getting trapped in local optima. In the past few years, the evolutionary algorithms have been greatly utilized in ANFIS modeling (Aliyari et al. 2007; Chen et al. 2007; Ho et al. 2009; Oliveira and Schirru 2009; Mollaiy-Berneti 2013).

The objective of this study is to develop an ANFIS automatically from numerical data with appropriate fuzzy rules, which offers a good tradeoff between the size of the rules base (complexity) and the accuracy of the model. Therefore, subtractive clustering is applied to determine the initial FIS. Subsequently, hybrid learning method is used to tune this initial FIS. To achieve the good complexity–accuracy tradeoff in the design of ANFIS, GA is coupled with SC to estimate the optimum value of cluster radius. To verify the applicability of the proposed model, it is used to predict industrial's electricity consumption in Iran utilizing socioeconomic indicators based on population, gross domestic product (GDP), number of customers, electricity price, and import and export.

The rest of the paper is organized as follows. Section 2 briefly describes the structure of ANFIS and presents the procedure of subtractive clustering and genetic algorithm for the model construction and optimization. The last part of this section has been devoted to describing the methodology to integrate the SC technique and GA for automatic generation of the initial fuzzy inference system structure. In Sect. 3, the numerical results from a real-world case study are presented and discussed. Finally, concluding remarks are given in Sect. 4.

2 Methodology

2.1 Adaptive neuro-fuzzy inference system

ANFIS is a class of adaptive multi-layer feed-forward networks, which consists of fuzzy layer, product layer, normalized layer, de-fuzzy layer, and summation layer. Figure 1 shows the architecture of a typical ANFIS with two inputs, two rules and one output for the first-order Sugeno fuzzy model, where each input is assumed to have two associated MFs. Fuzzy rule set with two fuzzy if–then rules is as follows:

Rule 1: if x is A_1 and y is B_1 then $z_1 = p_1 x + q_1 y + r_1$
Rule 2: if x is A_2 and y is B_2 then $z_2 = p_2 x + q_2 y + r_2$

As it is shown in Fig. 1, different layers of ANFIS have different nodes. The square and circle nodes are for adaptive nodes with parameters and fixed nodes without parameters, respectively. Different layers with their associated nodes are described below:

Layer 1 It consists of square nodes that make the membership grade for the input variables. The parameters in this layer are called premise parameters.
Layer 2 Every node in this layer calculates the firing strength of a rule via multiplying the incoming signals from the previous layer.
Layer 3 The ratio of the ith rule firing strength to the sum of all rules' firing strength is calculated in the third layer, generating the normalized firing strengths.

Layer 4 Each node in this layer performs multiplication of normalized firing strengths with the corresponding rule. The parameters in this layer are called consequent parameters.
Layer 5 The single node in this layer computes the overall output as the summation of all incoming signals.

The learning algorithm for ANFIS is a hybrid algorithm, the gradient descent and the least-squares method. The least-squares method is used to identify the consequent linear parameters through the training set, while gradient descent method is employed to update the premise non-linear parameters, through minimizing the overall quadratic cost function. The detailed algorithm and mathematical background of the hybrid-learning algorithm can be found in Jang (1993).

2.2 Subtractive clustering technique

Before start the ANFIS training, ANFIS needs to define the antecedent MFs and fuzzy rules. Due to automatic generation capability, it is possible to use subtractive clustering technique for making such decisions. An important advantage of using a SC technique to find rules is that the resultant rules are more tailored to the input data than they are in a FIS generated without clustering. This reduces the problem of combinatorial explosion of rules when the input data have a high dimension (Jang and Gulley 2000).

SC estimates the number of clusters and initial location of cluster centers in a set of data, and extracts the fuzzy rules by projection of the clusters onto the input space. This
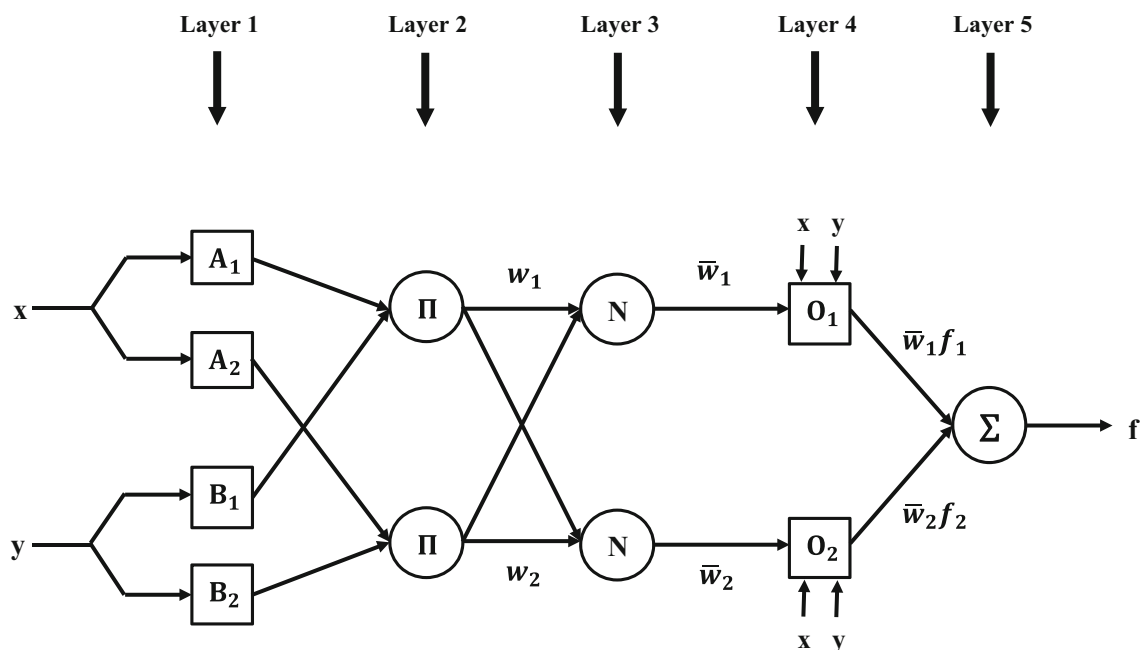

Fig. 1 The ANFIS architecture for a two-input T–S model with two rules

method operates by finding the point with the highest number of neighbors as center for a cluster based on the density of surrounding data points (Chiu 1994). A brief description of subtractive clustering method is as follows.

Consider a collection of m data points $\{x_1, x_2, \ldots, x_m\}$ in an N-dimensional space. Without loss of generality, all data points are assumed normalized in each dimension. Since each data point is a potential cluster center, the potential value for each data point is defined as

$$p_i = \sum_{j=1}^{m} \exp\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right), \tag{1}$$

where $\|\cdot\|$ denotes the Euclidean distance and r_a is a positive constant defining a normalized neighborhood data radius for each cluster. Data outside this radius have little influence on the potential. After computing the potential for each data point, the data point with the highest potential is chosen as the first cluster center. Assume x_1^* is the location of the first cluster center, and p_1^* its potential value. The potential of the remaining data point x_i^* is revised by

$$p_i \Leftarrow p_i - p_1^* \exp\left(-\frac{\|x_i - x_1^*\|^2}{(r_b/2)^2}\right), \tag{2}$$

where

$$r_b = \eta r_a, \tag{3}$$

where $\eta > 1$ is called the squash factor. The constant r_b defines a neighborhood that has measurable reductions in potential. Thus, the potential of each data point is reduced according to its distance from the first cluster center. The data points near the first cluster center will have significantly reduced potential, and consequently will have a lower chance of being selected as the next cluster center. After revising the potential calculation for each data point, the second cluster center is selected as the point with highest remaining potential. The process is then continued until a sufficient number of clusters are generated. In general, after the kth cluster center has been obtained, the potential of each data point is revised as follows:

$$p_i \Leftarrow p_i - p_k^* \exp\left(-\frac{\|x_i - x_1^*\|^2}{(r_b/2)^2}\right), \tag{4}$$

where x_k^* is the location of the kth cluster center and p_k^* is its potential value. After applying subtractive clustering, a sufficient number of clusters are identified. Each cluster represents a rule. The rules generated by this method are then optimized using ANFIS methodology, which uses hybrid-learning algorithm described in previous section.

2.3 Genetic algorithm

GA is a well-known and effective type of evolutionary computation technique working based on the idea of "survival of the fittest" and "natural selection". GA starts with a population of randomly generated solutions, codified by chromosomes. The chromosomes are represented as a vector of real numbers of which every entry includes one of the unknown parameters of the problem. Each chromosome in the population undergoes reproduction, crossover, and mutation, in each iteration, to produce a new population with the purpose of generating chromosomes that map to higher level of fitness.

Reproduction which is also called selection implies that the chromosome associated with fittest individuals can participate more than once, while less fit chromosomes may be completely suppressed leading to an increase in average fitness of the population.

Crossover is the process of exchanging the information carried by two chromosomes to produce new chromosomes for the next generation. Two chromosomes are selected from the mating pool randomly and named parents. The new chromosomes produced are entered to offspring pool. The offsprings may represent an unexplored point in the search space.

Mutation is a process of irregular alternation of some parts in a chromosome in the offspring pool by a random change. The purpose of mutation is (1) to perform a local search around the current solution, (2) to prevent a premature loss of important genetic material at a particular position, and (3) to maintain diversity in the population. For more information on GA details the interested readers are referred to DeJong (1988), Goldberg (1989), and Fogel (1995).

2.4 Proposed model

In a standard ANFIS, a grid-based definition of fuzzy terms and their membership functions is usually used to partition the input space into many local regions. All possible combinations of these fuzzy terms and their membership functions form the fuzzy rule base. As standard ANFIS has been applied to more complicated and high-dimensional systems, it invokes the problem of combinatorial explosion of rules, parameters, and data, namely the so-called curse of dimensionality (Liu and Li 2004).

In addition, this curse of dimensionality results in several negative consequences:

- Transparency and interpretability are reduced as humans are incapable of understanding hundreds or thousands of fuzzy rules and parameters.
- Overfitting learning often occurs. Too many parameters increase the chance of overfitting if the number of para-

meters in the network is equal to or larger than the total number of samples in the training set (Demuth et al. 2007).

- The requirements for more computation and more memory become excessive.

To make ANFIS usable in dealing with complex systems, the subtractive clustering approach is introduced to reduce the complexity of rule base in the ANFIS. As described in Sect. 2.2, the SC technique partitions the input–output data into clusters, and adapts the architecture of the ANFIS with the minimum number of rules for distinguishing the fuzzy quantities with respect to each cluster. By clustering, both the number of linguistic values of each input and the number of rules are the same as the number of clusters, which reduces the parameters of the ANFIS remarkably.

When dealing with SC, there are several essential parameters, namely, accept ratio, reject ratio, cluster radius, and squash factor, whose values need to be decided upon in advance. The discussions on the effects of parameters of the SC on fuzzy model performance are given in Demirli et al. (2003). The performance of the model is very sensitive to the cluster radius while accept and reject ratio do not have big influence (Liu and Li 2004). Here the squash factor, accept and reject ratio were considered 1.25, 0.5, and 0.15, respectively (suggested in Chiu 1994).

The GA properties make it suitable to use to design fuzzy systems. In this study, the cluster radiuses are optimized using a GA and ANFIS is called within the GA for evaluating the objective value of any candidate solution generated by GA. The objective function of the GA is to maximize the accuracy of predictions made by ANFIS model with minimum complexity, whose number of rules is controlled by parameters of a SC technique varying in GA.

The proposed flow diagram is illustrated in Fig. 2. It starts by randomly generating an initial population of chromosomes each as a candidate solution representing different sets of cluster radius. Each chromosome is then feed to SC technique to cluster input–output data and identify FIS. Initial FIS is optimized by hybrid-learning method in ANFIS to determine its fitness. The fitness of each chromosome is evaluated using Schwarz–Rissanen criterion (SRC) (Rissanen 1978; Schwarz 1978; Wang and Yen 1999), which is represented by

$$SRC(m) = \log(\widehat{\sigma}_E^2) + \frac{\log(n)m}{n}, \tag{5}$$

where m denotes the number of fuzzy rules, n is the number of training data, and $\widehat{\sigma}_E^2$ is the estimated variance of model residuals. Chromosomes are evolved by the selection, crossover and mutation such that fitness value is minimized as much as possible. When the fitness value fulfills the end condition,

the searching process is stopped and final FIS with optimum parameters is obtained.

3 Results and discussion

To demonstrate the effectiveness and applicability of the proposed ANFIS model, industrial's electricity consumption in Iran was used as the case study. A statistical analysis was performed on available data to select significant input variables and prepare data for model implementation.

3.1 Available data

Six variables in the period from 1967 to 2011 were available. The following socioeconomic indicators were considered for analysis: population, GDP, number of customers, electricity price, and import and export. The first four variables were obtained from Ministry of Energy (2012), and two others were obtained from Central Bank of the Islamic Republic of Iran (2012).

3.2 Data preprocessing and analysis

To investigate the strength of relationships between the variables with electricity demand, Pearson correlation factor was used. This factor ranges between $+1$ and -1. The positive sign indicate the relationship between the variables is direct, but negative sign means indirect relationship between variables. The closer the absolute value is to 1, the stronger the correlation between the two variables.

Measure of the Pearson correlation factor between each six variables and electricity demand was evaluated. The results are given in Table 1. It has been inferred that all variables have the strongest relationship with electricity demand and can be used as inputs in prediction model of electricity demand.

To avoid the saturation problem and consequently the low rate training of the models, the scaling of data is needed. Scaling data concentrate the dispersed data to a defined interval, usually taken over the intervals [0, 1] or [−1, 1]. The following scaling function that converts the real input values to the corresponding values in the interval of [0.1, 0.9] was used in this study:

$$x_n = 0.1 + (0.9 - 0.1) \times (x_i - x_{\min})/(x_{\max} - x_{\min}), \tag{6}$$

where x_n is the normalize value of x_i, x_i is the variable of the database, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values in the database, respectively.

To avoid overfitting, the available dataset was randomly divided into training and validation sets Jang (1996). 80 % of the dataset was used for training process, during which
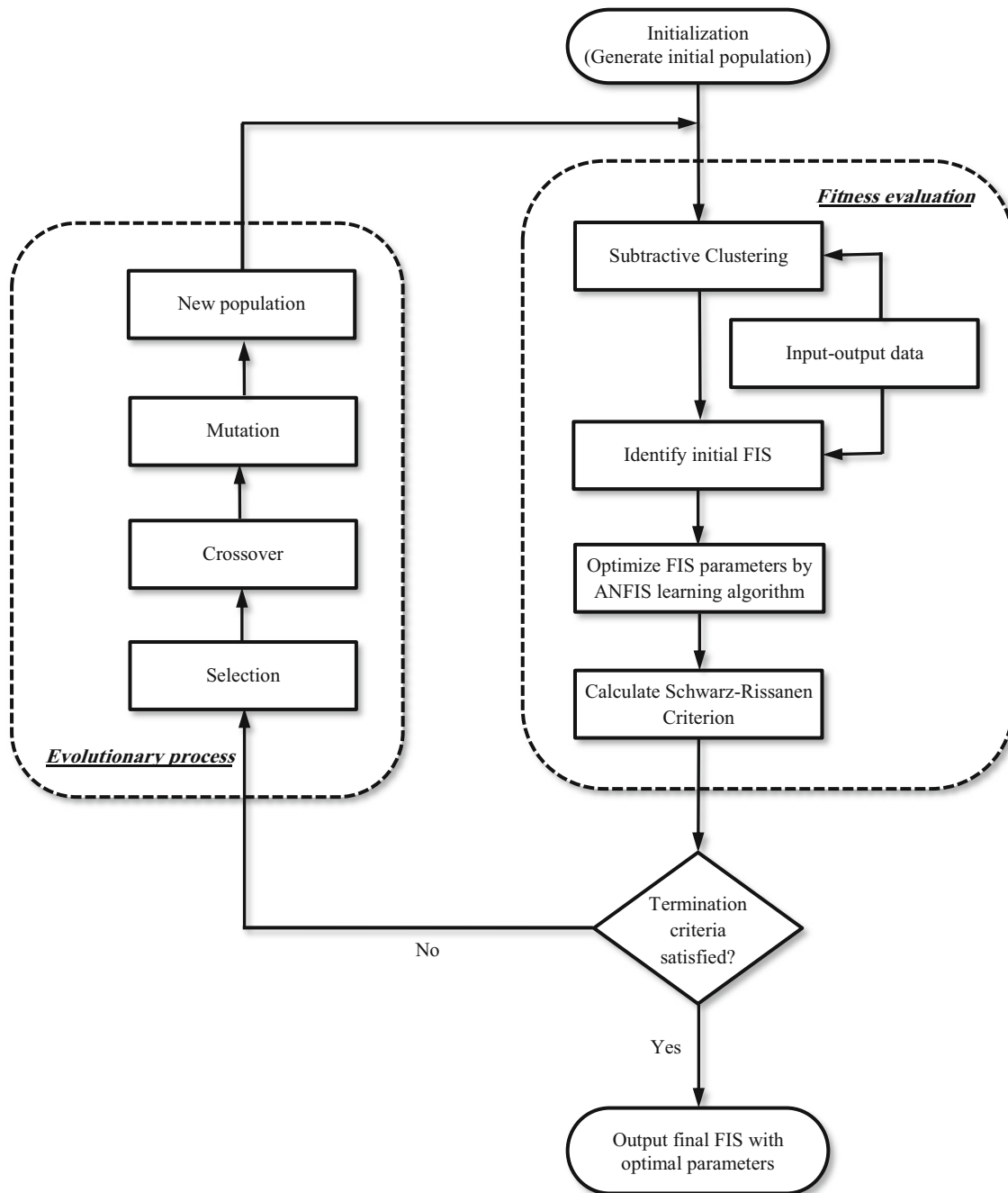
Fig. 2 The flow diagram of coupled SC-based ANFIS and GA

Table 1 Pearson correlation coefficient between electricity demand and input variables

Parameter	Pearson coefficient	Parameter	Pearson coefficient
Population	0.892	Electricity price	0.958
GDP	0.971	Import	0.909
Number of customer	0.984	Export	0.903

the fuzzy rules are generated and the MFs are tuned, while the remaining 20 % was used as validation data to check the generalization capability of the model.

3.3 Evaluation methods

The performance of the proposed model was assessed based on statistical parameters: mean absolute error (MAE),

Table 2 The rule base generated by proposed model

Rule no.	Rule
1	IF x_1 is A_{11} and x_2 is A_{12} and x_3 is A_{13} and x_4 is A_{14} and x_5 is A_{15} and x_6 is A_{16} THEN $y = 0.1745 + 0.0536\, x_1 + 0.0967\, x_2 - 5.0279\, x_3 + 0.1406\, x_4 + 0.5507\, x_5 + 0.5036\, x_6$
2	IF x_1 is A_{21} and x_2 is A_{22} and x_3 is A_{23} and x_4 is A_{24} and x_5 is A_{25} and x_6 is A_{26} THEN $y = 0.4279 + 2.0733\, x_1 - 0.7398\, x_2 + 0.6141\, x_3 - 0.7421\, x_4 - 0.082\, x_5 - 0.5109\, x_6$
3	IF x_1 is A_{31} and x_2 is A_{32} and x_3 is A_{33} and x_4 is A_{34} and x_5 is A_{35} and x_6 is A_{36} THEN $y = 1.5386 - 0.5426 x_1 + 0.1159\, x_2 + 0.5208\, x_3 + 0.2908\, x_4 - 0.0936\, x_5 - 0.6582\, x_6$

root-mean-square error (RMSE), and coefficient of determination (R^2). Expressions for these parameters are given as follows:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| y_i^p - y_i^a \right| \tag{7}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i^p - y_i^a)^2} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i^a - y_i^p)^2}{\sum_{i=1}^{N} (y_i^a - \bar{y}^a)^2}, \tag{9}$$

where N is the number of the training or validation samples, y_i^a, y_i^p, and $\bar{y}^a$ are the actual, predicted, and average of actual values, respectively.

3.4 Modeling and results

The ANFIS model was developed using fuzzy logic toolbox of MATLAB commercial software (2010). In the modeling process, identification method, MF type and learning technique were selected as subtractive clustering, Gaussian MFs, and hybrid learning, respectively. SC technique generates an initial FIS structure by making of clusters in the training data space and translation of these clusters into T–S rules. The identified initial first-order Sugeno fuzzy model was as follows:

R_1 : IF x_1 is A_{11} and x_2 is A_{12} and x_3 is A_{13} and x_4 is A_{14}

and x_5 is A_{15} and x_6 is A_{16}

THEN $y^1 = p_0^1 + p_1^1 x_1 + p_2^2 x_2 + p_3^3 x_3 + p_4^4 x_4 + p_5^5 x_5 + p_6^6 x_6$

.

.

.

R_k : IF x_1 is A_{k1} and x_2 is A_{k2} and x_3 is A_{k3}

and x_4 is A_{k4} and x_5 is A_{k5} and x_6 is A_{k6}

THEN $y^k = p_0^k + p_1^k x_1 + p_2^k x_2 + p_3^k x_3 + p_4^k x_4 + p_5^k x_5 + p_6^k x_6,$

Table 3 Cluster radius obtained by GA

|	x_1	x_2	x_3	x_4	x_5	x_6	y
r	0.8347	0.3358	0.9152	0.4403	0.8656	0.1520	0.8724

where x_1, x_2, x_3, x_4, x_5, and x_6 are population, GDP, number of customers, electricity price, and import and export, respectively. y^k is the consequent of the rule k, and p_0^1, p_1^1, p_2^2, p_3^3, p_4^4, p_5^5, and p_6^6 are the regression parameters. After setting the initial rule base, fine tuning of the parameters is required. Hybrid-learning algorithm with the combination of gradient descent technique and least-squares method was adopted in the fine tuning of the rules' premise MFs and consequent parameters, respectively.

To develop an appropriate SC-based ANFIS, GA with global searching capability was used to select optimum value of cluster radius for each dimension in the data space. The cluster radius is initialized as chromosome in GA, and then the global optimum is searched through the competition.

The evolutionary environments of the GA used in this study are population size was selected as 40, maximum iteration was set to 100, uniform crossover probability was set to 0.9 and uniform mutation probability was set to 0.01. These parameters, which are usually problem dependent, were selected based on the author's experience.

After applying the GA, high-efficient ANFIS that predicts the electricity demand without any loss of accuracy with only 3 rules was found. The obtained rule base is given in Table 2. Table 3 indicates the optimum cluster radius attained by GA. Figure 3 shows the projection MFs of this mechanism.

Table 4 provides a comparison analysis of the performance obtained by the proposed model with the respective performances of grid partitioning-based ANFIS, FCM-based ANFIS and SC-based ANFIS. Since the number of obtained rule by proposed model was three, number of cluster in FCM-based ANFIS was considered as three. As shown in this table, grid partitioning-based ANFIS is unable to perform mapping between input and output, because a large rule base was created in the FIS which makes MATLAB run out of memory, and it was obtained that prediction performance of the proposed model is better than the two other models, where both values of MAE and RMSE are smaller, coefficient of deter-
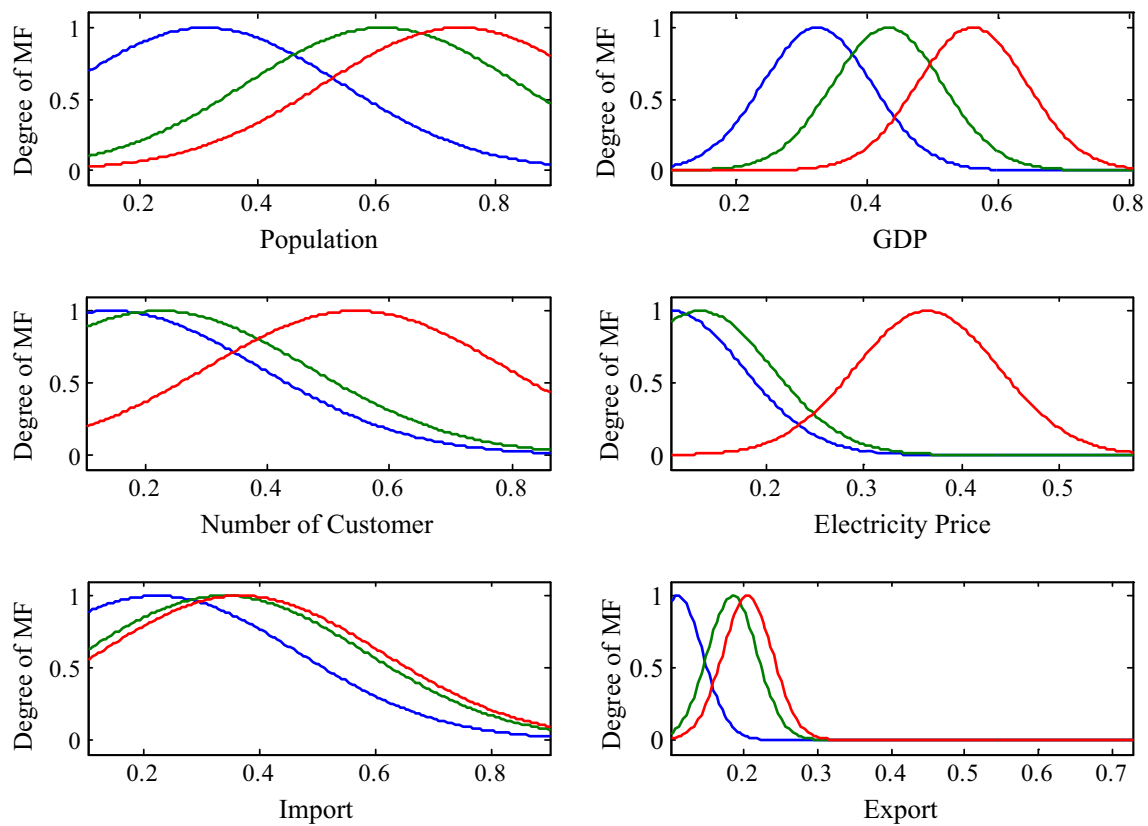
Fig. 3 Membership functions of six inputs

Table 4 Statistical results of grid partitioning-based ANFIS, FCM-based ANFIS, SC-based ANFIS and proposed model for validation phase

Model	No. of rules	MAE	RMSE	R^2
grid partitioning-based ANFIS	–	–	–	–
FCM-based ANFIS (with 3 cluster)	3	0.0237	0.0517	0.977
SC-based ANFIS (Default in MATLAB)	6	0.0323	0.0552	0.949
Proposed model	3	0.0132	0.0238	0.991

minations are also closer to unity, and number of rules is minimum.

In Fig. 4, the actual electricity demand and the demand predicted by these models were plotted in validation phase. Figure 5 shows a scatter plot of actual values against predicted values of the proposed model for validation set. It is observed from Tables 2 and 4, and Fig. 5 that the combination of irregular searching capability of GA and automatic rule generation of SC generates the ANFIS with the minimum number of the rules and best accuracy.

4 Conclusions

This paper has presented a framework for the automatic construction of an adaptive neuro-fuzzy inference system from numerical data based on the integration of subtractive

clustering technique and genetic algorithm. Genetic algorithm was used to resolve the dependence of the clustering process to influence radius of clusters. This algorithm found the optimum value of clusters radius by directly minimizing the SRC value. The applicability and capability of the proposed approach was investigated using the dataset of socioeconomic indicators and industrial's electricity demand in Iran and results were compared with grid partitioning-based ANFIS, fuzzy c-means-based ANFIS and subtractive clustering-based ANFIS. The comparisons revealed that the proposed model, through the integration of genetic algorithm and subtractive clustering technique, is remarkably effective in terms of both accuracy and the number of rules. This conclusion has been supported by the 0.0132 MAE value, 0.0238 RMSE value, and 0.991 R^2 value. The findings demonstrate that proposed model can serve as a reliable and simple tool to predict energy demand. This will give
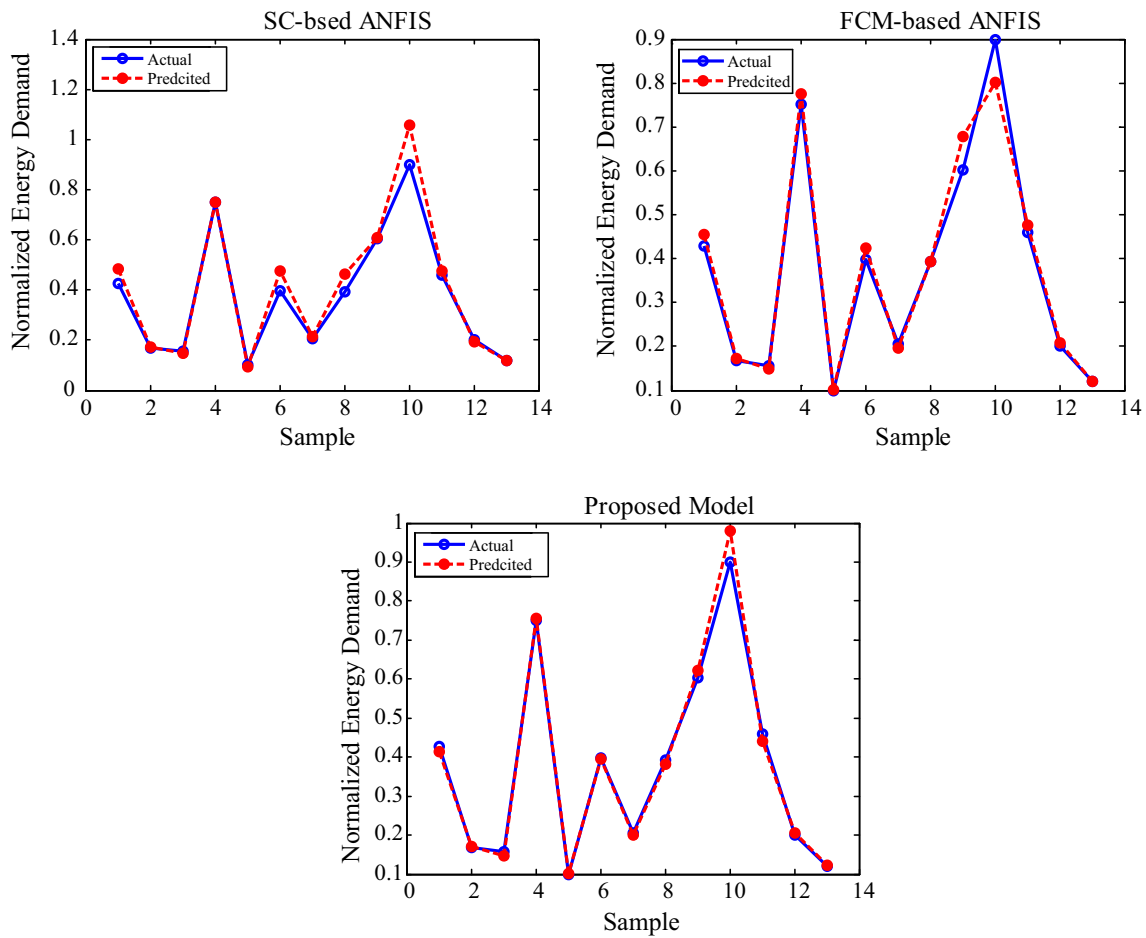
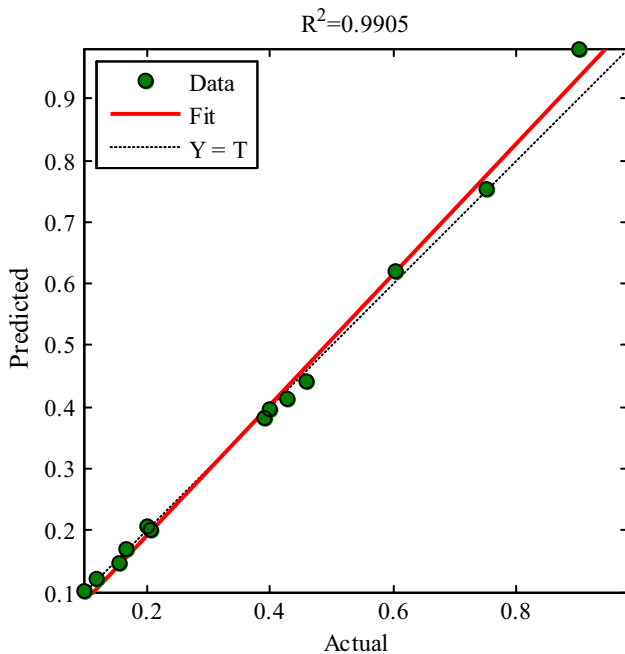**Fig. 4** Actual and predicted electricity demand for validation phase

Fig. 5 Scatter plot of actual and predicted electricity demand by proposed model for validation phase

constructive insight to policy makers in developing energy policies.

One problem when considering the gradient descent technique for tuning of the rules' premise MFs is tendency to get trapped in local optima. Also derived solutions by gradient descent technique are strongly dependent on the initial values of parameters. Attempts will be made in future to introduce self-constructing neuro-fuzzy system by evolutionary algorithm, which can find proper fuzzy rules and tune parameters of fuzzy rules (premise and consequent) simultaneously.

References

Akdemir B, Çetinkaya N (2012) Long-term load forecasting based on adaptive neural fuzzy inference system using real energy data. Energy Proc 14:794–799

Al-Ghandoor A, Samhouri M, Al-Hinti I, Jaber J, Al-Rawashdeh M (2012) Projection of future transport energy demand of Jordan using adaptive neuro-fuzzy technique. Energy 38:128–135

Aliyari M Sh, Teshnehlab M, Sedigh AK (2007) A novel hybrid learning algorithm for tuning ANFIS parameters using adaptive weighted PSO. In: Proceedings of the sixth IEEE international conference on fuzzy systems

Al-Shobaki S, Mohsen M (2008) Modeling and forecasting of electrical power demands for capacity planning. Energy Convers Manag 49(11):3367–3375

Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York

Cárdenas JJ, Romeral L, Garcia A, Andrade F (2012) Load forecasting framework of electricity consumptions for an intelligent energy management system in the user-side. Expert Syst Appl 39(5):5557–5565

Chen Y, Yang B, Abraham A, Peng L (2007) Automatic design of hierarchical Takagi–Sugeno type fuzzy systems using evolutionary algorithms. IEEE Trans Fuzzy Syst 15(3):385–397

Cheng CH, Wei LY (2010) One step-ahead ANFIS time series model for forecasting electricity loads. Optim Eng 2:303–317

Chiu S (1994) Fuzzy model identification based on cluster estimation. J Intell Fuzzy Syst 2(3):267–278

DeJong K (1988) Learning with genetic algorithms: an overview. Mach Learn 3:121–138

Demirli K, Cheng SX, Muthukumaran P (2003) Subtractive clustering based modeling of job sequencing with parametric search. Fuzzy Sets Syst 137:235–270

Demuth BH, Beale M, Hagan MT (2007) Neural network toolbox 5 user's guide. The MathWorks Incorporation, Natick

Economic time series database (2012) Central bank of the Islamic Republic of Iran, http://www.tsd.cbi.ir/DisplayEn/Content.aspx

Ekonomou L (2010) Greek long-term energy consumption prediction using artificial neural networks. Energy 35(2):512–517

Fadare DA (2009) Modelling of solar energy potential in Nigeria using an artificial neural network model. Appl Energy 86(9):1410–1422

Fogel DB (1995) Evolutionary computation: toward a new philosophy of machine intelligence. IEEE Press, New York

Geem ZW, Roper WE (2009) Energy demand estimation of South Korea using artificial neural network. Energy Policy 37(10):4049–4054

Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading

Ho WH, Tsai JT, Lin BT, Chou JH (2009) Adaptive network-based fuzzy inference system for prediction of surface roughness in end milling process using hybrid Taguchi-genetic learning algorithm. Expert Syst Appl 36:3216–3222

Hong WC (2009) Electric load forecasting by support vector model. Appl Math Model 33:2444–2454

Jang JSR (1993) ANFIS: adaptive network-based fuzzy inference systems. IEEE Trans Syst Man Cybern 23(3):665–685

Jang JSR (1996) Input selection for ANFIS Learning. In: Proceedings of IEEE international conference on fuzzy systems, New Orleans

Jang JSR, Gulley N (2000) Fuzzy logic toolbox: user's guide. The Math Works Inc., Prentice-Hall

Kumar U, Jain VK (2010) Time series models (grey-Markov, grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. Energy 35(4):1709–1716

Limanond T, Jomnonkwao S, Srikaew A (2011) Projection of future transport energy demand of Thailand. Energy Policy 33(5):2754–2763

Liu PY, Li HX (2004) Fuzzy neural network theory and application. World Scientific, Singapore, River Edge

MATLAB (2010) MATLAB Fuzzy logic toolbox reference manual, Version 7.10.0. The MathWorks Incorporation, Natick

Ministry of Energy (2013) Energy balance-sheet 2011. Energy report and statistics, Tehran, Iran [in Farsi]

Mollaiy-Berneti S (2013) A hybrid approach based on the combination of adaptive neuro-fuzzy inference system and imperialist competitive algorithm: oil flow rate of the wells prediction case study. Int J Comput Intell Syst 6(2):198–208

Oliveira MV, Schirru R (2009) Applying particle swarm optimization algorithm for tuning a neuro-fuzzy inference system for sensor monitoring. Prog Nucl Energy 51:177–183

Rissanen J (1978) Modeling by shortest data description. Automatica 14:465–471

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Tavanir holding company (2012) Statistical report on 45 years activities of Iran electric power industry, Deputy of human resources and research, management of information technology and statistics

Wang L, Yen J (1999) Extracting fuzzy rules for system modeling using a hybrid of genetic algorithms and Kalman filter. Fuzzy Sets Syst 101:353–362

Xia C, Wang J, McMenemy K (2010) Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks. Int J Electr Power Energy Syst 32:743–750

Yager R, Filev D (1994) Generation of fuzzy rules by mountain clustering. J Intell Fuzzy Syst 2(3):209–219

Ying LC, Pan MC (2008) Using adaptive network based fuzzy inference system to forecast regional electricity loads. Energy Convers Manag 49(2):205–211

Zahedi G, Azizi S, Bahadori A, Elkamel A, Alwi SRW (2013) Electricity demand estimation using an adaptive neuro-fuzzy network: a case study from the Ontario province-Canada. Energy 49:323–328

Zhang X, Mao Y (2009) The relationship between energy consumption and economic growth in china based on ANFIS, In: International workshop on intelligent systems and applications, ISA, Wuhan

Zhang M, Mu H, Li G, Ning Y (2009) Forecasting the transport energy demand based on PLSR method in China. Energy 34(9):1396–1400