Electric Power Load Short Term Forecasting

Rae'd Basbous, Dr. Labib Arafeh
Al-Quds Open University, Palestine
Al-Quds University, , Palestine
rbasbous@qou.edu , larafeh@eng.alquds.edu

Abstract— At this study, two kinds of models have been developed, namely the Single Input Single Output, SISO, and Multiple Input Single Output, MISO. These two developed approaches depend on the Fuzzy based techniques including integrated and adaptive Neuro-Fuzzy approaches, and have been compared to represent the Short Term Load Forecasting, STLF, models.

Different models for SISO and MISO have been developed using the training data, such as, Sugeno Fuzzy Inference System with different optimization techniques including Hybrid and Back-propagation optimization techniques, Sugeno model using the Subtractive Clustering, and finally Sugeno cascaded model using Subtractive Clustering and Hybrid optimization technique.

The developed models have been integrated with a stand-alone application with Graphical User Interface, GUI. The developed Electric Power Load Forecasting System, EPLFS, can be accessed online to predict the power load.

The preliminary and promising results indicate the suitability and adequacy of the developed models depending on the Fuzzy approach to solve the short term load forecasting using the time and weather variables

***Keywords*—**ANFIS, STLF, Subtractive Clustering, Sugeno.

I. Introduction

Load forecasts (LF) is an important component of power system operation and planning involving prognosis of the future level of demand to serve as the basis for supply-side and demand side [1]. Precise load forecasting helps the electric utility to make unit commitment decisions, reduce spinning reserve capacity and schedule device maintenance plan properly [2].

LF can be divided into three main categories according to [3]. These categories are Long-Term Load Forecasting, LTLF, Mid-Term Load Forecasting, MTLF, and Short-Term Load Forecasting, STLF.

STLF cover an interval ranging from an hour to a week [3]. For STLF several factors should be considered, such as time factors, and weather data. STLF is important for different functions such as unit commitment, economic dispatch, energy transfer scheduling and real time control.

In this research, different modeling techniques for the short term load forecasting problem have been explored. Different measures have been used to check the adequacy of the developed models, these measures include the Correlation Coefficient (CC), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE).

Two kinds of models have been developed, namely the Single Input Single Output, SISO, and Multiple Input Single Output, MISO. These two developed approaches depend on the Fuzzy based techniques including integrated and adaptive Neuro-Fuzzy approaches, and have been compared to represent the STLF models. Real historical data profiles for two years (2006 and 2007) have been used to develop and test these proposed models. These data profiles were provided by Jerusalem District Electric Company (JDECO), and the Palestinian Meteorology Office (PMO).

A pre-processing stage has been accomplished for the collected data. The bad data and outliers in the collected data identified and removed. In order to develop and test the models, we have divided the data using a cross validation algorithm to training and testing datasets (75% of the available historical data profiles have been used for training and 25% for testing).

In Developing the SISO models the time parameter was considered as an input, while in MISO models three inputs have been considered namely, Time, High and

Rae'd Basbous is with AlQuds Open University, Jerusalem, Palestine; AlQuds Educational Area; e-mail: rbasbous@ qou.edu.

Dr. Labib Arafeh is with AlQuds University, Faculty of Engineering, Computer Engineering Department; e-mail: larafeh@eng.alquds.edu.

low Temperatures. For the two kinds of models, the power load was the output.

Different models for SISO and MISO have been developed using the training data, such as, Sugeno FIS with different optimization techniques including Hybrid and Back-propagation optimization techniques, Sugeno model using the Subtractive Clustering, and finally Sugeno cascaded model using Subtractive Clustering and Hybrid optimization technique to improve the models outcome. These models have developed using the Fuzzy toolbox in Matlab 2008.

These models have been integrated with a stand alone application with GUI called "Electric Power Load Forecasting System, EPLFS". This application has been developed using the Matlab 2008 GUI toolbox. Load forecasting can be done using this system, so one can load datasets saved in a text file, obtain the forecasted load using the developed models, plot the predicted power loads and the actual ones if known, and evaluate the predicted output by calculating the various measures used including the CC, MAPE and RMSE.

The EPLFS has been tested using the obtained power load historical profile for the year 2008 and used as the actual load. The system has been used to predict the load for one day and one week ahead using the developed models.

The organization of his paper is as follows. Section II describes the data provided. In addition, the analysis and of data and preprocessing is also presented. Section III presents and reviews the currently available fuzzy-based modeling techniques. Section IV covers the implementation and development steps that were followed to explore the use of soft computing approaches for STLF. The results of the developed models presented in section V. Section VI summarizes and concludes the paper.

II. Data Description

A. Data Sources

Developing any supervised-based soft computing model needs pairs of data (inputs and output), and in order to have a reliable STLF models that best represents the trends of these input and output data, we need reasonable actual sets of data composed of the electric power load as an output for a certain time during a day with known weather conditions as an inputs for a specific line that provides a chosen area.

The sources of the available datasets profiles are Jerusalem District Electric Company (JDECO) and Palestinian Meteorology Office (PMO). The provided power historical data profile includes the time and the corresponding power load at that time, while the weather historical data profile includes humidity, highest temperature, and lowest temperature for each day.

B. Data Preprocessing

The selection of the training datasets from the available data significantly affects the forecasting results, and to achieve a reliable and a more comprehensive approach to load forecasting, the days which have similar load and historical temperature values should be chosen to train (develop) the models [4].

Many factors affect the success of model training on a given task. The representation and quality of the instance data is first and foremost [5]. If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more difficult.

In our research, pre-processing stages have been accomplished for the collected data. These stages are as shown in Fig. 1. They are, obtaining historical profiles, input variables selection, bad data and outliers detecting and removing, time formatting, and cross validation.

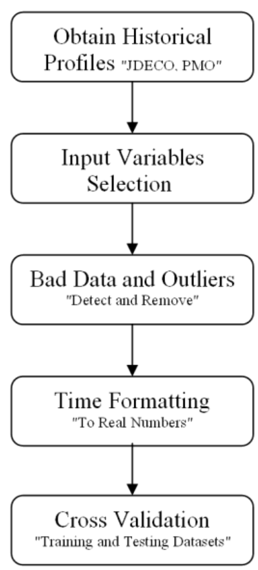


Fig. 1: Data Preprocessing Procedure Stages.

In the second stage, the time, power load, and temperature elements (low and high temperatures) have been selected to train the models. After that an existing algorithm for detecting and removing the outliers from the historical profiles has been used, and a manual procedure has been followed for detecting and removing the bad data such as the zero loads. Time formatting in the fourth stage is necessary since the input to the models should be a real number format and not in a time format (hour: minutes) as provided to us. Finally, a cross validation technique has been applied to divide the datasets into training and testing datasets.

Fig. 2 depicts a sample for detecting the outliers for (two months 7/2006, 7/2007) which was used to develop our July models before removing the outliers.

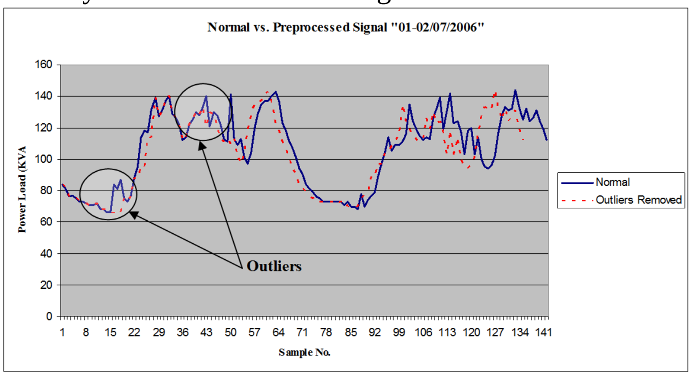


Fig. 2: The Original Dataset Before and After Removing the Outliers.

The figure shows the original signal for the first two days of the datasets (01-02/07/2006). Three clear outliers which have been detected and removed were circled.

C. *Extra Testing Datasets from the Year 2008*

In order to test our developed models using new unseen datasets, new profiles have been obtained for the year 2008 from JDECO and PMO for using the developed models to predict the power load for one day and one week. A small two samples selected to test the developed models. The first one is for one week from July (01-07/07/2008) to test the general July and Summer MISO models. The second datasets are for one week from May (0-07/05/2008) to test the general May and Spring MISO models.

III. Neuro Fuzzy Inference Systems

A. *Fuzzy Inference Systems*

Fuzzy Inference Systems (FISs) are also known as fuzzy rule-based systems [3], fuzzy model, fuzzy expert system, and fuzzy associative memory. This is a major unit of a fuzzy logic system. The decision-making is an important part in the entire system. The FIS formulates suitable rules and based upon the rules the decision is made.

The basic FIS can take either fuzzy inputs or crisp inputs, but the outputs it produces are almost always fuzzy sets. When the FIS is used as a controller, it is necessary to have a crisp output. Therefore in this case defuzzification method is adopted to best extract a crisp value that best represents a fuzzy set.

FIS consists of a fuzzification interface, a rule base, a database, a decision-making unit, and finally a defuzzification interface. A FIS with five functional block described in Fig. 3.

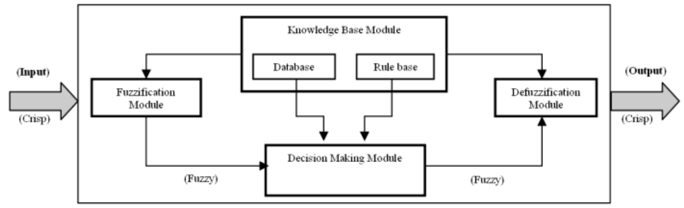


Fig. 3: Fussy Inference System [10].

The most important two types of fuzzy inference method are Mamdani's fuzzy inference method, which is the most commonly seen inference method. This method was introduced by Mamdani and Assilian in 1975 [6]. Another well-known inference method is the so-called Sugeno or Takagi–Sugeno–Kang method of fuzzy inference process. This method was introduced by Sugeno et al. in 1985 [7]. This method is also called as TS method.

A typical fuzzy rule in a Sugeno fuzzy model has the format [8]:

$$IF\ x\ is\ A\ and\ y\ is\ B\ THEN\ z = f(x, y) \qquad (1)$$

Where *AB* are fuzzy sets in the antecedent; $Z = f(x, y)$ is a crisp function in the consequent. Usually $f(x, y)$ is a polynomial in the input variables x and y, but it can be any other functions that can appropriately describe the output of the system within the fuzzy region specified by the antecedent of the rule.

For STLF a typical rule in a MISO Sugeno fuzzy model with three inputs (Time, HiTemp, LowTemp) and one output (PowerLoad), has the form [8]:

If Time is $Time_j$ *and Hi-Temp is* $HiTemp_k$ *and Low-Temp is* $LowTemp_l$*, then*

$$PowerLoad = p_i\, Time_j + q_i\, HiTemp_k + r_i LowTemp_l + s_i \qquad (2)$$

Where ($_j$) represent the time input MF, ($_k$) represent the high temperature input MF, and ($_l$) represent the low temperature input MF. The terms p_i, q_i, r_i, s_i, indicate the consequent parameters which determined through the training process.

B. Adaptive Neuro Fuzzy Inference Systems

In fuzzy modeling, the membership functions and rule base are generally determined by trial-and-error approaches. Although this approach is straightforward, the determination of best fitting boundaries of membership functions and number of rules are very difficult. In order to calibrate the membership functions and rule base in fuzzy modeling, the neural networks have been employed by researchers [9]-[14]. This system has been called fuzzy neural, neuro-fuzzy or adaptive network based system. The key properties of neuro-fuzzy systems are the accurate learning and adaptive capabilities of the neural networks, together with the generalization and fast-learning capabilities of fuzzy logic systems.

ANFIS constructs a fuzzy inference system (FIS) whose membership function parameters are adjusted using either a backpropagation algorithm alone, or in combination with a least squares method. This allows the fuzzy systems to learn from the data they are modeling. The purpose of ANFIS is to integrate the best features of Fuzzy Systems and Neural networks.

The Least-Squares (LSQ) optimization algorithm [15]-[17] is a mathematical optimization technique that attempts to find a function which closely approximates a given dataset. It tries to minimize the sum of the squares of the ordinate differences between points generated by the function and corresponding points in the dataset.

The Hybrid Learning (HL) algorithm [10] and [18], which combines the Gradient Descent (GD) and the LSQ algorithms, is one of the widely used algorithm in the literature to identify the parameters of the ANFIS. In the HL algorithm procedure, there are two passes which are called forward pass and backward pass. In the forward pass, functional signals go forward until the de-fuzzy layer and the consequent parameters are identified by the LSQ algorithm. In the backward pass, the error rates propagate backward and the premise parameters are updated by the GD algorithm.

Back-propagation learning algorithm, or propagation of error, is a common method of teaching artificial neural networks how to perform a given task. It is a very powerful method to adjust the weights of the neural network It was first described by Paul Werbos in 1971 which he published in his doctoral thesis [19], but it wasn't until 1986, through the work of [20], when Rumelhart et al, rediscovered this technique, that it gained recognition, and it led to a “renaissance” in the field of artificial neural network research.

C. Data Clustering

Data Clustering is considered an interesting approach for finding similarities in data and putting similar data into groups. Clustering partitions a data set into several groups such that the similarity within a group is larger than that among groups [18]. The idea of data grouping, or clustering, is simple in its nature and is close to the human way of thinking; whenever we are presented with a large amount of data, we usually tend to summarize this huge number of data into a small number of groups or categories in order to further facilitate its analysis.

Four of the most representative off-line clustering techniques frequently used in [3]:

1. K-means (or Hard C-means) Clustering,
2. Fuzzy C-means Clustering,
3. Mountain Clustering, and
4. Subtractive Clustering.

Mountain Clustering, proposed by Yager and Filev [18]. This technique calculates a mountain function (density function) at every possible position in the data space, and chooses the position with the greatest density value as the center of the first cluster. It then destructs the effect of the first cluster mountain function and finds the second cluster center. This process is repeated until the desired number of clusters has been found.

Subtractive clustering, proposed by Chiu [18]. This technique is similar to mountain clustering, except that instead of calculating the density function at every possible position in the data space, it uses the positions of the data points to calculate the density function, thus reducing the number of calculations significantly.

IV. Implementation and EPLFS Development

A. Implementation

In system modeling and identification, the important steps are to identify structure and parameters of the system based on the available data. The structure identification itself can be considered as two types, identification of the input variables of the model and the input–output relation. Most of the modeling approaches consider the input variables as a known priori and hence only the input and output relation has to be found [23].

In order to deeply study the effect of the temperature in the Short Term Load Forecasting (STLF), two kinds of models have been developed, SISO and MISO models. For the SISO models that shown in Fig. 4 the time has been used as the input for the models and the power load at that time has been used as the output.

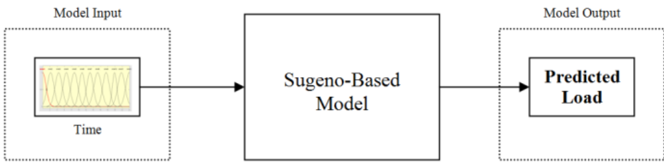


Fig. 4: SISO Sugeno FIS Model Architecture

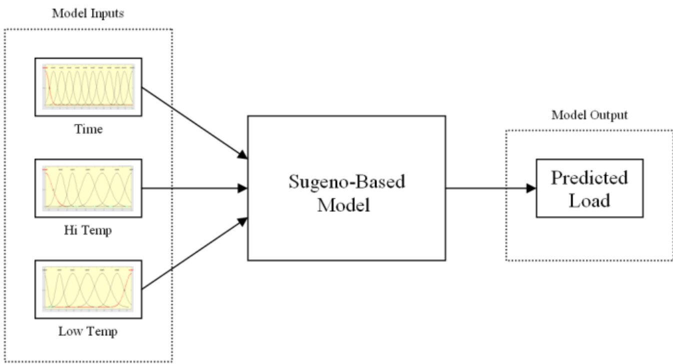


Fig. 5: MISO Sugeno FIS Model Architecture

In the MISO models which shown in Fig. 5, three variables (Time (T), High Temperature (HT), and the Low Temperature (LT) for that day) have been used as an input for the developed models, and the power load at that time was considered as the model output.

The results in Table I show that the input variable temperatures have a considerable effect on load forecasting. The results in the table are for the two kinds of models, SISO and MISO Sugeno models with hybrid optimization. The adequacy of the two developed models has been checked using the CC and two error measures; MAPE, and the RMSE.

Table I: Temperature Effect on Load Forecasting.

Type of Model	No. of MFs	Dataset	Errors Measures		
			CC	*MAPE*	*RMSE*
MISO	12 7 7	Training	0.9815	0.0257	0.0769
		Testing	0.9719	0.0324	0.1715
SISO	12	Training	0.9106	0.0667	0.1662
		Testing	0.9085	0.0656	0.3045

The proposed models are to be trained with the obtained historical data profiles from JDECO and PMO before testing them. The first step for training a model is obtaining an accurate historical data. In addition, data should be chosen that is relevant to the model. Several models have been developed such as, Sugeno with different optimization techniques, Subtractive Clustering, and finally Subtractive Clustering cascaded with Hybrid optimization technique to improve the developed models.

Fig. 6 illustrates a general developing "training" block diagram of our models. It consists of three main stages.

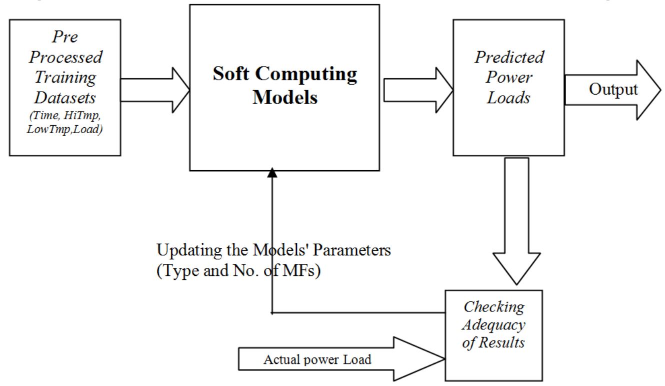


Fig. 6: A General Block Diagram for Developing/Training Soft Computing Models

These stages can be summarized as follow:

1. The first stage is pre-processing the input datasets for the system. These datasets include four elements; three of them are inputs; namely, the time, the high temperature, and the low temperature of the day and one output (the actual loads).
2. The second stage is concerned with various soft computing models that have been developed. SISO/MISO models will be developed using different techniques (Hybrid and Back-propagation optimization techniques, Subtractive Clustering and finally by cascading two models). The same datasets (Training and Testing) have been used in developing all the models.
3. The third stage checks the adequacy of the developed models to demonstrate their performance.

Three measures according to Basbous in [3] have been used to effectively check the adequacy of results, and these measures illustrated bellow:

1. *The CC measure between actual and predicted power loads*. It indicates the strength and direction of a linear relationship between the forecasted and actual loads and calculated by [8]:

$$CC_{xy=}\sqrt{1-\frac{\sum_{i=1}^{N}(y_i-x_i))^2}{\sum_{i=1}^{N}(y_i-y)^2}} \quad ,(3)$$

where y_i : is the ith actual data,

y : is the average of all actual data,

x_i : is the ith predicted data.

N : is the number of data points under consideration.

2. *The Mean Absolute Percentage Error (MAPE)*, which has been traditionally used to measure accuracy in load forecasting [2]. It captures the proportionality between the forecast error and the actual load. The MAPE is calculated by [2]:

$$\text{MAPE}= \sum_{i=1}^{N}\left|\frac{y_i - x_i}{y_i}\right| * \frac{100}{N}\% \qquad ,(4)$$

3. The Root Mean Square Error (RMSE), which is used to evaluate the error (differences) between the forecasted and actual loads. The general form of the RMSE equation for the actual power loads (Y) and the predicted ones (X) in given by [24]:

$$\text{RMSE}= \frac{\sqrt{\sum_{i=1}^{N}(y_i - x_i)^2}}{(N-1)} \qquad ,(5)$$

In order to obtain the best results from the developed models, the model parameters (type and number of membership functions) need to be updated and determined manually to be fixed for the all models as shown in Fig. 6. The predicted loads will be measured against the actual marks and the system parameters will be altered to find the best outcome. We have been tried several Membership Functions (MFs) including: Gaussian Curve, Generalized Bell, Trapezoidal and Triangular. Table II lists the results that obtained form a July SISO model with different MFs and the same parameters (No. of MFs, hybrid optimization technique, training datasets and testing datasets).

Table II: Results for a July SISO Model with Different Types of MFs.

Type of MF	No. of MFs	Dataset	Training Dataset Errors		
			CC	*MAPE*	*RMSE*
Gaussian Curve	12	Training	0.9104	0.0667	0.1664
		Testing	0.9094	0.0654	0.3032
Trapezoidal		Training	0.9100	0.0669	0.1667
		Testing	0.9088	0.0656	0.3041
Triangular		Training	0.9094	0.0672	0.1672
		Testing	0.9079	0.0661	0.3055
Generalized Bell		Training	0.9106	0.0667	0.1662
		Testing	0.9085	0.0656	0.3045

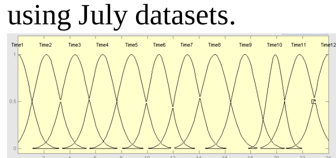
As listed in Table II, there is no major difference between the outputs (predicted loads) regarding the type of the membership function. However, the MF that produced the best results is found to be the Generalized Bell (GBell).

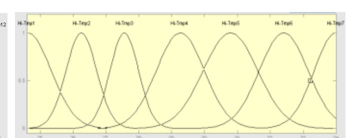
Table III lists the results that obtained form a July MISO model with different number of MFs and the same parameters (Type of MFs (GBell), hybrid optimization technique, training datasets and testing datasets).

Table III: Results for a July MISO Model with Different Number of MFs.

No. of MFs	Dataset	Errors		
		CC	*MAPE*	*RMSE*
6 6 6	Training	0.9252	0.0593	0.1526
	Testing	0.9307	0.0570	0.2666
8 6 6	Training	0.9511	0.0466	0.1242
	Testing	0.9528	0.0449	0.2212
12 6 6	Training	0.9521	0.0452	0.1229
	Testing	0.9511	0.0448	0.2250
12 7 7	Training	***0.9815***	***0.0257***	***0.0769***
	Testing	***0.9719***	***0.0324***	***0.1715***
12 8 8	Training	0.9561	0.0428	0.1179
	Testing	0.9548	0.0440	0.2167

It is clear from Table III that the best results obtained when we have been used 12 MFs for the time input, and 7 MFs for the temperature inputs (High and Low). According to the results shown in Table III, we will fix the number of the MF's in the proposed models to 12 MF's for the time input and 7 MF's for the Low and High temperatures inputs. Fig. 7 represents the inputs MFs that we have been used in building a MISO model using July datasets.


(a): Time Input MFs


(b): High Temperature Input MFs

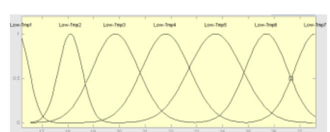

(c): Low Temperature Input MFs

Fig. 7: MISO Model Inputs (Time, High Temperature, Low Temperature) MFs

Table IV shows the results obtained from a two July MISO Sugeno model with hybrid optimization the first one developed using all the training datasets available.

Table IV: Results for a July MISO Model Before and After Removing the Outliers.

Datasets Used	Dataset	Errors		
		CC	*MAPE*	*RMSE*
All the Datasets	Training	0.9414	0.0510	0.1336
	Testing	0.9255	0.0568	0.2585
Outliers Removed	Training	0.9815	0.0257	0.0769
	Testing	0.9719	0.0324	0.1715

The second one developed using the training datasets after removing the outliers. The results obtained show the effect of the outliers to the accuracy of the models. The model that has been developed before removing the outliers shows a CC value between the actual and predicted loads for the training datasets of 0.9414 while the CC value for the model that has been developed after removing the outliers is 0.9815.

Using the parameters that give us the best results and the outliers have been removed from the datasets, two SISO and MISO models have been developed using datasets from the month of July and May. In addition to that, two SISO and MISO models have been developed using datasets from Spring (month of April and May) and Summer (month of July and August) seasons. Several models with different optimization techniques have been developed for these types of models.

B. Stand Alone Application " EPLFS"

These models were integrated within a stand alone application with GUI. The developed Electric Power Load Forecasting System "EPLFS" is as shown in Fig. 8; the figure demonstrates the forecasted power loads for a testing datasets. The three lists in the system present the times, forecasted loads, and the actual loads. Three different measures appear in the bottom of the right corner, the CC, MAPE, and RMSE.

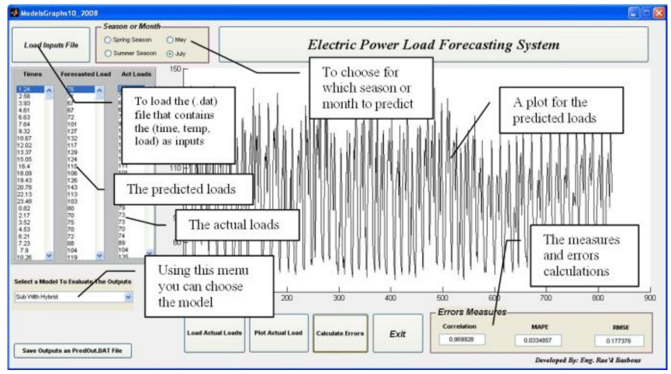


Fig. 8: The EPLF System: Showing a Plot for the Predicted Loads in a Certain Hours.

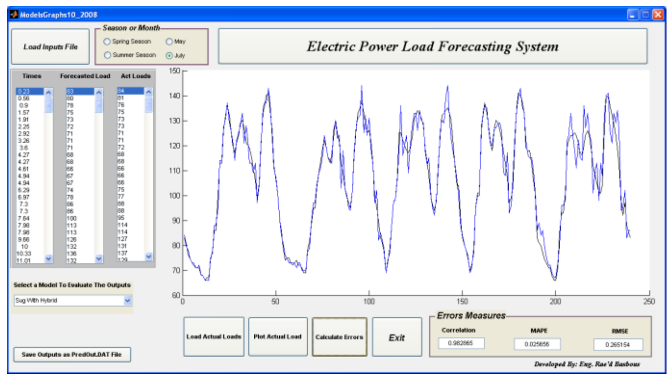


Fig. 9: The EPLF System: Showing a Plot for the Predicted Loads vs. Actual Ones.

Using EPLFS we can load the datasets, evaluate the predicted output using the developed models, plot the actual and predicted load, and calculate several measures including the CC, MAPE and RMSE.

Fig. 8 represents a snapshot for the EPLFS when used to predict the power loads for certain times with temperatures parameters.

Fig. 9 represents a snapshot for the EPLFS when used to predict the power loads for certain times with temperatures parameters. A plot for the actual and predicted power loads are seen in the snapshot. The blue dotted line represents the actual loads and the black continuous one represents the predicted power loads.

To load an input and actual load files to the EPLFS, it should be in (.DAT) format. For the input file it should be with three columns. The first column representing the times formatted as mentioned in the previous chapter. The second and third columns representing the temperature parameters (High and Low) respectively. For the file that containing the actual loads to be compared with the predicted ones, the loads should be in one column.

V. Results and Comparisons

In this research and using the available historical datasets profiles, we have been started by developing eight Sugeno models with hybrid optimization technique. Four of the models are SISO models and the other four are MISO models. These models have been used to predict the power load in specific months (May, July) or general model to be used in seasons (Spring, Summer). Another eight models have been developed using the same datasets but with the back-propagation optimization technique. After that the Subtractive clustering has been used to construct a Sugeno models for the same months and seasons. Then, the Subtractive clustering with the Hybrid optimization technique have been used to construct a cascaded model to improve and enhance the results obtained from the previous models.

As mentioned before, the training data sets and the models parameters (number and type of MF, number of rules, and cluster radius) have been fixed and used for the proposed models. For example 12 MF for the Time input and 7 MF for the Low and High temperatures of the type G-Bell have been fixed for all the models with Hybrid and Back-propagation optimization techniques. Furthermore, for the models that have been constructed using the Subtractive Clustering, a cluster radius of the value 0.1 has been fixed and used to construct these models.

A. Results of the Developed Models

Table V below lists the average CC measures for all the developed models using Hybrid, Back-propagation, Subtractive Clustering and Subtractive clustering with Hybrid optimization (cascaded model). It is clear from the table that the cascaded model has produced the best results with an average equal to (0.97) for the MISO models followed by the models that have been developed using the subtractive clustering with an average correlation equal to (0.95). The results of the cascaded models have been obtained using less number of rules compared with the models that have been developed using the Hybrid and Back-propagation optimization techniques.

Table V: The Correlations Measures Average for all the Developed Models

Optimization Method	SISO		MISO	
	Training	*Testing*	*Training*	*Testing*
Hybrid	0.87770	0.87323	0.95483	0.94928
B-Prop	0.82280	0.81780	0.90940	0.90448
Subtractive	0.87880	0.87428	0.95925	0.94813
Sub. with Hybrid	0.87935	0.87345	0.97238	0.96158

The developed cascaded model with cluster radius equal to 0.1 produced 313 rules, while with Hybrid and Back-propagation optimization techniques the number of rules that have been produced equal to the multiplication of inputs membership functions (12 MFs for the time * 7 MFs for the Hi-Temp * 7 MFs for the Low-Temp) which is equal to 588 rules.

In addition to the CC, two different error measures, the MAPE and the RMSE have been used to examine and show the adequacy of the developed models and its outcome. The CC measures the agreements between the actual and predicted power loads, while the error measures RMSE and MAPE give an indication how the performance of the developed models are.

Table VI shows the average results of the error measure MAPE for all the models. It is clear from the table that the lowest values are for the MISO models developed using the cascaded models with an average of the MAPE equal to (0.03). This low value reflects the highest CC that achieved from these models as shown in the Table V. You can notice that the MISO models have the lowest MAPE values over the SISO models because of the temperature parameters effect on the power load. The same thing has been noticed in the CC measures since the MISO models produced the best results and have the highest CC values.

Table VI: The Average MAPE Measures for all the Developed Models

Optimization Method	SISO		MISO	
	Training	*Testing*	*Training*	*Testing*
Hybrid	0.08685	0.08793	0.05068	0.05410
B-Prop	0.10453	0.10510	0.07778	0.07735
Subtractive	0.08618	0.08738	0.04623	0.05335
Sub. with Hybrid	0.08603	0.08763	0.03878	0.04668

Table VII shows the average results of the error measure RMSE for all the models. This measure show the adequacy of the developed models too in addition to the MAPE measures. The same thing for the RMSE results as in the MAPE results achieved where the cascaded model have the best results (the lowest RMSE values). The developed MISO models with the temperature parameters produce the lowest RMSE values over the SISO models. It is clear from the table and the figure that the lowest values are for the MISO models developed using the cascaded models with an average of the RMSE equal to (0.08).

Table VII: The Average RMSE Measures for all the Developed Models

Optimization Method	SISO		MISO	
	Training	*Testing*	*Training*	*Testing*
Hybrid	0.17580	0.31280	0.10505	0.19883
B-Prop	0.20608	0.36710	0.15453	0.27873
Subtractive	0.17463	0.31113	0.09908	0.18628
Sub. with Hybrid	0.17423	0.31240	0.08310	0.16485

A graphical representation for the average CC and error measures (RMSE and MAPE) are shown in Fig. 10. The three measures are combined together in the same graph to take a clear look to the behavior of these measures. A relation can be concluded from the graph which is: an increasing in the CC leads to decrease in the error measures (RMSE and MAPE). As an example for this, from the above tables when the average CC for the MISO cascaded models are equal to (0.97) the corresponding average MAPE and RMSE values for the cascaded model are equal to (0.03 and 0.08) respectively. The second case is when the average CC measure for the SISO cascaded models is equal to (0.87) the corresponding average error measures MAPE and RMSE for the cascaded model are equal to (0.08 and 0.17) respectively.

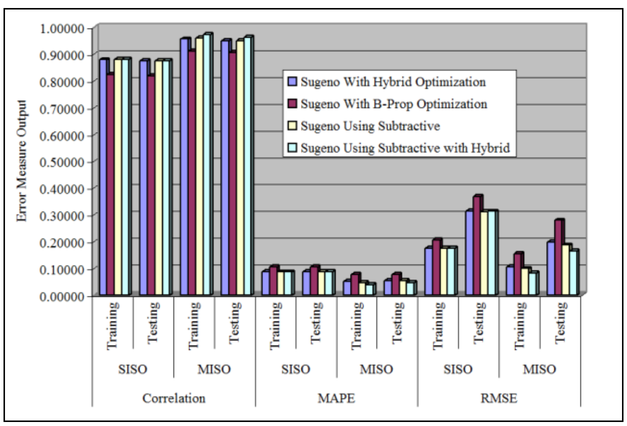


Fig. 10: The Average Correlation Measures Against the Error Measures (MAPE and RMSE) for all the developed SISO and MISO models

Fig. 10 can be summarized by the following points:

1. The developed cascaded models (Subtractive Clustering with Hybrid optimization) produced the highest results. The average CC between the actual and predicted loads ranging between 0.95 and 0.98 with average equal to 0.97.
2. The developed models with Back-propagation optimization techniques have the lowest CC compared to the other models and similarly for the two error measures. The correlation coefficient ranging between 0.89 and 0.91 with average equal to 0.90.
3. The developed models using the Subtractive Clustering can be enhanced when subject to the Hybrid optimization technique. The average correlation for the developed models using the Subtractive Clustering is equal to 0.95 and after applying the Hybrid optimization technique it is enhanced to an average equal to 0.97.
4. Finally, we can notice that the highest the CC the lowest the MAPE and RMSE. For example, the average CC of the cascaded model is equal to 0.97, the average MAPE is equal to 0.03, and the average RMSE is equal to 0.08. While, the average CC of the developed models using the Back-propagation optimization technique is equal to 0.90, the average MAPE is equal to 0.07, and the average RMSE is equal to 0.15.

B. One Day and One Week Ahead Prediction Using Unseen Datasets from the Year 2008

As mentioned in chapter four, new historical profiles have been obtained from JDECO and PMO for the year 2008. The selected datasets (one day and one week from July and May) have been used to test the EPLFS for new unseen datasets. These datasets have not been considered in the cross validation process that has been applied in developing our models. The first day of May and July (01/05, 01/07) has been chosen in order to use the system to predict the load for one day. To test the system for predicting the load for a period of one week the first week of May and July (01-07/05, 01-07/07) has been considered.

These datasets have been not used in the cross validation for developing the models. The average CC that has been obtained for one day prediction is equal to 0.94 and the average MAPE is equal to 0.058. The average correlation in case of one week prediction is equal to 0.93 while the average MAPE is equal to 0.059. Table VIII and Table IX show the average correlation and error measures for one day and one week power load prediction using the developed models.

From these two tables it is clear that the best results have been obtained from the general May model in case of one day and one week prediction. The lowest average CC has been obtained from the general Spring model because of the wide range of the model input parameters (the wide variations of low and high temperatures).

These results show the accuracy of the developed models to predict the power loads for the new unseen datasets one day and one week ahead.

Table VIII: The Average CC and Error Measures for One Day Prediction

Model	*CC*	*MAPE*	*RMSE*
July	0.9680	0.0407	0.2242
Summer	0.9435	0.0506	0.3187
May	0.9794	0.0350	0.1902
Spring	0.8738	0.1078	0.4874
Average	***0.9412***	***0.0585***	***0.3051***

Table IX: The Average CC and Error Measures for One Week Prediction

Model	*CC*	*MAPE*	*RMSE*
July	0.9428	0.0533	0.2872
Summer	0.9255	0.0576	0.3513
May	0.9578	0.0464	0.2850
Spring	0.9047	0.0809	0.4308
Average	***0.9327***	***0.0595***	***0.3385***

Fig. 11 and Fig. 12 below show the results obtained from the EPLFS for one day and one week prediction using the developed models.

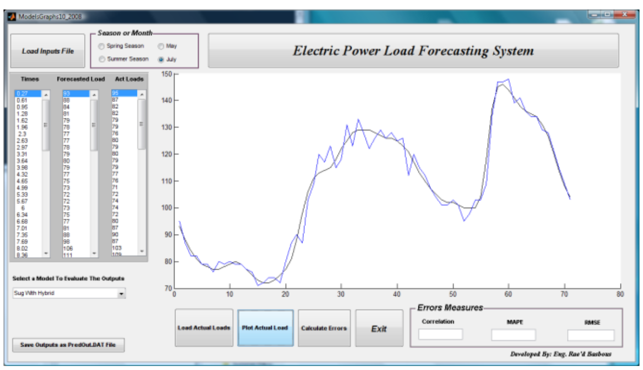


Fig. 11: One Day Prediction using the EPLFS for New Unseen Datasets.

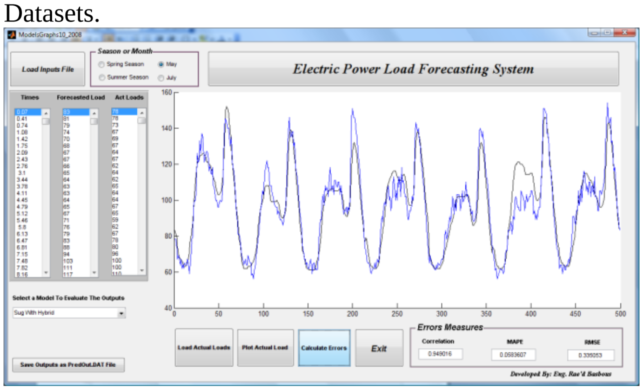


Fig. 12: One Week Prediction using the EPLFS for New Unseen Datasets.

C. Comparison with Other Studies

A plenty of works can be found in the STLF field. Some of these works are mentioned here briefly. Many papers that have been published recently in the refereed journals are considered and the ones whose main interests are STLF by soft computing methods are taken into account.

It is important to mention that different datasets and different approaches have been used in developing these models. However, we are trying to compare the obtained average CC and MAPE as a measure of errors. Furthermore, the same equations of the CC and MAPE that have been presented in (3) and (4) used in these papers to check the adequacy of the developed models. All the CC and MAPE results listed as a real number with fractions instead of using the percentage sign.

Hwang [25] described a new practical knowledge-based expert system (called LoFY) for short-term load forecasting equipped with graphical user interfaces. Also, various forecasting models like trending, multiple regression, artificial neural networks, fuzzy rule-based model, and relative coefficient model have been included to increase the forecasting accuracy. The simulation based on historical sample data shows that the forecasting accuracy is improved when compared to the results from the conventional methods. Through the fuzzy rule-based approach, the forecasting accuracy at special days has been improved remarkably. The average MAPE results found for this system is 0.020.

Khan [1] presented a comparative study of six soft computing models namely multilayer perceptron networks, Elman recurrent neural network, radial basis function network, Hopfield model, fuzzy inference system and hybrid fuzzy neural network for the hourly electricity demand forecast of Czech Republic. The soft computing models were trained and tested using the actual hourly load data obtained from the Czech Electric Power Utility for seven years (January 1994 – December 2000). A comparison of the proposed techniques is presented for predicting 48 hourly demands for electricity. Simulation results indicate that hybrid fuzzy neural network and radial basis function networks are the best candidates for the analysis and forecasting of electricity demand for the experimented data, with the following MAPEs: For weekday forecast, 0.010 by radial basis function networks, 0.009 by fuzzy neural network; and for weekend forecast, 0.013 by radial basis function networks, 0.020 by fuzzy neural network.

A modeling technique based on the fuzzy curve notion is proposed by Papadakis [26] to generate fuzzy models for STLF. Different forecast models are developed for each day type in every season. The model is considered as a fuzzy neural network described in terms of a parameter vector and is trained using a genetic algorithm with enhanced learning and accuracy attributes. The performances of the developed fuzzy models are tested using load data of the Greek interconnected power system. They achieve a MAPE of 0.0167 with the data from year 1995.

A feed-forward neural network with a back-propagation algorithm is presented by Bhattacharyya [27] for three types of short-term electric load forecasting: daily peak (valley) load, hourly load and the total load. The forecast has been made for the northern areas of Vietnam using a large set of data on peak load, valley load, hourly load and temperature. The data were used to train and calibrate the artificial neural network, and the calibrated network was used for load forecasting. The results obtained from the model show that the application of neural network to short-term electric load forecasting problem is very useful with quite accurate results. The method has given the best performance with 0.9427 average CC and 0.108 average MAPE.

A Fuzzy Logic (FL) expert system is integrated with Artificial Neural Networks (ANN) for a more accurate short-term load forecast is presented by Tamimi [28]. The 24 hour ahead forecasted load is obtained through two steps. First, a FL module maps the highly nonlinear relationship between the weather parameters and their impact on the daily electric load peak. Second, 12 ANN modules are trained using historical hourly load and weather data combined with the FL output data, to perform the final forecast. Comparisons made between this model, an ANN model, and an Autoregressive Moving Average (ARMA) model were show the efficiency and accuracy of this new approach. The average MAPE for these methods is found equal to 0.029.

From the results mentioned above, it is clearly noticed that the soft computing methods provide a promising solution to the STLF problem. In addition, combining or integrating more than one method together leads to an enhancement to the proposed models. For example Tamimi [28] combined the NN with FL, and [25] developed a forecasting system and include it with different forecasting models to increase the forecasting accuracy. Furthermore, the lowest results over the proposed models that haven mentioned above found in the models with Back-propagation optimization proposed by Bhattacharyya [27]. This agrees with the results obtained from our developed Sugeno FIS models with Back-propagation optimization.

Comparing our results with these systems we can see that our developed models produced satisfactory results using the temperature parameters only to predict the electric load without taking in account the other weather parameters or the type of the day or any other conditions. In our developed models, the CC for one day ahead prediction for the unseen datasets from the year 2008 ranges between (0.87 and 0.97) with an average value 0.94; the corresponding MAPE ranges between (0.03 and 0.10) with an average 0.05. Whereas; the obtained CC for one week ahead prediction for the same datasets ranges between (0.90 and 0.95) with an average value 0.93; the corresponding MAPE ranges between (0.04 and 0.08) with an average 0.05.

An improvement to the results that have been obtained from our developed models can be achieved when a hourly temperature and weather data profiles are available. Also, an average high and low temperature of the day which has been considered in the MISO models leads to reduce the error measures between the actual and predicted power loads compared to SISO models.

VI. Summery and Conclusion

The general objective of this work is to explore the use of soft computing and artificial intelligence approaches to develop Short Term Load Forecasting (STLF) system that predict the power load for one day up to one week ahead in specific month or specific season. The introduction of NF modeling approaches to the area of load forecasting has been presented. The basic concepts of ANFIS, Hybrid Learning, BP Learning, and Data Clustering have been reviewed in the previous sections. The implementation of and NF approaches to model the relationships between Temperature, Time, and Power Load has been introduced.

Real JDECO power line in Bier Nabala village and PMO historical data profiles for two years (2006 and 2007) have been collected and used to develop and test the various models. Developing the models for load forecasting has been applied firstly using the available datasets. Two kinds of models have been developed, Single Input Single Output (SISO) models, and Multiple Inputs Single Output (MISO) models. Three main inputs (Time (T), High Temperature (HT), and Low Temperature (LT)) and one output (Power Load (PL)) have been considered in building these models. For the

SISO models, only the time has been considered as the input for the models and the power load at that time has been used as the output.

It has been found that the temperature is a major input parameter on STLF. Models, that do not utilize temperature measurements in training, produce quite larger errors than the ones exploiting them as input parameters. The correlation has been improved from 91% for the SISO model to about 98% for the MISO model as demonstrated in Table (5.1) using the same parameters (number of MFs, type of MFs, and cluster radius). These results clearly reveal the effect of the temperature parameters on predicting the power load.

Fuzzy Inference System (FIS) with different optimization techniques have been used to develop our models. Firstly we started by developing a SISO and MISO models using ANFIS with hybrid optimization technique. Then a SISO and MISO models have been developed using Back-propagation optimization technique. After that the Subtractive clustering has been used to develop such models. Finally cascaded models have been developed for STLF by constructing the models using the Subtractive clustering and then learning these models using the hybrid optimization techniques to achieve more accurate models.

While testing these models using the testing datasets that has been isolated before the training stage using the developed cross validation algorithm, the average obtained CC between the actual and predicted power loads for all the developed SISO models ranges between (0.81 and 0.87), with an average CC value 0.85; The corresponding average MAPE that ranges between (0.08 and 0.10) with an average value of 0.09, and average RMSE that ranges between (0.31 and 0.36), with an average value of 0.32. Whereas; the obtained average CC for the MISO models ranges between (0.90 and 0.96) with an average CC value 0.94; The corresponding average MAPE that ranges between (0.04 and 0.07) with an average value of 0.05, and average RMSE that ranges between (0.16 and 0.27) with an average value 0.20. This demonstrates the adequacy of adopting these types of approaches to STLF problem as well as the improvements of models' forecasting performance when taking the weather into consideration.

It was noticed that the forecasting performance has been furtherlly improved by the MISO cascaded models, while maintaining all other factors including MFs types and numbers, and cluster radius. This improvement is notices as an improvement in the obtained CC that results from the MISO cascaded models which ranges between (0.95 and 0.97) and with an average CC value of 0.96 for all the developed models as compared with the forecasting performance of CC that ranges between (0.90 and 0.94) when developing the models using the other optimization techniques; The corresponding MAPE ranges between (0.03 and 0.06) with an average value 0.04, and RMSE ranges between (0.07 and 0.20) with an average value 0.16 compared to average MAPE ranges between (0.05 and 0.07), and average RMSE ranges between (0.18 and 0.27) for the other optimization techniques.

These models have been integrated with a stand alone application with GUI. The developed Electric Power Load Forecasting System "EPLFS" can be accessed online through either a local area network, or using a web server. The EPLFS has been tested using the obtained power load historical profile for the year 2008 and used as the actual load. The system has been used to predict the load for one day and one week ahead using the developed models. The CC for one day ahead prediction ranges between (0.87 and 0.97) with an average value 0.94; The corresponding MAPE ranges between (0.03 and 0.10) with an average 0.05, and RMSE ranges between (0.19 and 0.48) with an average 0.30. Whereas; the obtained CC for one week ahead prediction ranges between (0.90 and 0.95) with an average value 0.93; The corresponding MAPE ranges between (0.04 and 0.08) with an average 0.05, and RMSE ranges between (0.28 and 0.43) with an average 0.33.

Finally, different works in the field of STLF using different techniques accomplished by other researchers have been compared with our developed models. These works show the ability of the soft computing techniques to represent the STLF, and agree with our results that the Back-propagation optimization technique produced the lowest results. Our over all results indicates the suitability and adequacy of the developed models to solve the short term load forecasting problem using the time and weather variables.

Acknowledgment

The authors wish to thank Eng. Mansour Nassar and Eng. Salah Alqam from the strategic planning department in Jerusalem District Electricity Company (JDECO) for opening up the load consumption database for this scientific research and also to Mr. Isam Al-Seifi from Palestinian Meteorology Organization (PMO) for providing the temperature measurements.

References

[1] Khan, M., Abraham, A., Ondrusek, C. (2001), 'Soft Computing for Developing Short Term Load Forecasting Models in Czech Republic', Hybrid Information Systems, pp. 207-222, 2001.

[2] McSharry, P. (2006), 'Evaluation of Short-Term Forecasting Methods for Electricity Demand in France', RTE-VT workshop, Paris, May 29-30, 2006.

[3] Basbous, R. (2009), 'Electric Power Load Short Term Forecasting', M.Sc. Thesis, Al Quds University, 2009.

[4] Peng, M., Hubele, N., Karady, G. (1992), 'Advancement in the Application of Neural Networks for Short-Term Load Forecasting', IEEE Transactions on Power Systems, Vol.[7], pp. 250 - 257 , 1992.

[5] Pyle, D. (1999), Data Preparation for Data Mining, Morgan Kaufmann Publishers, Los Altos, CA, 1999.

[6] Mamdani, E., Assilian, S. (1975), An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, International Journal of Man-Machine Studies, Vol. [7], No.(1), pp.1–13, 1975.

[7] Sugeno, M., Takagi, T. (1985), Fuzzy Identification of Systems and its Application to Modeling and Control, IEEE Transactions on Systems and Machine Cybernetics, Vol. [15], pp.116–132, 1985.

[8] Arafeh, L., Singh, S., Putatunda, S. (1999), 'A Neuro Fuzzy Logic Approach to Material Processing', IEEE Transactions on Systems, Man, and Cybernetics, Part C, Vol. [29], No.(3), pp. 362-370, 1999.

[9] Jang, J. (1992), Neuro-Fuzzy Modeling: Architectures, Analyses and Applications, Ph.D. Thesis, University of California, Berkeley, 1992.

[10] Jang, J. (1993), ANFIS: Adaptive-Network-Based Fuzzy Inference System, IEEE Transactions on Systems and Machine Cybernetics, Vol. [23], No.(3) pp. 665–685, 1993.

[11] Rong, L., Wang, Z. (1996), An Algorithm of Extracting Fuzzy Rules Directly from Numerical Examples by Using FNN, Proceedings of the IEEE International Conference on Systems and Machine Cybernetics, Beijing, China, pp. 1067–1072, 14–17 October 1996.

[12] Ouyang, C., Lee, S. (2000), A Hybrid Algorithm for Structure Identification of Neuro-Fuzzy Modeling, Proceedings of the IEEE International Conference on Systems and Machine Cybernetics, Nashville, Tennessee, pp. 3611–3616, 8–11 October, 2000.

[13] Rojas, I., Pomares, H., Ortega, J., Prieto, A. (2000), A Self-Organized Fuzzy System Generation from Training Examples, IEEE Transactions on Fuzzy Systems, Vol. [8], 23–36, 2000.

[14] Guler, I., Ubeyli, E. (2004), Application of Adaptive Neuro-Fuzzy Inference System for Detection of Electrocardiographic Changes in Patients with Partial Epilepsy using Feature Extraction, Expert Systems with Applications, Vol. [27], pp. 323–330, 2004.

[15] Levenberg, K. (1944), A Method for the Solution of Certain Nonlinear Problems in Least Squares, Quart Appl Math , Vol. [2], pp. 164–168, 1944.

[16] Marquardt, D. (1963), An Algorithm for Least-Squares Estimation of Nonlinear Parameters, Journal of the Society for Industrial and Applied Mathematics., Vol. [11], pp. 431–441, 1963.

[17] Dennis, J. (1977), Nonlinear Least-Squares, State of the Art in Numerical Analysis, Academic Press, Orlando, FL, 1977.

[18] Jang, J., Sun, C. and Mizutani, E. (1997), Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, Prentice-Hall, Englewood Cliffs, NJ, Chapter One, Chapter Eight, and Chapter Twelve, pp. 7-9 and pp. 219-336, 1997.

[19] Werbos, P. (1974), Beyond Regression: New Tools for Prediction and analysis in the behavioral sciences, PhD Thesis, Harvard University, 1974.

[20] Rumelhart, D., Hinton, G., Williams, R. (1986), Learning Internal Representations by Error Propagation, Parallel Distributed Processing - Explorations in the Microstructure of Cognition, Vol. [1], Chapter 8, pp. 318-364, The MIT Press: Cambridge , 1986.

[21] Karayiannis, N. (1996), Hybrid Learning Schemes for Fast Training of Feed-Forward Neural Networks, Department of Electrical and Computer Engineering, University of Houston, USA, Mathematics and Computers in Simulation, Vol. [41], pp. 13–28, 1996.

[22] Hartigan, J., Wong, M. (1979), "A K-Means Clustering Algorithm," Applied Statistics, Vol. [28], pp. 100-108, 1979.

[23] Satish, B., Swarup, K., Srinivas, S., Hanumantha, A. (2004), Effect of Temperature on Short Term Load Forecasting using an Integrated ANN, Electric Power Systems Research, Vol. [72], pp. 95-101, 2004.

[24] Oriqat, Y. (2007), 'Modeling Techniques Applied to Short Essay Auto-Grading Problem', M.Sc. Thesis, Al-Quds University, pp. 46-47, 2007.

[25] Hwang, K., Kim, G. (2001), 'A Short-Term Load Forecasting Expert System', Proceedings of The Fifth Russian-Korean International Symposium on Science and Technology, pp. 112 – 116, 2001.

[26] Papadakis, S. (1998), A Novel Approach to Short-Term Load Forecasting using Fuzzy Neural Network, IEEE Transactions on Power Systems, Vol. [13], No. (2), pp. 480–492, 1998.

[27] Bhattacharyya, S., Thanh, L. (2004), Short-Term Electric Load Forecasting Using an Artificial Neural Network: Case of Northen Vietnam, International Journal of Energy Research, Vol. [28], pp. 463-472, 2004.

[28] Tamimi, M., Egbert, R. (2000), Short Term Electric Load Forecasting via Fuzzy Neural Collaboration, Electric Power Systems Research, Vol. [56], pp. 243–248, 2000.

Rae'd Basbous received the B.Sc. Degree in Computer Systems Engineering from Palestine Polytechnic University, Palestine, in 2000, and the M.S. Degree in Electronic and Computer Engineering from AlQuds University, Palestine, in 2009.

Labib Arafeh, Associate Prof. in Computing, Held his PhD from University of Manchester Institute of Science & Technology in 1992, M.Sc. from University of Essex in 1984, and B.E. from Bangalore University in 1980.