We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX — CLARIVATE ANALYTICS — INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

A Multi Adaptive Neuro Fuzzy Inference System for Short Term Load Forecasting by Using Previous Day Features

Zohreh Souzanchi Kashani
Young Researchers Club, Mashhad Branch,
Islamic Azad University, Mashhad
Iran

1. Introduction

Load forecasting had an important role in power system design, planning and development and it is the base of economical studies of energy distribution and power market. The period of load forecasting can be for one year or month (long-term or medium-term) and for one day or hour (short-term) [1, 2, 3, and 4].

For short-term load forecasting several factors should be considered, such as time factors, weather data, and possible customers' classes. The medium- and long-term forecasts take into account the historical load and weather data, the number of customers in different categories, the appliances in the area and their characteristics including age, the economic and demographic data and their forecasts, the appliance sales data, and other factors [17].

The time factors include the time of the year, the day of the week, and the hour of the day. There are important differences in load between weekdays and weekends. The load on different weekdays also can behave differently. For example, in Iran, Fridays is weekends, may have structurally different loads than Saturdays through Thursday. This is particularly true during the summer time. Holidays are more difficult to forecast than non-holidays because of their relative infrequent occurrence.

Several techniques have been used for load forecasting that among its common methods we can refer to linear-regression model, ARMA, BOX-Jenkis[5] and filter model of Kalman, expert systems [6] and ANN [1-4,7]. According to load-forecasting complex nature, however its studying by linear techniques cannot meet the need of having high accuracy and being resistant. Adaptive neural-fuzzy systems can learn and build any non-linear and complex record through educational input-output data.

Then neural-fuzzy systems have many applications in studying load forecasting and power systems according to the non-linear and complex nature of power nets. Among them we can refer to load-peak forecasting and daily network load-curve forecasting.

The east of Iran power plant consumed load information was used for simulation of consumed load forecasting system. The effect of weather forecasting information in

consumed load was considered by entering Mashhad climate information gathered from weather forecasting department of the province.

2. A review of previous works

Certain days load model (formal and informal vacations) is completely different from load model of working days of week (Saturday to Wednesday), but is very similar to its near Fridays. Short-term load forecasting by using fuzzy system cannot have a good function for load forecasting of days by itself since load model of special off-days has a big difference to a usual days. As usual days load model is different regarding the surface and the shape of curve, therefore we need an expert system for adjusting the primary forecasting which apply necessary information for results correction by using an expert person's experience.

On the other hand the power price is a signal with high frequency at competitive market; multi season changes, calendar effect weekends and formal vacation) and the high percentage of unusual prices are mostly during periods of demand increase [8]. The behavior of load curve for different week days is different and in sequential weeks is similar to each other. In this paper, authors use ARMA[1] and ANFIS[2] models for power signal forecasting. A compound method is also suggested in [9] based on neural network that forecast power price and load simultaneously. In [10, 11], PSO[3] has been used for forecasting that in these papers it is in the form of long-term. In [12], the method of neural network learning and SVR[4] is presented in order to a faster forecasting. A local learning method is introduced here and KNN[5] is used for model optimizing. In [13], power load model is also mentioned as a non-linear model and a method is suggested that has the capability of non-linear map.

This paper purpose is introducing SVR with a new algorithm for power load forecasting. SVR and ANN are used for error reduction.

2.1 Short-term load forecasting methods

As we use short term load forecasting in our method, review some important methods here.

A large variety of statistical and artificial intelligence techniques have been developed for short-term load forecasting [17].

Similar-day approach. This approach is based on searching historical data for days within one, two, or three years with similar characteristics to the forecast day. Similar characteristics include weather, day of the week, and the date. The load of a similar day is considered as a forecast. Instead of a single similar day load, the forecast can be a linear combination or regression procedure that can include several similar days. The trend coefficients can be used for similar days in the previous years.

Regression methods. Regression is the one of most widely used statistical techniques. For electric load forecasting regression methods are usually used to model the relationship of

---

[1] Autoregressive moving average
[2] Adaptive Neural- Fuzzy Inference System
[3] Particle swarm optimization
[4] Support vector regression
[5] K-nearest neighbor

load consumption and other factors such as weather, day type, and customer class. Engle et al. [18] presented several regression models for the next day peak forecasting. Their models incorporate deterministic influences such as holidays, stochastic influences such as average loads, and exogenous influences such as weather. References [19], [20], [21], [22] describe other applications of regression models to loads forecasting.

Time series. Time series methods are based on the assumption that the data have an internal structure, such as autocorrelation, trend, or seasonal variation. Time series forecasting methods detect and explore such a structure. Time series have been used for decades in such fields as economics, digital signal processing, as well as electric load forecasting. In particular, ARMA (autoregressive moving average), ARIMA (autoregressive integrated moving average), ARMAX (autoregressive moving average with exogenous variables), and ARIMAX (autoregressive integrated moving average with exogenous variables) are the most often used classical time series methods. ARMA models are usually used for stationary processes while ARIMA is an extension of ARMA to non-stationary processes. ARMA and ARIMA use the time and load as the only input parameters. Since load generally depends on the weather and time of the day, ARIMAX is the most natural tool for load forecasting among the classical time series models. Fan and McDonald [23] and Cho et al. [24] describe implementations of ARIMAX models for load forecasting. Yang et al. [25] used evolutionary programming (EP) approach to identify the ARMAX model parameters for one day to one week ahead hourly load demand forecast. Evolutionary programming [26] is a method for simulating evolution and constitutes a stochastic optimization algorithm. Yang and Huang [27] proposed a fuzzy autoregressive moving average with exogenous input variables (FARMAX) for one day ahead hourly load forecasts.

Neural networks. The use of artificial neural networks (ANN or simply NN) has been a widely studied electric load forecasting technique since 1990 [28]. Neural networks are essentially non-linear circuits that have the demonstrated capability to do non-linear curve fitting. The outputs of an artificial neural network are some linear or nonlinear mathematical function of its inputs. The inputs may be the outputs of other network elements as well as actual network inputs. In practice network elements are arranged in a relatively small number of connected layers of elements between network inputs and outputs. Feedback paths are sometimes used. In applying a neural network to electric load forecasting, one must select one of a number of architectures (e.g. Hopfield, back propagation, Boltzmann machine), the number and connectivity of layers and elements, use of bi-directional or uni-directional links, and the number format (e.g. binary or continuous) to be used by inputs and outputs, and internally.

The most popular artificial neural network architecture for electric load forecasting is back propagation. Back propagation neural networks use continuously valued functions and supervised learning. That is, under supervised learning, the actual numerical weights assigned to element inputs are determined by matching historical data (such as time and weather) to desired outputs (such as historical electric loads) in a pre-operational "training session". Artificial neural networks with unsupervised learning do not require pre-operational training. Bakirtzis et al. [29] developed an ANN based short-term load forecasting model for the Energy Control Center of the Greek Public Power Corporation. In the development they used a fully connected three-layer feed forward ANN and back propagation algorithm was used for training.

Input variables include historical hourly load data, temperature, and the day of the week. The model can forecast load profiles from one to seven days. Also Papalexopoulos et al. [30] developed and implemented a multi-layered feed forward ANN for short-term system load forecasting. In the model three types of variables are used as inputs to the neural network: season related inputs, weather related inputs, and historical loads. Khotanzad et al. [31] described a load forecasting system known as ANNSTLF. ANNSTLF is based on multiple ANN strategies that capture various trends in the data. In the development they used a multilayer perceptron trained with the error back propagation algorithm. ANNSTLF can consider the effect of temperature and relative humidity on the load. It also contains forecasters that can generate the hourly temperature and relative humidity forecasts needed by the system. An improvement of the above system was described in [32]. In the new generation, ANNSTLF includes two ANN forecasters, one predicts the base load and the other forecasts the change in load. The final forecast is computed by an adaptive combination of these forecasts. The effects of humidity and wind speed are considered through a linear transformation of temperature. As reported in [32], ANNSTLF was being used by 35 utilities across the USA and Canada. Chen et al. [4] developed a three layer fully connected feed forward neural network and the back propagation algorithm was used as the training method. Their ANN though considers the electricity price as one of the main characteristics of the system load. Many published studies use artificial neural networks in conjunction with other forecasting techniques (such as with regression trees [26], time series [33] or fuzzy logic [34]).

Expert systems. Rule based forecasting makes use of rules, which are often heuristic in nature, to do accurate forecasting. Expert systems, incorporates rules and procedures used by human experts in the field of interest into software that is then able to automatically make forecasts without human assistance.

Expert system use began in the 1960's for such applications as geological prospecting and computer design. Expert systems work best when a human expert is available to work with software developers for a considerable amount of time in imparting the expert's knowledge to the expert system software. Also, an expert's knowledge must be appropriate for codification into software rules (i.e. the expert must be able to explain his/her decision process to programmers). An expert system may codify up to hundreds or thousands of production rules. Ho et al. [35] proposed a knowledge-based expert system for the short-term load forecasting of the Taiwan power system. Operator's knowledge and the hourly observations of system load over the past five years were employed to establish eleven day types. Weather parameters were also considered. The developed algorithm performed better compared to the conventional Box-Jenkins method. Rahman and Hazim [36] developed a site-independent technique for short-term load forecasting. Knowledge about the load and the factors affecting it are extracted and represented in a parameterized rule base. This rule base is complemented by a parameter database that varies from site to site. The technique was tested in several sites in the United States with low forecasting errors.

The load model, the rules, and the parameters presented in the paper have been designed using no specific knowledge about any particular site. The results can be improved if operators at a particular site are consulted.

Fuzzy logic. Fuzzy logic is a generalization of the usual Boolean logic used for digital circuit design. An input under Boolean logic takes on a truth value of "0" or "1". Under fuzzy logic

an input has associated with it a certain qualitative ranges. For instance a transformer load may be "low", "medium" and "high". Fuzzy logic allows one to (logically) deduce outputs from fuzzy inputs. In this sense fuzzy logic is one of a number of techniques for mapping inputs to outputs (i.e. curve fitting).

Among the advantages of fuzzy logic are the absence of a need for a mathematical model mapping inputs to outputs and the absence of a need for precise (or even noise free) inputs. With such generic conditioning rules, properly designed fuzzy logic systems can be very robust when used for forecasting. Of course in many situations an exact output (e.g. the precise 12PM load) is needed. After the logical processing of fuzzy inputs, a "defuzzification" process can be used to produce such precise outputs. References [37], [38], [39] describe applications of fuzzy logic to electric load forecasting.

Support vector machines. Support Vector Machines (SVMs) are a more recent powerful technique for solving classification and regression problems. This approach was originated from Vapnik's [40] statistical learning theory. Unlike neural networks, which try to define complex functions of the input feature space, support vector machines perform a nonlinear mapping (by using so-called kernel functions) of the data into a high dimensional (feature) space. Then support vector machines use simple linear functions to create linear decision boundaries in the new space. The problem of choosing an architecture for a neural network is replaced here by the problem of choosing a suitable kernel for the support vector machine [41]. Mohandes [42] applied the method of support vector machines for short-term electrical load forecasting. The author compares its performance with the autoregressive method. The results indicate that SVMs compare favorably against the autoregressive method. Chen et al. [43] proposed a SVM model to predict daily load demand of a month. Their program was the winning entry of the competition organized by the EU Load NITE network. Li and Fang [44] also used a SVM model for short-term load forecasting.

3. Consumed load model

The load forecasting art is in selecting the most appropriate way and model for and the closest ones to the existing reality of the network among different methods and models of load forecasting, by studying and analyzing the last procedure of load and recognizing the effective factors sufficiently and maximizing each of them, and then in this way it forecasts different time periods required for the network with an acceptable approximation. It should be accepted that there is always some error in load forecasting due to the accidental load behavior but never this error should go further than the acceptable and tolerable limit. Relative accuracy has a particular importance in load forecasting in power industry. Especially when load forecasting is the basis of network development planning and power plant capacity. Since, any forecasting with open hand causes extra investment and the installation capacity to be useless and vice versa any forecasting less than real needs, faces the network with shortage in production and damages the instruments due to extra load.

Consumed load model is influenced by different parameters such as weather, vacations or holidays, working days of week and etc. in order to build a short-term load forecasting system, we should consider the influence of different parameters in load forecasting, which it can be full field by a correct selection of system entries. Selection of these parameters depends on experimental observations and is influenced by the environment conditions and is determined by trial and error.

4. Reviewing the predictability of time series by the help of lyapunov exponent[6]

Chaos is a phenomenon that occurs in many non-linear definable systems which show a high sensitivity to the primary conditions and semi random behavior. These systems will remain stable in the chaotic mode if they provide the Lyapunov exponent equations.

4.1 Background

Detecting the presence of chaos in a dynamical system is an important problem that is solved by measuring the largest Lyapunov exponent. Lyapunov exponents quantify the exponential divergence of initially close state-space trajectories and estimate the amount of chaos in a system.[50]

Over the past decade, distinguishing deterministic chaos from noise has become an important problem in many diverse fields, e.g., physiology [51], economics [52]. This is due, in part, to the availability of numerical algorithms for quantifying chaos using experimental time series. In particular, methods exist for calculating correlation dimension (D2) [53], Kolmogorov entropy [54], and Lyapunov characteristic exponents. Dimension gives an estimate of the system complexity; entropy and characteristic exponents give an estimate of the level of chaos in the dynamical system.

The Grassberger-Procaccia algorithm (GPA) [53] appears to be the most popular method used to quantify chaos. This is probably due to the simplicity of the algorithm [55] and the fact that the same intermediate calculations are used to estimate both dimension and entropy.

However, the GPA is sensitive to variations in its parameters, e.g., number of data points [56], embedding dimension [56], reconstruction delay [57], and it is usually unreliable except for long, noise-free time series. Hence, the practical significance of the GPA is questionable, and the Lyapunov exponents may provide a more useful characterization of chaotic systems.

For time series produced by dynamical systems, the presence of a positive characteristic exponent indicates chaos. Furthermore, in many applications it is sufficient to calculate only the largest Lyapunov exponent ($\lambda1$). However, the existing methods for estimating $\lambda1$ suffer from at least one of the following drawbacks: (1) unreliable for small data sets, (2) computationally intensive, (3) relatively difficult to implement. For this reason, we have developed a new method for calculating the largest Lyapunov exponent. The method is reliable for small data sets, fast, and easy to implement. "Easy to implement" is largely a subjective quality, although we believe it has had a notable positive effect on the popularity of dimension estimates.

For a dynamical system, sensitivity to initial conditions is quantified by the Lyapunov exponents. For example, consider two trajectories with nearby initial conditions on an attracting manifold. When the attractor is chaotic, the trajectories diverge, on average, at an exponential rate characterized by the largest Lyapunov exponent [58]. This concept is also generalized for the spectrum of Lyapunov exponents, λi (i=1, 2, ..., n), by considering a small

---

[6] Lyapunov exponent

n-dimensional sphere of initial conditions, where n is the number of equations (or, equivalently, the number of state variables) used to describe the system. As time (t) progresses, the sphere evolves into an ellipsoid whose principal axes expand (or contract) at rates given by the Lyapunov exponents.

The presence of a positive exponent is sufficient for diagnosing chaos and represents local instability in a particular direction. Note that for the existence of an attractor, the overall dynamics must be dissipative, i.e., globally stable, and the total rate of contraction must outweigh the total rate of expansion. Thus, even when there are several positive Lyapunov exponents, the sum across the entire spectrum is negative.

Wolf et al. [59] explain the Lyapunov spectrum by providing the following geometrical interpretation. First, arrange the n principal axes of the ellipsoid in the order of most rapidly expanding to most rapidly contracting. It follows that the associated Lyapunov exponents will be arranged such that

$$\lambda_1 > \lambda_2 > \ldots > \lambda_n$$

where λ_1 and λ_n correspond to the most rapidly expanding and contracting principal axes, respectively. Next, recognize that the length of the first principal axis is proportional to $e^{\lambda 1 t}$; the area determined by the first two principal axes is proportional to $e^{(\lambda 1 + \lambda 2)t}$; and the volume determined by the first k principal axes is proportional to $e^{(\lambda 1 + \lambda 2 + \cdots + \lambda k)t}$. Thus, the Lyapunov spectrum can be defined such that the exponential growth of a k-volume element is given by the sum of the k largest Lyapunov exponents. Note that information created by the system is represented as a change in the volume defined by the expanding principal axes. The sum of the corresponding exponents, i.e., the positive exponents, equals the Kolmogorov entropy (K) or mean rate of information gain [58]:

$$K = \sum_{\lambda i > 0} \lambda i$$

When the equations describing the dynamical system are available, one can calculate the entire Lyapunov spectrum. The approach involves numerically solving the system's n equations for n+1 nearby initial conditions. The growth of a corresponding set of vectors is measured, and as the system evolves, the vectors are repeatedly reorthonormalized using the Gram-Schmidt procedure. This guarantees that only one vector has a component in the direction of most rapid expansion, i.e., the vectors maintain a proper phase space orientation. In experimental settings, however, the equations of motion are usually unknown and this approach is not applicable. Furthermore, experimental data often consist of time series from a single observable, and one must employ a technique for attractor reconstruction, e.g., method of delays [60], singular value decomposition.

As suggested above, one cannot calculate the entire Lyapunov spectrum by choosing arbitrary directions for measuring the separation of nearby initial conditions. One must measure the separation along the Lyapunov directions which correspond to the principal axes of the ellipsoid previously considered. These Lyapunov directions are dependent upon the system flow and are defined using the Jacobian matrix, i.e., the tangent map, at each point of interest along the flow [58]. Hence, one must preserve the proper phase space orientation by using a suitable approximation of the tangent map. This requirement, however, becomes unnecessary when calculating only the largest Lyapunov exponent.

If we assume that there exists an ergodic measure of the system, then the multiplicative ergodic theorem of Oseledec [61] justifies the use of arbitrary phase space directions when calculating the largest Lyapunov exponent with smooth dynamical systems. We can expect that two randomly chosen initial conditions will diverge exponentially at a rate given by the largest Lyapunov exponent [62]. In other words, we can expect that a random vector of initial conditions will converge to the most unstable manifold, since exponential growth in this direction quickly dominates growth (or contraction) along the other Lyapunov directions. Thus, the largest Lyapunov exponent can be defined using the following equation where d(t) is the average divergence at time t and C is a constant that normalizes the initial separation:

$$d(t) = Ce^{\lambda_1 t}$$

For experimental applications, a number of researchers have proposed algorithms that estimate the largest Lyapunov exponent [55,59], the positive Lyapunov spectrum, i.e., only positive exponents [59], or the complete Lyapunov spectrum [58]. Each method can be considered as a variation of one of several earlier approaches [59] and as suffering from at least one of the following drawbacks: (1) unreliable for small data sets, (2) computationally intensive, (3) relatively difficult to implement. These drawbacks motivated our search for an improved method of estimating the largest Lyapunov exponent.

4.2 Calculation of lyapunov exponent for time series

In order to calculate Lyapunov exponent for those systems which their equation is not determined and their time series is not available, different algorithm is suggested [45-49].

The algorithm proposed by Wolf [48], seeks the time series of close points in the phase space. These points went round the phase space or got divergent rapidly. Close points in the same direction are selected.

The differential coefficient is in the direction of the maximum development and their average logarithm on the route of phase space yields the biggest Lyapunov exponent. Suppose that series of $x_0, x_1, x_2, \dots x_k$ is available and the interval between them is obtained as $t_n - t_0 = n_\tau$ that τ is the interval between two successive measurement. If the system has chaotic behavior, we can explain divergence of the adjacent routes based on the difference range between them, as following.

$$d_0 = |x_j - x_i| \tag{1}$$

$$d_1 = |x_{j+1} - x_{i+1}| \tag{2}$$

$$.$$
$$.$$
$$.$$

$$d_n = |x_{j+n} - x_{i+n}| \tag{3}$$

It is supposed that d_n will increase exponential by n increase:

$$d_n = d_0 e^{\lambda n} \tag{4}$$

So by calculating its logarithm we have:

$$\lambda = \frac{1}{n} \text{Ln} \frac{d_n}{d_0} \tag{5}$$

There should be at least one Lyapunov exponent bigger than zero to have chaos, the existence of positive value of λ means the chaotic behavior of system. Therefore, in order to Table 1 we can expect system to forecast.

Winter	Fall	Summer	Spring
0.07563	0.05428	0.0444	0.0523

Table 1. Lyapunov exponent for seasons of one year

5. Preparing the input data

First step in the process of electricity load forecasting is to provide last information of the system load being studied. After preparing the input data matrix, it is turn of classification. The reason of this classification is the existence of completely determined models in different days that were referred to in many references. Among different days of weeks, Saturday to Thursday which are working days in Iran, have the same load model. Fridays have also their own particular model and have a low level of load. Special days have a completely different model, too. So it seems necessary at the first look that each of these classes should be analyzed separately. We consider 2 groups of features that refer to previous days; 2, 7, and 14 day ago, and 2, 3, 4 day ago.

6. Adaptive neural- Fuzzy inference system

ANFIS, proposed by Jang [14, 15], is an architecture which functionally integrates the interpretability of a fuzzy inference system with adaptability of a neural network. Loosely speaking ANFIS is a method for tuning an existing rule base of fuzzy system with a learning algorithm based on a collection of training data found in artificial neural network. Due to the less tunable use of parameters of fuzzy system compared with conventional artificial neural network, ANFIS is trained faster and more accurately than the conventional artificial neural network. An ANFIS which corresponds to a Sugeno type fuzzy model of two inputs and single output is shown in Fig. 1. A rule set of first order Sugeno fuzzy system is the following form:

Rule i: If x is Ai and y is Bi then fi = $p_i x + q_i y + r_i$.

ANFIS structure as shown in Figure 1 is a weightless multi-layer array of five different elements [15]:

- Layer 1: Input data are fuzzified and neuron values are represented by parameterized membership functions;
  - $O_{l,i}$ is the output of the ith node of the layer l.
  - Every node i in this layer is an adaptive node with a node function

$$O_{1,i} = \mu A_i(x) \text{ for } i = 1, 2, \text{ or}$$

$$O_{1,i} = \mu B_{i-2}(x) \text{ for } i = 3, 4$$

- x (or y) is the input node i and A_i (or B_{i-2}) is a linguistic label associated with this node
- Therefore $O_{1,i}$ is the membership grade of a fuzzy set (A1,A2,B1,B2).
- Typical membership function:

$$\mu A(x) = \frac{1}{1+|\frac{x-ci}{ai}|2bi}$$

- a_i, b_i, c_i is the parameter set.
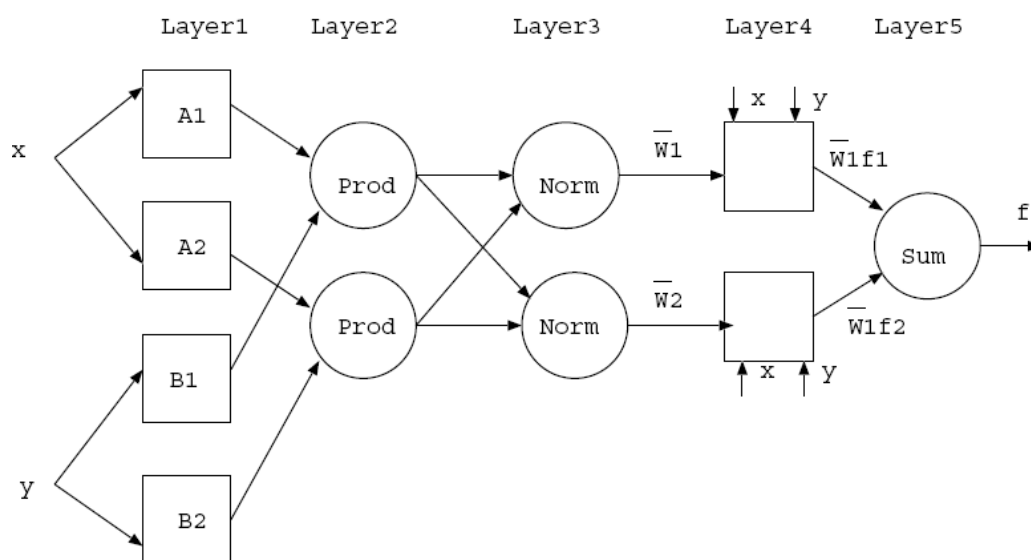- Parameters are referred to as premise parameters.

Fig. 1. ANFIS architecture

- Layer 2: The activation of fuzzy rules is calculated via differentiable T-norms (usually, the soft-min or product);
  - Every node in this layer is a fixed node labeled Prod.
  - The output is the product of all the incoming signals.
  - $O_{2,i} = w_i = \mu A_i(x) \cdot \mu B_i(y)$, $i = 1, 2$
  - Each node represents the fire strength of the rule
  - Any other T-norm operator that perform the AND operator can be used
- Layer 3: A normalization (arithmetic division) operation is realized over the rules matching values;
  - Every node in this layer is a fixed node labeled Norm.
  - The ith node calculates the ratio of the ith rulet's firing strenght to the sum of all rulet's firing strengths.
  - $O_{3,i} = \bar{w}_i = \frac{w_i}{w_1+w_2}$, i = 1, 2
  - Outputs are called normalized firing strengths.
- Layer 4: The consequent part is obtained via linear regression or multiplication between the normalized activation level and the output of the respective rule;
  - Every node i in this layer is an adaptive node with a node function:

$$O_{4,1} = \overline{w}_i f_i = \overline{w}_i(p_x + q_i y + r_i)$$

- $\overline{w}_i$ is the normalized firing strenght from layer 3.
- $\{p_i, q_i, r_i\}$ is the parameter set of this node.
- These are referred to as consequent parameters.
- Layer 5: The NFN output is produced by an algebraic sum over all rules outputs.
  - The single node in this layer is a fixed node labeled *sum*, which computes the overall output as the summation of all incoming signals:
  - overall output = $O_{5,1} = \sum \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$

The main objective of the ANFIS design is to optimize the ANFIS parameters. There are two steps in the ANFIS design. First is design of the premise parameters and the other is consequent parameter training. There are several methods proposed for designing the premise parameter such as grid partition, fuzzy C-means clustering and subtractive clustering. Once the premise parameters are fixed, the consequent parameters are obtained based on the input-output training data. A hybrid learning algorithm is a popular learning algorithm used to train the ANFIS for this purpose.

- ANFIS uses a hybrid learning algorithm to identify the membership function parameters of single-output, Sugeno type fuzzy inference systems (FIS).
- There are many ways of using this function.
- Some examples:

[FIS,ERROR] = ANFIS(TRNDATA)

[FIS,ERROR] = ANFIS(TRNDATA,INITFIS)

7. The proposed method for power consumed load forecasting

Since fuzzy methods and systems were presented for using in different applications, researchers noticed that making a fuzzy powerful system is not a simple work. The reason is that finding suitable fuzzy rules and membership functions is not a systematic work and mainly requires many trails and errors to reach to the best possible efficiency. Therefore the idea of using learning algorithms was proposed for fuzzy systems. Meanwhile learning of fuzzy network proposed them as the first goals for being unified in fuzzy methods in order to make the development and usage process of fuzzy systems automatic for different applications. Function estimation by using the learning methods is proposed in neural networks and neural-fuzzy networks.

In the suggested methods we forecast load consume and its improvement by the help of the offered method. One of the famous neural-fuzzy systems for function estimation is ANFIS model. We used this system for power consumed load forecasting in this paper too, but with this difference that we used one separate adaptive neural-fuzzy system for each season of the year. Although at the time of training these systems data overlapping is considered, because data of each season of the year is not completely independent and there is some similarities between the first days of a season with its previous season regarding the amount of load consumption. Figure 2 shows the diagram of multi adaptive neural-fuzzy system (multi ANFIS).

As it is shown too, in the Figure 2, we us a switch for any subsystem of a season be thought in lieu of that season. Therefore the time of system training and testing will decrease and the entrance of extra data is prevented.
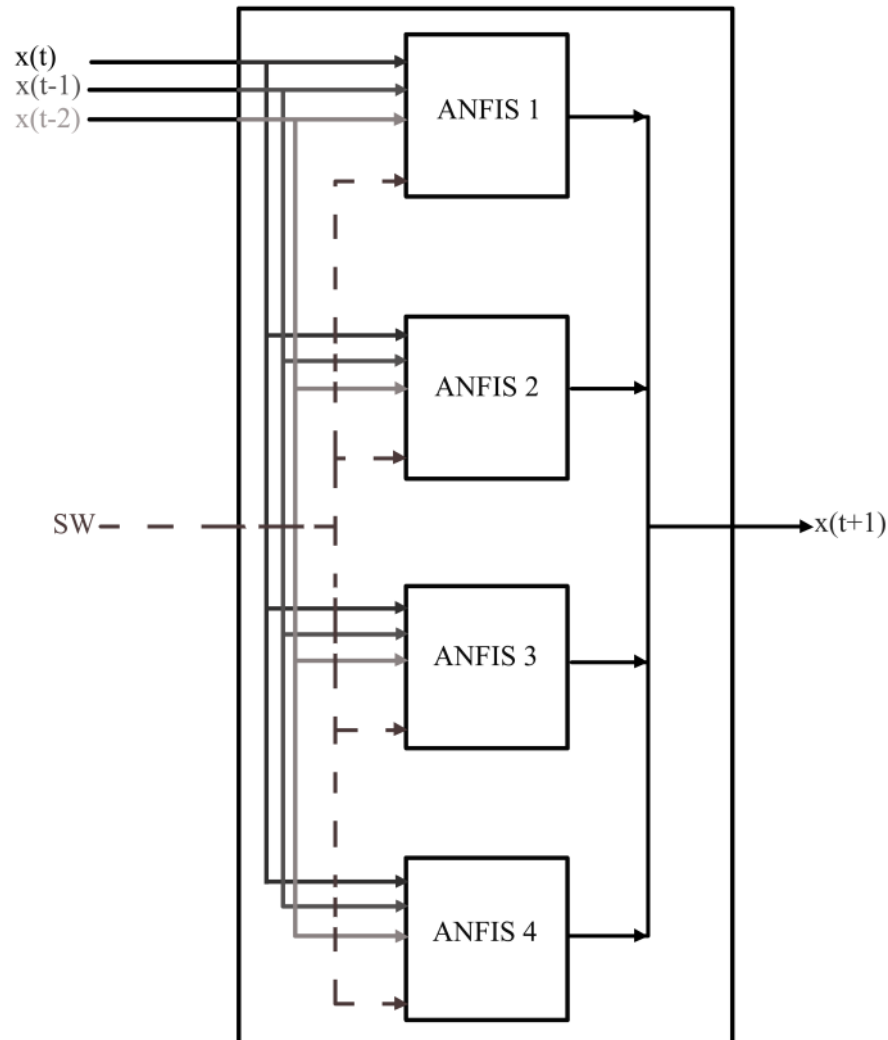

Fig. 2. Implemented Diagram of Multi ANFIS

8. Result

In the proposed method we classified day into two categories. We divide the season days into two groups of working days (Saturday to Thursday) and holidays that their load consumption is different from other days.

Here we also calculated the output of Multi ANFIS based on the features of previous day, one time with 2, 7, and 14 day ago and another time with 2, 3, and 4 day ago. You can see the results in Table 2 and 3.

The amount of the accuracy of the performance of any of calculation methods in load forecasting is determined by measuring the obtained values of system model and comparing it with real data.

Mean Absolute Percentage Error (MAPE) is used for studying the performance of every
mentioned method with the data of related test. MAPE is determined by following relation:

$$\text{MAPE} = 1/N \left(\sum_{i=1}^{N} APE_i \right) \tag{6}$$

$$\text{APE} = |(V(\text{forecast}) - V(\text{actual})) / V(\text{actual})| * 100\% \tag{7}$$

MAPE for working days	load consumption forecasting
1.5409	Spring
2.1869	Summer
2.4575	Fall
1.5116	Winter

Table 2. Power load consumption forecasting for the working days (saturday to thursday)
with 2, 3, and 4 day ago

MAPE for working days	Load consumption forecasting
0.9602	Spring
0.8568	Summer
1.1392	Fall
1.3015	Winter

Table 3. Power load consumption forecasting for the working days (saturday to thursday)
with 2, 7, and 14 day ago

As it is obvious of the above Tables, making working days separate from holidays with
using previous days features (2, 7,and 14 day ago) yields a better result, in load
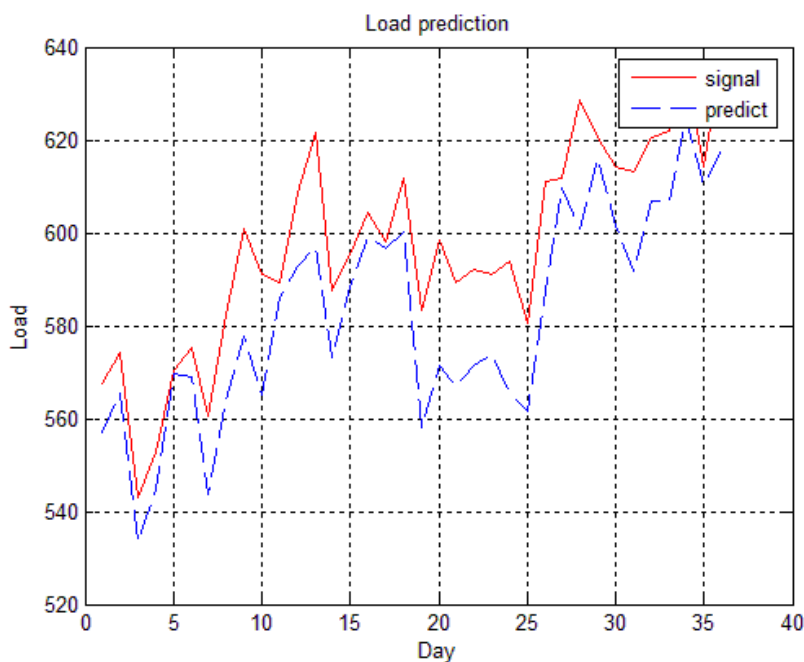consumption forecasting.

Fig. 3. Power load forecasting for Working days (Saturday to Thursday) of fall with features
of 2, 3, and 4 day ago

Also in order to compare, the diagram of daily load forecasting curves for fall through both groups is shown in Figures 3 and 4. It should be mentioned that MATLAB software is used for load forecasting and simulation.
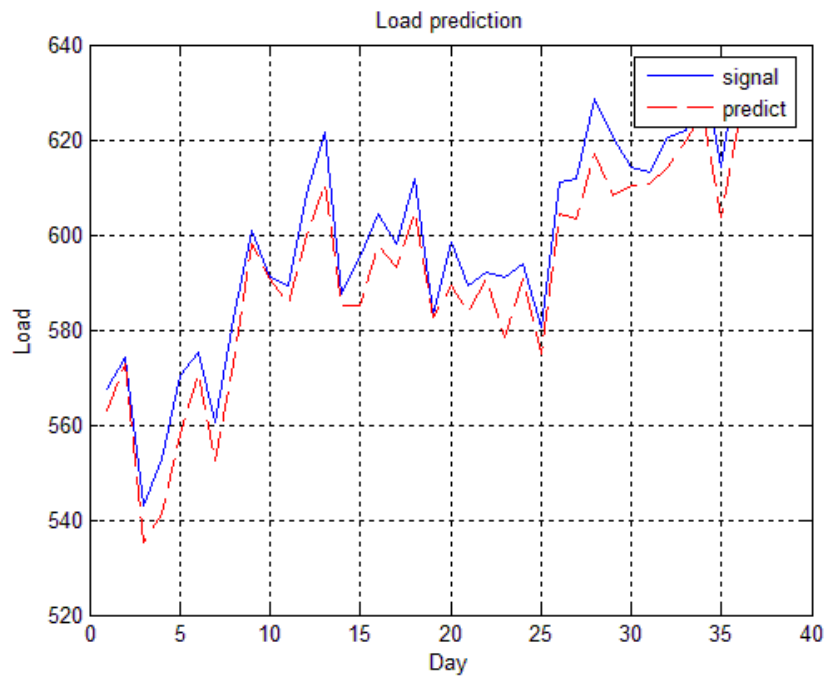

Fig. 4. Power load forecasting for Working days (Saturday to Thursday) of fall with features of 2, 7, and 14 day ago

9. Conclusion and suggestion

Comparing mentioned methods above shows that separation of working days from holidays has a better result in load consumption forecasting. As shown in Figure 5 we can
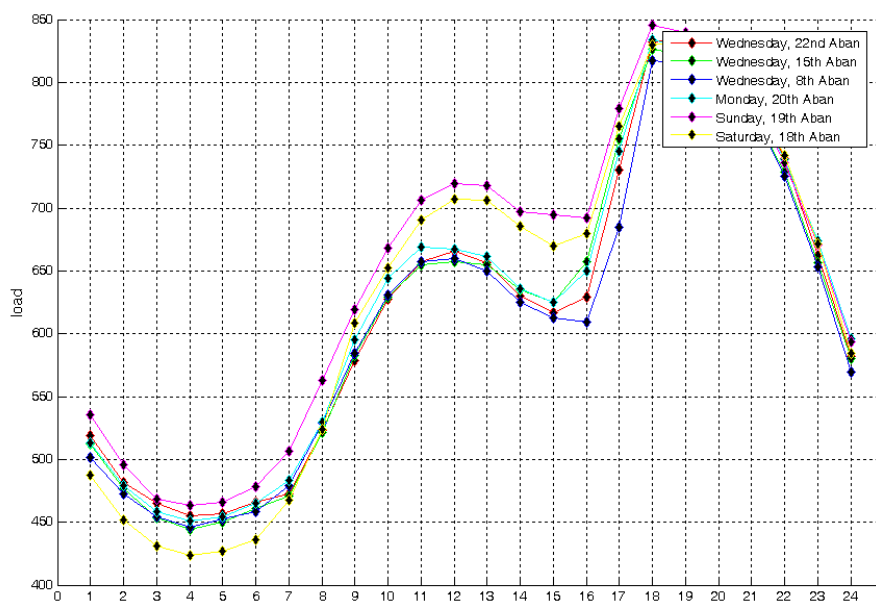
Fig. 5. Compare of the feature of 2, 7 and 14 day ago with 2, 3 and 4 day ago

see that using the features of 2, 7 and 14 day ago are better than 2, 3 and 4 day ago. A cyan and yellow line are refer to 3 and 4 day ago. We can see that these features cannot have good effect on load forecasting.

According to this that in most proposed methods load consumption time series data is used; it seems that we can obtain better results by using time series data of one or more parameters effective in load consumption [16] also with load consumption time series. Accurate load forecasting is very important for electric utilities in a competitive environment created by the electric industry deregulation.

10. References

[1] Huang, S.J. and K.R. Shih, 2003. Short term load forecasting via ARMA model identification including non- Gaussian process consideration. IEEE Trans. Power Syst., 18: 673-679.

[2] Kandil Nahi, Rene Wamkeue, Maarouf saad and Semaan Georges, 2006. An efficient approach for short term load forecasting using artificial neural networks. Int. J. Electric Power Energy syst., 28: 525-530.

[3] Mandal Paras, Tomonobu Senjyu, Naomitsu Urasaki, Toshihisa Funabashi, 2006. A neural network based several hours ahead electric load forecasting using similar days approach. Int. J. Elect. Power Energy Syst., 28: 367-373.

[4] Topalli Ayca Kumluca, Ismet Erkmen and Ihsan Topalli, 2006. Intelligent short term load forecasting in Turkey. Int. J. Electric. Power Energy Syst., 28: 437-447.

[5] Moghram, I. and S. Rahman, 1989. Analysis and evaluation of five short term load forecasting techniques. IEEE Trans. Power Syst., 4: 1484-1491.

[6] Rahman, S. and R. Bhatnager, 1988. An expert system based algorithm for short term load forecast. IEEE Trans. Power Sys., pp: 392-399.

[7] Senjyu, T., H. Takara, K. Uezato and T. Funbasi, 2002. One hour ahead load forecasting using neural network. IEEE Trans. Power sys., 17: 113-118.

[8] F. J. Nogales, J. Contreras, A. J. Conejo and R. Espinola, "Forecasting Next-Day Electricity Prices by Times Series Models", IEEE Transaction On Power Systems, Vol. 17, no.2, May 2002.

[9] Amjadi N, Daraeepour A, "Mixed price and load forecasting of electricity markets by a new iterative prediction method", Electric power systems research, 2009.

[10] Niu D, Li J, Liu D, "Middle-long power load forecasting based on particle swarm optimization", computers and mathematics with applications, 2009.

[11] Alrashidi M, Elnaggar K, "Long term electric load forecasting based on particle swarm optimization", applied enery, 2009.

[12] He W, "Forecasting electricity load with optimized local learning models", Electrical power and energy systems, 2008.

[13] Hong W-C, "Electric load forecasting by support vector model", Applied Mathematical Modelling, 2009.

[14] Jang, J.S.R., C. T. Sun and E. Mizutani. Neuro-Fuzzy and Soft Computing, Prentice Hall, Upper Saddle River, NJ, USA. 1997.

[15] Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system, IEEE Trans. on System, Man, and Cybernetics, vol. 23, pp.665-685, May/June, 1993.

[16] Souzanchi-K Z, Yaghobi M, Akbarzadeh-T M, "Modeling and Forecasting Short_term Electricity Load based on Multi Adaptive Neural-Fuzzy Inference System by Using Temterature", 2nd International conference on Signal Processing Systems(ICSPS), 2010.

[17] E A. Feinberg, D Genethliou. LOAD FORECASTING, APPLIED MATHEMATICS FOR POWER SYSTEMS Load Forecasting

[18] R.F. Engle, C. Mustafa, and J. Rice. Modeling Peak Electricity Demand. Journal of Forecasting, 11:241–251, 1992.

[19] O. Hyde and P.F. Hodnett. An Adaptable Automated Procedure for Short-Term Electricity Load Forecasting. IEEE Transactions on Power Systems, 12:84–93, 1997.

[20] S. Ruzic, A. Vuckovic, and N. Nikolic. Weather Sensitive Method for Short-Term Load Forecasting in Electric Power Utility of Serbia. IEEE Transactions on Power Systems, 18:1581–1586, 2003.

[21] T. Haida and S. Muto. Regression Based Peak Load Forecasting using a Transformation Technique. IEEE Transactions on Power Systems, 9:1788–1794, 1994.

[22] W. Charytoniuk, M.S. Chen, and P. Van Olinda. Nonparametric Regression Based Short-Term Load Forecasting. IEEE Transactions on Power Systems, 13:725–730, 1998.

[23] J.Y. Fan and J.D. McDonald. A Real-Time Implementation of Short-Term Load Forecasting for Distribution Power Systems. IEEE Transactions on Power Systems, 9:988–994, 1994.

[24] M.Y. Cho, J.C. Hwang, and C.S. Chen. Customer Short-Term Load Forecasting by using ARIMA Transfer Function Model. Proceedings of the International Conference on Energy Management and Power Delivery, 1:317–322, 1995.

[25] H.T. Yang, C.M. Huang, and C.L. Huang. Identification of ARMAX Model for Short-Term Load Forecasting: An Evolutionary Programming Approach. IEEE Transactions on Power Systems, 11:403–408, 1996.

[26] D.B. Fogel. An Introduction to Simulated Evolutionary Optimization. IEEE Transactions on Neural Networks, 5:3–14, 1994.

[27] H.T. Yang and C.M. Huang. A New Short-Term Load Forecasting Approach using Self-Organizing Fuzzy ARMAX Models. IEEE Transactions on Power Systems, 13:217–225, 1998.

[28] M. Peng, N.F. Hubele, and G.G. Karady. Advancement in the Application of Neural Networks for Short-Term Load Forecasting. IEEE Transactions on Power Systems, 7:250–257, 1992.

[29] A.G. Bakirtzis, V. Petridis, S.J. Kiartzis, M.C. Alexiadis, and A.H. Maissis. A Neural Network Short-Term Load Forecasting Model for the Greek Power System. IEEE Transactions on Power Systems, 11:858–863, 1996.

[30] A.D. Papalexopoulos, S. Hao, and T.M. Peng. An Implementation of a Neural Network Based Load Forecasting Model for the EMS. IEEE Transactions on Power Systems, 9:1956–1962, 1994.

[31] A. Khotanzad, R.A. Rohani, T.L. Lu, A. Abaye, M. Davis, and D.J. Maratukulam. ANNSTLF–A Neural-Network-Based Electric Load Forecasting System. IEEE Transactions on Neural Networks, 8:835–846, 1997.

[32] A. Khotanzad, R.A. Rohani, and D. Maratukulam. ANNSTLF– Artificial Neural Network Short-Term Load Forecaster–Generation Three. IEEE Transactions on Neural Networks, 13:1413–1422, 1998.

[33] T.W.S. Chow and C.T. Leung. Nonlinear Autoregressive Integrated Neural Network Model for Short-Term Load Forecasting. IEE Proceedings on Generation, Transmission and Distribution, 143:500–506, 1996.

[34] S.E. Skarman and M. Georgiopoulous. Short-Term Electrical Load Forecasting using a Fuzzy ARTMAP Neural Network. Proceedings of SPIE, 181–191, 1998.

[35] K.L. Ho, Y.Y. Hsu, F.F. Chen, T.E. Lee, C.C. Liang, T.S. Lai, and K.K. Chen. Short-Term Load Forecasting of Taiwan Power System using a Knowledge Based Expert System. IEEE Transactions on Power Systems, 5:1214–1221, 1990.

[36] S. Rahman and O. Hazim. Load Forecasting for Multiple Sites: Development of an Expert System-Based Technique. Electric Power Systems Research, 39:161–169, 1996.

[37] S.J. Kiartzis and A.G. Bakirtzis. A Fuzzy Expert System for Peak Load Forecasting: Application to the Greek Power System. Proceedings of the 10th Mediterranean Electrotechnical Conference, 3:1097– 1100, 2000.

[38] V. Miranda and C. Monteiro. Fuzzy Inference in Spatial Load Forecasting. Proceedings of IEEE Power Engineering Winter Meeting, 2:1063–1068, 2000.

[39] S.E. Skarman and M. Georgiopoulous. Short-Term Electrical Load Forecasting using a Fuzzy ARTMAP Neural Network. Proceedings of SPIE, 181–191, 1998.

[40] V.N. Vapnik. The Nature of Statistical Learning Theory. NewYork, Springer Verlag, 1995.

[41] N. Christiani and J.S. Taylor. An Itroduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, 2000.

[42] M. Mohandes. Support Vector Machines for Short-Term Electrical Load Forecasting. International Journal of Energy Research, 26:335–345, 2002.

[43] B.J. Chen, M.W. Chang, and C.J. Lin. Load Forecasting using Support Vector Machines: A Study on EUNITE Competition 2001. Technical report, Department of Computer Science and Information Engineering, National Taiwan University, 2002.

[44] Y. Li and T. Fang. Wavelet and Support Vector Machines for Short-Term Electrical Load Forecasting. Proceedings of International Conference on Wavelet Analysis and its Applications, 1:399– 404, 2003.

[45] Gencay R, Dechert W.D., (1992), "An Algorithm for the n-Lyapunov exponents of an n-dimensional unknown dynamical system" Physica D 59:142 57.

[46] Nychka D.W., Ellner S., (1992), "Finding chaos in a noisy system", J R Stat Soc B., 54:399-426.

[47] J. C. Sprott, "Chaos and Time-Series Analysis", Oxford University Press, USA.

[48] Wolf, A., Swift, J.B., Swinney, H., L., and Vastono, J.A., (1985), "Determining Lyapunov exponent from a time series", Physical D 16, PP. 285-317.

[49] Michael T. Rosenstein, James J. Collins, and Carlo J. De Luca," A practical method for calculating largest Lyapunov exponents from small data sets"

[50] M Rosenstein, J Collins, C Luca, A practical method for calculating largest Lyapunov exponents from small data sets, 1992

[51] G. W. Frank, T. Lookman, M. A. H. Nerenberg, C. Essex, J. Lemieux, and W. Blume, Chaotic time series analysis of epileptic seizures, Physica D 46 (1990) 427.

[52] P. Chen, Empirical and theoretical evidence of economic chaos, Sys. Dyn. Rev. 4 (1988) 81.

[53] P. Grassberger, and I. Procaccia, Characterization of strange attractors, Phys. Rev. Lett. 50 (1983) 346.

[54] P. Grassberger, and I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, Phys. Rev. A 28 (1983) 2591.

[55] S. Ellner, A. R. Gallant, D. McCaffrey, and D. Nychka, Convergence rates and data requirements for Jacobian-based estimates of Lyapunov exponents from data, Phys. Lett. A 153 (1991) 357.

[56] J. B. Ramsey, and H.-J. Yuan, The statistical properties of dimension calculations using small data sets, Nonlinearity 3 (1990) 155.

[57] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp, Singular-value decomposition and the Grassberger-Procaccia algorithm, Phys. Rev. A 38 (1988) 3017.

[58] J.-P. Eckmann, and D. Ruelle, Ergodic theory of chaos and strange attractors, Rev. Mod. Phys. 57 (1985) 617.

[59] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Determining Lyapunov exponents from a time series, Physica D 16 (1985) 285.

[60] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45 (1980) 712.

[61] V. I. Oseledec, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc. 19 (1968) 197.

[62] G. Benettin, L. Galgani, and J.-M. Strelcyn, Kolmogorov entropy and numerical experiments, Phys. Rev. A 14 (1976) 2338.

Fuzzy Inference System - Theory and Applications

Edited by Dr. Mohammad Fazle Azeem

ISBN 978-953-51-0525-1

Hard cover, 504 pages

Publisher InTech

Published online 09, May, 2012

Published in print edition May, 2012

This book is an attempt to accumulate the researches on diverse inter disciplinary field of engineering and management using Fuzzy Inference System (FIS). The book is organized in seven sections with twenty two chapters, covering a wide range of applications. Section I, caters theoretical aspects of FIS in chapter one. Section II, dealing with FIS applications to management related problems and consisting three chapters. Section III, accumulates six chapters to commemorate FIS application to mechanical and industrial engineering problems. Section IV, elaborates FIS application to image processing and cognition problems encompassing four chapters. Section V, describes FIS application to various power system engineering problem in three chapters. Section VI highlights the FIS application to system modeling and control problems and constitutes three chapters. Section VII accommodates two chapters and presents FIS application to civil engineering problem.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Zohreh Souzanchi Kashani (2012). A Multi Adaptive Neuro Fuzzy Inference System for Short Term Load Forecasting by Using Previous Day Features, Fuzzy Inference System - Theory and Applications, Dr. Mohammad Fazle Azeem (Ed.), ISBN: 978-953-51-0525-1, InTech, Available from: http://www.intechopen.com/books/fuzzy-inference-system-theory-and-applications/a-multi-adaptive-neuro-fuzzy-inference-system-for-short-term-load-forecasting-by-using-previous-day-

INTECH

open science | open minds