ELECTRICITY CONSUMPTION IN THE INDUSTRIAL SECTOR OF JORDAN: APPLICATION OF MULTIVARIATE LINEAR REGRESSION AND ADAPTIVE NEURO-FUZZY TECHNIQUES

M. Samhouri  
Industrial Engineering  
Department  
The Hashemite University  
samhouri@hu.edu.jo

A. Al-Ghandoor  
Industrial Engineering  
Department  
The Hashemite University  
ghandoor@hu.edu.jo

R. Fouad  
Industrial Engineering  
Department  
The Hashemite University  
rhfouad@yahoo.co.uk

Abstract

In this study two techniques, for modeling electricity consumption of the Jordanian industrial sector, are presented: (i) multivariate linear regression and (ii) neuro-fuzzy models. Electricity consumption is modeled as function of different variables such as number of establishments, number of employees, electricity tariff, prevailing fuel prices, production outputs, capacity utilizations, and structural effects. It was found that industrial production and capacity utilization are the most important variables that have significant effect on future electrical power demand. The results showed that both the multivariate linear regression and neuro-fuzzy models are generally comparable and can be used adequately to simulate industrial electricity consumption. However, comparison that is based on the square root average squared error of data suggests that the neuro-fuzzy model performs slightly better for future prediction of electricity consumption than the multivariate linear regression model. Such results are in full agreement with similar work, using different methods, for other countries.

Keywords: Energy Consumption, Linear Regression, Neuro-fuzzy, Arab Countries

1. Introduction

Jordan is considered among low-middle income countries, within the Middle East Region, with an average income per capita of about $7,000, in 2007, and its population reached 6 million inhabitants (Department of Statistics (DoS), 2007). It suffers from a chronic lack of adequate supplies of natural resources including fresh water, crude oil and other commercial minerals. Thus, Jordan depends heavily on imports of crude oil, refined products and natural gas from neighboring Arab countries as main sources of energy. Its current imports of around 100,000 barrels of crude oil per day are placing the country under extreme economic pressures, especially with increasing unit price of oil in the international market. The annual energy bill has been rapidly increasing over the past few years due to high rates of population and economic growth combined with the consecutive increase in oil price. Consequently, there has been a growing concern about energy consumption and its adverse impact on the economy and environment, with special focus on the industrial sector, because its contribution accounted for about one third of final energy and electricity consumption.

The industrial sector was probably affected the most by the economic and technological changes that the country has witnessed during the past three decades. For example, in 1985, there were about 4,546 industrial facilities and workshops, and 43,313 employees (DoS, 1986). Two decades later, the industrial sector has grown to include about 13,357 facilities and employed approximately 150 thousand workers, most of them are Jordanians (DoS, 2005). Such enormous increase in the number of facilities and produced products has contributed to an increase on energy and electricity demands. In 2006, electricity consumption reached nearly 8,965 GWh with shares of 31, 34, 15, 15, 3, and 2% for industrial, residential, commercial sectors, water pumping, street light, and for other purposes (such as agricultural sector) respectively (Ministry of Energy and Mineral Resources (MEMR), 2007).

Energy modeling is a subject of widespread current interest among engineers and scientists concerned with problems of energy supply and demand in different sectors of the economy (Al-Ghandoor et al., 2008 a, b, & c; Jaber et. al., 2008, Al-Ghandoor et al., 2005; Jaber et al., 2004; Jaber et al., 2003). It is a tool that has capability of making useful contributions to planning and future policy formulation of energy sector (Dincer and Dost, 1996). In spite of the existence of several studies attempting to analyze current and future energy requirements for different sectors and industries in Jordan (Jaber et al., 2001; Jaber, 2002; Akash and Mohsen, 2003; Tamimi, 1993), there is still a need for more detailed studies and tailor-made models to analyze and explain driving forces behind changes in electricity consumption in the industrial sector and other sectors of the economy. Therefore, this paper will use statistical and artificial intelligence methods to simulate and identify main drivers behind growth in electricity consumption in the
Jordanian industrial sector in order to support the ongoing and future energy-related research aiming to improve the general situation and efficiency of this sector. Such a model would help energy planners to understand the implications of changes in the exogenous variables when underlying relationships are fairly stable; this would be of great help in designing better national energy policies and strategies.

Traditionally, regression analysis has been the most popular modeling technique in predicting energy consumption (Al-Ghandoor et al., 2008 a, b & c; Al-Ghandoor et al., 2005, Flores et al., 2004, Al-Garni et al., 1994). Recently, with the rapid development of data modeling, alternative approaches such as neural networks and neuro-fuzzy methods have become more popular and easier to operate in different areas (Mrinal and Chitralekha, 2006; Onqasakal and Chayakulkheere, 2005). However, their applications, especially the neuro-fuzzy method, are rarely explored in forecasting and/or analysis of energy consumption. This study demonstrates and compares the application of multivariate regression and neuro-fuzzy models in predicting electricity consumption in industry in Jordan.

This paper was organized in four sections, including the introduction. The methodology adopted, in this study, is described in the following section, results and discussions are presented in section three, and the paper ends with some concluding remarks.

2. Methodology

Several models have been previously proposed to model electricity and energy consumptions, ranging from relatively simple ones, such as that proposed by Al-Shehri (2000) which consists of only a single equation model for industrial energy consumption, to more complex models, such as that developed by the Energy Information Administration (EIA, 2000) which consists of several modules that interact with each other through an integration module. A survey of energy consumption models can be found in (Jebaraj and Iniyan, 2006; Worrell et al., 2004). This study will use the analysis of variance (ANOVA) that is based on least square method to identify the most important variables that affect energy consumption of in the industrial sector. The key variables will then be utilized to develop different models that are based on multivariate linear regression and neuro-fuzzy analyses.

2.1 Variables

In the open literature, previous research papers have taken into consideration variables such as energy costs, production levels, number of employees, and number of establishments (Donnelly, 1987; Flores et al., 2004). This study utilizes advantages of such experience and introduces other new important variables such as structural effect and the capacity utilization factor. These variables are briefly explained in the following section, along with reasons for their inclusion in the current proposed model.

2.1.1 Structural effect ($G_p/G_N$)

At a given level of output, as demand changes towards more electricity-intensive industries, electricity demand will rise accordingly. In contrast, a shift towards less electricity-intensive industries would have an opposite effect: causing net reduction on final electricity consumption. To capture this effect in the model, the manufacturing sector is disaggregated into two industrial clusters: electricity intensive (GI) and non-electricity intensive (GN) clusters. The ratio between the gross outputs of these two clusters is included as a variable in this model.

2.1.2 Electricity prices ($E_S$)

When the unit price of electricity is increased, the industry is expected to respond by using electricity more efficiently, i.e. by applying energy conservation and management measures.

2.1.3 Fuel prices ($F_S$)

At present diesel and heavy fuel oil (HFO) are the most common substitutes for electricity in most industries, in Jordan, especially for manufacturing processes such as heating and drying applications. It is worth noting that neither imported nor locally produced natural gas is available, yet, for industry. It is expected that when diesel and HFO prices are increased, most industries will respond by switching to electricity and vice versa. The weighted average of diesel and HFO prices is included in the current model.

2.1.4 Establishments ($E_S$)

Since each facility consumes certain amount of electricity to enable its production, as the number of facilities increases, industrial electricity consumption is expected to increase.

2.1.5 Capacity utilization ($C_U$)

This variable represents how efficiently industrial establishments will utilize the installed production capacity. This variable is included in this study simply because as capacity utilization factor increases, the available resources, including electricity and energy, are employed more
efficiently. Furthermore, as final production increased, then losses per unit produced would be much less.

2.1.6 Number of Employees (EM)

This variable can be viewed from two different perspectives. To some degree, technology has replaced human labor in the majority of industries, and therefore electricity consumption may increase due to the increased on automated production lines and systems. On the other hand, each employee requires additional electricity in the form of conditioned air, hot water, lighting, etc., and thus electricity consumption would decrease as modern machines replace old fashioned manually operated ones.

2.1.7 Gross output (G)

A wide range of products are manufactured, in Jordan, ranging from food products to chemical and mining industries. Therefore, it is extremely difficult to rely on the physical products to indicate the production level. For this reason, the monetary value of goods is used instead of the number of goods in this investigation. Thus, it is expected that the total value of final products will have a proportional relationship with electricity consumption in the industrial sector: higher monetary value of products would require greater rates of electricity consumption in order to be generated.

2.2 Data sources

Historical data, during the period 1985-2004, are utilized to build the current simulation model for electricity consumption in the industrial sector. Electricity consumption data were obtained from the Ministry of Energy and Mineral Resources online database (MEMR, 2007). Economy data such as gross output, number of establishments and employees were taken from annual statistical reports by the Department of Statistics (DoS, 1986-2005). Electricity tariff, which shows the current structure and prices of unit of electricity for different groups of consumers, was taken from published annual reports of different governmental institutions (Jordan Electricity Authority (JEA), 1986-1996; National Electricity Power Company (NEPCO), 1997-2005, Electricity Regulatory Commission (ERC), 2003-2005). Fuel prices were taken from annual reports of MEMR (MEMR, 1986-2005). Other required data about structural effect and capacity utilization variables are not available directly, but have been estimated in a previous research paper (Al-Ghandoor et al., 2008c). As mentioned previously, there are two clusters of industries in Jordan, i.e. (i) the electricity-intensive industries include mining and quarrying, textiles, papers, basic metals, plastics, and chemical industries, and (ii) non-electricity-intensive industries include all other industries in the manufacturing sector. Knowing this, the ratio between the gross outputs of these two clusters (GI/GN) can be determined.

In Jordan, unfortunately, no information is available regarding capacity utilization of different industries; therefore, the following analysis is adopted in this study to estimate the relative capacity utilization from one year to another: for a given industry, the ratio of total gross output to the number of facilities within the industry has been calculated for each year (during the study period from 1985 to 2004). The maximum ratio is taken as an indication of maximum capacity utilization and is given a value of one. The capacity utilization for the remaining years are then calculated relative to the obtained maximum ratio, by dividing the ratio obtained for a given year to the maximum ratio. This analysis has been carried out for all industries, then the overall industrial sector capacity utilization has been calculated as a weighted average of capacity utilization for all industries within the industrial sector. Table 1 shows the complete set of data used in this research.

Table 1. Data set for the Jordanian industrial sector electricity consumption model

<table>
<thead>
<tr>
<th>Year</th>
<th>E (TJ)</th>
<th>G/rGs</th>
<th>ES (JD/TJ)</th>
<th>FS (JD/TJ)</th>
<th>ES</th>
<th>CU (%)</th>
<th>EM</th>
<th>G (10' JD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>3.25</td>
<td>0.42</td>
<td>7.55</td>
<td>1.77</td>
<td>4.54</td>
<td>39</td>
<td>43.3</td>
<td>952</td>
</tr>
<tr>
<td>1986</td>
<td>3.26</td>
<td>0.33</td>
<td>6.87</td>
<td>1.74</td>
<td>4.87</td>
<td>33</td>
<td>40.5</td>
<td>786</td>
</tr>
<tr>
<td>1987</td>
<td>3.82</td>
<td>0.33</td>
<td>6.23</td>
<td>1.59</td>
<td>5.08</td>
<td>29</td>
<td>41.8</td>
<td>870</td>
</tr>
<tr>
<td>1988</td>
<td>3.74</td>
<td>0.40</td>
<td>6.12</td>
<td>1.63</td>
<td>5.19</td>
<td>38</td>
<td>41.6</td>
<td>900</td>
</tr>
<tr>
<td>1989</td>
<td>3.94</td>
<td>0.49</td>
<td>5.88</td>
<td>2.00</td>
<td>5.60</td>
<td>52</td>
<td>48.7</td>
<td>1322</td>
</tr>
<tr>
<td>1990</td>
<td>4.21</td>
<td>0.45</td>
<td>7.80</td>
<td>2.00</td>
<td>6.22</td>
<td>55</td>
<td>51.6</td>
<td>1454</td>
</tr>
<tr>
<td>1991</td>
<td>4.25</td>
<td>0.44</td>
<td>7.82</td>
<td>2.00</td>
<td>8.12</td>
<td>55</td>
<td>51.4</td>
<td>1543</td>
</tr>
<tr>
<td>1992</td>
<td>4.83</td>
<td>0.35</td>
<td>8.69</td>
<td>2.44</td>
<td>9.36</td>
<td>63</td>
<td>70.3</td>
<td>2084</td>
</tr>
<tr>
<td>1993</td>
<td>5.14</td>
<td>0.30</td>
<td>8.70</td>
<td>2.53</td>
<td>9.04</td>
<td>58</td>
<td>71.4</td>
<td>2063</td>
</tr>
<tr>
<td>1994</td>
<td>5.46</td>
<td>0.28</td>
<td>8.66</td>
<td>2.55</td>
<td>8.64</td>
<td>66</td>
<td>99.6</td>
<td>2493</td>
</tr>
<tr>
<td>1995</td>
<td>6.00</td>
<td>0.29</td>
<td>8.62</td>
<td>2.58</td>
<td>8.97</td>
<td>68</td>
<td>103.1</td>
<td>2792</td>
</tr>
<tr>
<td>1996</td>
<td>6.38</td>
<td>0.31</td>
<td>10.11</td>
<td>2.58</td>
<td>8.95</td>
<td>63</td>
<td>106.4</td>
<td>2784</td>
</tr>
<tr>
<td>1997</td>
<td>6.47</td>
<td>0.30</td>
<td>10.09</td>
<td>2.59</td>
<td>8.97</td>
<td>67</td>
<td>108.2</td>
<td>2935</td>
</tr>
<tr>
<td>1998</td>
<td>6.84</td>
<td>0.29</td>
<td>10.07</td>
<td>2.63</td>
<td>9.03</td>
<td>65</td>
<td>110.2</td>
<td>3058</td>
</tr>
<tr>
<td>1999</td>
<td>6.89</td>
<td>0.31</td>
<td>10.09</td>
<td>2.64</td>
<td>10.61</td>
<td>63</td>
<td>116.0</td>
<td>3096</td>
</tr>
<tr>
<td>2000</td>
<td>7.10</td>
<td>0.27</td>
<td>10.09</td>
<td>2.64</td>
<td>11.50</td>
<td>60</td>
<td>123.3</td>
<td>3167</td>
</tr>
<tr>
<td>2001</td>
<td>7.28</td>
<td>0.28</td>
<td>10.11</td>
<td>2.90</td>
<td>12.05</td>
<td>65</td>
<td>130.9</td>
<td>3404</td>
</tr>
<tr>
<td>2002</td>
<td>7.89</td>
<td>0.26</td>
<td>10.37</td>
<td>3.11</td>
<td>12.48</td>
<td>70</td>
<td>136.6</td>
<td>3699</td>
</tr>
<tr>
<td>2003</td>
<td>8.31</td>
<td>0.25</td>
<td>10.64</td>
<td>3.45</td>
<td>12.92</td>
<td>72</td>
<td>143.0</td>
<td>3884</td>
</tr>
<tr>
<td>2004</td>
<td>8.92</td>
<td>0.24</td>
<td>10.68</td>
<td>3.64</td>
<td>13.35</td>
<td>87</td>
<td>149.3</td>
<td>4902</td>
</tr>
</tbody>
</table>

2.2 Multivariate linear regression model

Regression analysis is widely used to analyze multifactor data by building appropriate mathematical equation that relates the response (i.e.
variable of interest) to a set of predictor or independent variables, as shown in Equation 1 which is the proposed multivariate linear regression model. The starting point of this analysis is to define the response variable (here it is electricity consumption in the industrial sector) and the important factors (here these are the structural effects, electricity and fuel prices, number of establishments, capacity utilizations, number of employees, and production outputs) that might be relevant to explain the response’s behavior.

\[
(E)_t = \mu_0 + \mu_{I,N}(\frac{G_I}{G_N}) + \mu_4(EX) + \mu_5(FS) + \mu_6(R) + \epsilon_t
\]

where \( E \) is the electricity consumption, \( \mu_0 \) the regression model intercept, \( \mu_{I,N} \) the regression model coefficient of the gross output ratio between the outputs of intensive (I) and non intensive (N) industrial clusters, \( \mu_j \) represent the regression model coefficients (\( j = 1,2,3,4,5, 6 \)), \( t \) represents the year, and \( \epsilon_t \) the difference between the actual and the predicted electricity consumption.

### 2.4 Adaptive neuro-fuzzy model

Neuro-fuzzy is an associative memory system that consists of fuzzy nodes instead of simple input and output nodes, and it uses neural network learning functions to refine each part of the fuzzy knowledge separately. Learning in a separated network is faster than learning in a whole network (Jang, 1993). An adaptive neuro-fuzzy inference system (ANFIS) is a fuzzy inference system implemented in the framework of an adaptive neural network. By using a hybrid learning procedure, ANFIS can construct an input-output mapping based on both human-knowledge as fuzzy IF-THEN rules and stipulated input-output data pairs for neural networks training. ANFIS architecture is shown in Figure 1, where \( x \) and \( y \) are the inputs, \( f \) is the output, \( A_i \) and \( A_2n \) are the input membership functions, \( w_i \) and \( w_2n \) are the rules firing strengths.

ANFIS is an architecture which is functionally equivalent to a Sugeno-type fuzzy rule base. It is a method for tuning an existing rule base with a learning algorithm based on a collection of training data. This allows the rule base to adapt. Training data is used to teach the neuro-fuzzy system by adapting its parameters (which in essence are fuzzy set membership function parameters) and using a standard neural network algorithm which utilizes a gradient search, such that the mean square output error is minimized. From the ANFIS architecture, shown in Figure 1, it is observed that for given values of premise parameters, the overall output can be expressed as a linear combination of the consequent parameters.

ANFIS modeling and prediction of electricity consumption starts by obtaining a data set (input-output data points) and dividing it into training and validating data sets. The training data set is used to find the initial premise parameters for the membership functions by equally spacing each of the membership functions. A threshold value for error between the actual and desired output is determined. The consequent parameters are computed using the least squares method. Then, an error for each data pairs is found. If this error is larger than the threshold value, the premise parameters are updated using the back propagation neural networks. This process is terminated when the error becomes less than the threshold value. Then, the testing data points are used to compare the model with actual system for validating purposes.

### 2.5 Models comparison criteria

After identifying the most important variables using ANOVA, five models of different data sets were developed using the multivariate linear regression analysis. The same data set is then utilized to develop another five models using the neuro-fuzzy analysis. Each data set is obtained by randomly dividing the whole data into two groups: training group which is used to build the model and testing group which is used to validate the model. The square root of average squared error (RASE) of testing group, given by Equation 2 (Tso and Yau, 2007), is used as performance measure in the comparison process between the multivariate linear regression and neuro-fuzzy models.

\[
RASE = \sqrt{\text{ASE}} = \sqrt{\text{SSE} / n}
\]
\[ SSE = \sum_{i=1}^{n} (E - \hat{E})^2 \]  

(3)

Where \( SSE \) is the sum of squared error, \( n \) the number of observations of testing group, \( E \) the actual electricity consumption, \( \hat{E} \) the predicted electricity consumption, and \( ASE \) the average squared error.

3. Results and discussions

3.1 Variable selection

A software package (Minitab) is applied to test the significance of each variable described before using the least square method (the ANOVA analysis). Table 2 shows the ANOVA results. In order to decide whether or not a variable is significant, the p-value associated with each parameter has been estimated and is provided in the same table. In this study, the variable that has a p-value smaller than 0.05 is considered as an important variable. Therefore, electricity unit price, fuel prices, number of establishments, number of employees, and the structural effect variables were not significant at the 0.05 level.

3.2 Multivariate linear regression analysis

Five models have been built but now using only the significant variables; i.e., the gross output and capacity utilization factor. In each model, fourteen data points have been selected randomly to build the model. Table 3 summarizes obtained results of the five models.

### Table 2. ANOVA results For the Jordanian Industrial electricity Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (TJ)</td>
<td>3731.0</td>
<td>0.000</td>
</tr>
<tr>
<td>( G/G_0 )</td>
<td>471.1</td>
<td>0.605</td>
</tr>
<tr>
<td>( ES/(JD/TJ) )</td>
<td>-0.06649</td>
<td>0.286</td>
</tr>
<tr>
<td>( FS/(JD/TJ) )</td>
<td>0.0119</td>
<td>0.968</td>
</tr>
<tr>
<td>( ES )</td>
<td>0.01396</td>
<td>0.750</td>
</tr>
<tr>
<td>( CU(%) )</td>
<td>-48.108</td>
<td>0.000</td>
</tr>
<tr>
<td>( EM )</td>
<td>-0.02248</td>
<td>0.055</td>
</tr>
<tr>
<td>( G/(1000 JD) )</td>
<td>0.0030649</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: *: \( R^2 \) = 99.6%, Adjusted \( R^2 \) = 99.3%, Predicted \( R^2 \) = 98.55%.

*In this model and all following models, years 1990 and 2004 were deleted from the model since they have large residuals, even after some data transformation and centering.

### Table 3. Regression Summary Outputs for the Jordanian Industrial electricity Models

<table>
<thead>
<tr>
<th>Coeff.</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (TJ)</td>
<td>2952.4</td>
<td>3067.2</td>
<td>3049.2</td>
<td>2938.8</td>
<td>2859.2</td>
</tr>
<tr>
<td>( CU(%) )</td>
<td>-34.062</td>
<td>-34.526</td>
<td>-34.924</td>
<td>-32.483</td>
<td>-32.457</td>
</tr>
<tr>
<td>( G/(1000 JD) )</td>
<td>0.001974</td>
<td>0.001949</td>
<td>0.001956</td>
<td>0.001944</td>
<td>0.001975</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>99.4%</td>
<td>99.5%</td>
<td>99.4%</td>
<td>99.3%</td>
<td>99.4%</td>
</tr>
<tr>
<td>Adjusted ( R^2 )</td>
<td>99.2%</td>
<td>99.4%</td>
<td>99.3%</td>
<td>99.1%</td>
<td>99.3%</td>
</tr>
<tr>
<td>Predicted ( R^2 )</td>
<td>98.92%</td>
<td>99.09%</td>
<td>98.95%</td>
<td>98.78%</td>
<td>98.98%</td>
</tr>
</tbody>
</table>

All coefficients are significant at the 0.05 level. Values in bracket are the VIF.

3.2.1 Model adequacy check

In order to verify the multivariate linear regression model, its adequacy should be checked. First of all, the ANOVA tool, used in the multivariate linear regression analysis to test the validity and significance of the model, is based on some assumptions, such as the residuals having constant variance and being normally distributed. A graphical analysis of the residuals can be used to check the validity of such assumptions (Montgomery and Runger, 2007). Figure 2 (a) and (b) shows the residual plots of one of the previous five models. As can be seen from this figure, the analysis demonstrates satisfactory results since the residuals are contained within a horizontal band (the constant variance assumption is satisfied) and since the cumulative normal distribution is approximately a straight line (i.e. the normality assumption is also satisfied). The residual analysis for the remaining four models demonstrates also satisfactory results. Secondly, no leverage and influence points were detected in all models. Thirdly, larger variation inflation factors (VIFs), usually larger than 10 (Montgomery et al., 2007), indicate that the associated regression coefficients are poorly estimated because of multicollinearity. The later suggests near-linear dependencies among the regression variables. As can be shown from Table 3, the variation inflation factor (VIF) for the parameters is less than 10; this is an indication that multicollinearity does not exist in these models. Fourthly, each model seems to represent its data’s behavior acceptably since most popular measures of goodness-of-fit are satisfied, i.e. the coefficient of multiple determination (R2), the adjusted R2, and the predicted R2 statistics are very high for all models as can be seen from Table 3. Finally, all coefficients in five models have the expected signs, and their magnitudes seem to be reasonable. From the preceding tests, one can conclude that each
suggested model does not violate main assumptions, and represents its data accurately.

The system is well-trained to model the actual electrical energy consumption. The ANFIS-predicted electricity consumption is shown in Figure 5 as a surface plot of electricity consumption as a function of the gross output (G) and capacity utilization (CU).

Different types of membership functions (MF) of the inputs and output were tested to train the ANFIS prediction system. A four Gaussian-type MF for capacity utilization (CU) input and two Gaussian-type MF for gross output (G) input, resulted in high accurate modeling results and minimum training and validation errors. The final (MF) were tuned and updated by the ANFIS model to achieve a good mapping of the input variables to the electricity consumption output. Figure 6 shows the final fuzzy inference system (FIS) used to predict the electrical energy consumption.

### 3.3 Adaptive neuro-fuzzy analysis

The fuzzy logic toolbox of Matlab 7 was used to obtain deemed results. A total of 26 nodes and 8 fuzzy rules were used to build the fuzzy systems for modeling the Jordanian industrial electrical consumption.

#### 3.3.1 ANFIS prediction of electricity energy consumption

The five data sets, mentioned previously, are now used to build another five models that are based on neuro-fuzzy analysis. Following is the results for one of these models. The neural network training for building a fuzzy model for electricity consumption used 14 training data points, and 800 learning epochs. Figure 3 shows the training curve of ANFIS with root mean square error (RMSE) of 23.5 (i.e., almost 0.5%). A comparison between the actual and ANFIS predicted electricity consumption values after training is shown in Figure 4, which shows that
Figure 6. The final fuzzy inference system (FIS) for predicting electricity energy consumption.

3.4 Models comparison

The multivariate linear regression and ANFIS prediction models for Jordanian industrial electricity consumption were validated by feeding into the models the data points that were not used in the model's building. The RASE values of the five models using the multivariate linear regression and neuro-fuzzy techniques are given in Table 4.

In general, as given in Table 4, the neuro-fuzzy technique achieved better results with a lower average value of RASE than the multivariate linear regression technique. Despite this, the difference between prediction accuracy with neuro-fuzzy and regression is not significant and almost have the same performance. This is basically because the electrical energy data does not have highly nonlinear nature, where intelligent methods prove much better than other conventional techniques. Consequently, the performance of both techniques was shown to be almost the same.

Table 4. RASE and average error of multivariate linear regression and neuro-fuzzy models.

<table>
<thead>
<tr>
<th>Model</th>
<th>RASE (TJ)</th>
<th>Neuro-Fuzzy RASE (TJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>132.15</td>
<td>94.75</td>
</tr>
<tr>
<td>Model 2</td>
<td>176.54</td>
<td>126.75</td>
</tr>
<tr>
<td>Model 3</td>
<td>168.19</td>
<td>175.00</td>
</tr>
<tr>
<td>Model 4</td>
<td>121.00</td>
<td>133.00</td>
</tr>
<tr>
<td>Model 5</td>
<td>143.80</td>
<td>69.75</td>
</tr>
<tr>
<td>Average</td>
<td>148.34</td>
<td>119.85</td>
</tr>
</tbody>
</table>

4. Conclusions

Prediction of electricity consumption plays an important role in future energy-related research and planning strategies. A multivariate linear regression method and an adaptive neuro-fuzzy inference system are used to model and predict the electrical energy consumption of the Jordanian industrial sector for the period 1985-2004. The significant factors affecting electricity consumption have been found to be industry's capacity utilization and gross output. The electricity energy consumption values predicted by multivariate linear regression and by ANFIS models are compared to the actual measured values in order to determine the error of both prediction techniques and validate the results. The following conclusions can be drawn from the above analysis:

1) A comparison between the multivariate linear regression and neuro-fuzzy models shows that the neuro-fuzzy model performs slightly better. However, difference is not significant and both models can be considered generally comparable.

2) ANFIS analysis shows that the capacity utilization, CU, input is more important in electricity prediction than the gross output, G. This is because the final MF of the CU are considerably different from the initial ones, while for the gross output, a very little tuning of initial MF parameters is occurred after ANFIS training. This shows the importance of ANFIS model in optimizing the electricity consumption process.

3) The present study shows that ANFIS is a technique that can be used efficiently to model and predict electricity energy consumption. It is believed that this approach can be applied to identify many other parameters in different fields.

References


market. Engineering intelligent systems 2005; 13 (1); 53-59.

