# Application of ANN for Estimation of Power Demand of Villages in Sulaymaniyah Governorate

A. Majeed, P. Ali

*Abstract*—Before designing an electrical system, the estimation of load is necessary for unit sizing and demand-generation balancing. The system could be a stand-alone system for a village or grid connected or integrated renewable energy to grid connection, especially as there are non–electrified villages in developing countries. In the classical model, the energy demand was found by estimating the household appliances multiplied with the amount of their rating and the duration of their operation, but in this paper, information exists for electrified villages could be used to predict the demand, as villages almost have the same life style. This paper describes a method used to predict the average energy consumed in each two months for every consumer living in a village by Artificial Neural Network (ANN). The input data are collected using a regional survey for samples of consumers representing typical types of different living, household appliances and energy consumption by a list of information, and the output data are collected from administration office of Piramagrun for each corresponding consumer. The result of this study shows that the average demand for different consumers from four villages in different months throughout the year is approximately 12 kWh/day, this model estimates the average demand/day for every consumer with a mean absolute percent error of 11.8%, and MathWorks software package MATLAB version 7.6.0 that contains and facilitate Neural Network Toolbox was used.

*Keywords*—Artificial neural network, load estimation, regional survey, rural electrification.

## I. Introduction

FOR simplification of complex energy systems, energy modules are used. It is difficult to analyze such complex systems with a lot of variables virtually; hence, this analysis is carried out through mathematical equations. Thereafter, decision makers could optimize their decisions in terms of system planning and designing [1].

As a major topic, load forecasting has always been popular for energy system planning and operation. The forecasted values of load will affect the commitment of units and security assessment, which in turn will affect the operational costs and system security.

There exist two different types of forecasting systems, the traditional models and the modern techniques. The traditional, such as peak load models and load shape models, use statistical models by employing time series and regression analysis [2]. Recently, tremendous focus and interest have been given to the artificial intelligence technique and its uses for load forecasting application, such as for short-term load forecasting [3], [4], and, different load types are forecasted in [5]; likewise, ANN

Parikhan Ali is with the Sulaimani University, Iraq (e-mail: parikhan.ali@univsul.edu.iq).

in combination with wavelet transform are used in [6], furthermore, many medium and long-term loads are estimated in [7], [8]. However, little research has been conducted in the field of energy habits of rural households. For this purpose, the Genetic Algorithm-Support vector regression (GA-SVR) predicting model is developed to predict village electrical load [9], as well as methods used to determine hourly activity load curves for use in energy models. Conditional demand analysis (CDA) curves have been derived from surveys conducted in, and the logging of, 15 rural villages. Appliance curves are developed using the CDA curves and penetration levels; these could be combined and used to develop the total activity curves. These activity curves can then be used to specify the demand for the relevant energy consuming activities for a rural village [1].

This paper proposes a prediction of average energy consumption each two months for every consumer living in villages using the information collected through a regional survey of electrified villages.

## II. Architecture of ANN

ANN are normally organized into layers of processing units (neurons) with multiple connections; these connections generally have adaptable parameters which modify the signals that pass along them [10]. The formulation of the ANN model requires a number of decisions related to different model parameters such as model structure, training algorithm and input variables. Multilayer feed forward neural networks are most commonly chosen as they have proved their good performance [11].

## III. Neural Network Based Demand Model

There are different factors that affect the design of a neural network which affect the demand model, they are: input variables, structure of the neural network and training and validation algorithms [11].

## IV. Selection of Input Variables

The factors which have an impact on the demand could be selected as the input variables in the structure of the neural network. These factors vary with respect to the characteristics of the consumers and may change from one utility to another.

The fundamental premise in selecting input variables is to use a parsimonious set of demand affecting factors. Only those factors that are globally relevant, i.e., affect demand in a similar manner throughout the entire spectrum of time and weather conditions, should be included [11]. As it could be

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:2, 2018

seen from the literature, the significance of the correlation between a potential input variables and the demand can be determined using some statistical tests based on the linear or nonparametric correlation coefficients. Regression analysis quantifies the correlation between the demand and potential explanatory variables. This paper used the information from a regional survey, as shown in Table I in the appendix, as the inputs, while the availability of electricity differ from one village to another, the average number of hours that electricity is available during the year is selected as another input. And finally, a code consisting of three digits is created to represent each two months throughout a year according to the dates, as shown in Table I, where the output represents the average energy consumption for each two month block.

TABLE I
A CODE GENERATION FOR EVERY TWO MONTHS DURING A YEAR

| Months | Code |
|---|---|
| January & February 2010 | 000 |
| March & April 2010 | 001 |
| May & June 2010 | 010 |
| July & August 2010 | 011 |
| September & October 2010 | 100 |
| November & December 2010 | 101 |

V. DESIGN OF NEURAL NETWORK STRUCTURE

To design a multilayer feed forward network, one needs to select the numbers of neurons in a hidden layer and the hidden layer numbers, the number of neurons in the input and output layers, the connection type between the layers and the activation function for the neurons. In practice, a fully connected network with one hidden layer is a reasonable choice.

The selected input and output neurons depend on the variables affecting the input and output, respectively. But, the number of hidden layer neurons is dependent on the learning capacity of the network. The selection of the number of hidden neurons is the essential issue in the network design. A network with too few hidden neurons will not be capable modeling the demand accurately. Likewise, a lot of neurons in the hidden layer may result in memorizing the training data, and consequently, the performance of the network will be poor when a new data are applied. Also, there is a linear relation between the hidden layer size and the training time [11] and the activation value determines the actual output from the output function unit [12]. In this paper, one hidden layer with different numbers of neurons is tested, in which the minimum error could be achieved by using eight neurons in the hidden layer, as shown in Fig. 1. The activation functions chosen are tan sigmoidal for hidden layer and pure linear in the output layer, respectively. The MathWorks software package MATLAB version 7.6.0, which contains a Neural Network Toolbox that facilitates creation, training, validating and testing of NNs, is used.
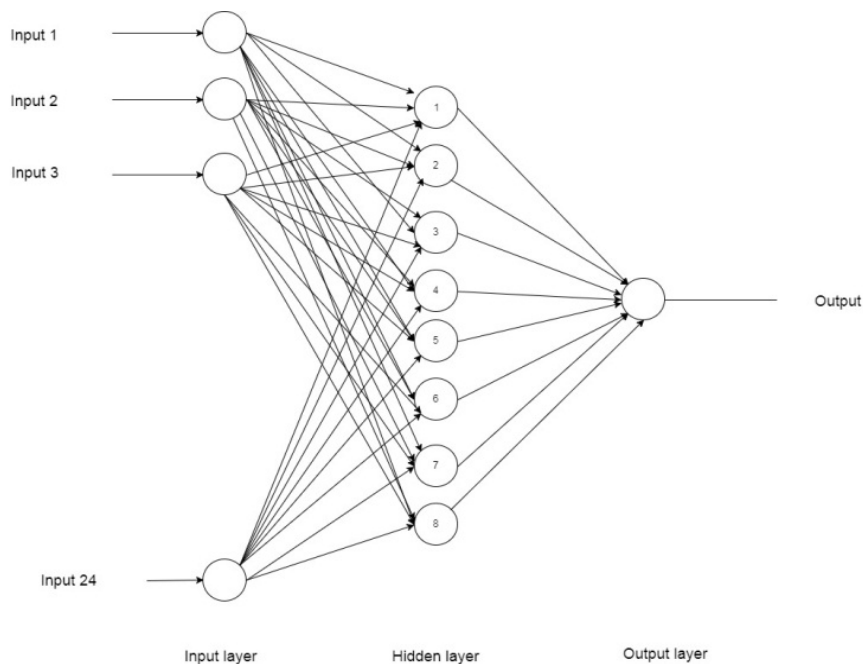


Fig. 1 A multi-Layer Feed forward Network

VI. TRAINING AND VALIDATION

In order to obtain the desired output, training will be processed, which is a determination for the network parameters (weights), which is dubbed as a training set. The multilayer feed forward network could be a supervising train, in which the weights are adapted in such a manner that there is a desired output vector corresponding to an input vector [11]. The performance function for feed-forward networks is mean square error (MSE) referring to (1), which is the average square error between the network outputs and the target outputs.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:2, 2018

$$MSE = \frac{1}{l}\sum_{i=1}^{l}(f(x_i) - y_i)^2 \qquad (1)$$

where $l$ the number of data points is, $f(x_i)$ is the predicted

value and $y_i$ is the actual value.

Back propagation is the usual training algorithm, in which the output is calculated for a given input. Then the network outputs are compared to the targets [13]. Training data for a neural network is used to model the demand of a given customer for a particular time, as in Table I. There are several versions of back propagation algorithms applied for weight adjustment; this paper used the Leven berg-Marquardt algorithm.

Since acceptable training errors do not always guarantee similar network performance for a different set of data; for example, due to the lack of representativeness of the training set or the improperly selected network size, it is necessary to validate the network performance after it is trained [11]. The performance of the model is calculated from the test samples.

## VII. Data Normalization

Data are normalized in order to make the input factors more comparable and to control the effect that is caused by a factor in the network that may be spiked. It helps improve the numerical stability of the training process [11]-[14]. There are many different normalization equations to scale the data; however, the equation used to normalize data in this paper is (2):

$$x_{new} = 0.1 + \left[ 0.8 * \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \right] \qquad (2)$$

where $x_{old}$ is the old value, $x_{new}$ is the new value (normalized

value), $x_{min}$ and $x_{max}$ are the minimum and maximum of the original data range [15].

## VIII. Case Study and Results

A regional survey used to collect information from four different villages; hence, 283 samples of data could be generated, and because the energy selling sector collect the data of energy consumption every two months for each consumer, we had to create a code for each two months as in Table I. Thus, 24 variables are used as inputs for training to predict the average energy consumption corresponding to the months in the input, while the output, which is used as a target, are gained from the administration office of Piramagrun. This energy consumption collection is for a period of one year (2009/12/31 - 2010/12/31), considering the four seasons in which different changes of load due to weather is taken into account through the use of appliances, this is done by adjusting the weights during the learning process. A total of 203 samples are presented to the network during training, and the network is adjusted according to its error, 60 samples are used for validation and 20 samples of the data are used for testing to provide an independent measure of network performance. The mean absolute percentage error (MAPE) is used to evaluate performance of the prediction. Hence, refer to (3) the MAPE for test samples was 15.7%. Fig. 2 shows a curve for actual and predicted average energy consumed for 20 consumers in different months for 283 samples of data.

Then the 283 samples are purified from data that induce large error to 172 samples and trained for better network parameter adaption, while 100 samples are used for training, 52 samples are for validation and 20 samples are for testing, and then the MAPE reduced to 13.1%. Fig. 3 shows the actual and predicted data for average energy consumed of 20 samples in different months for 172 samples of data.

Then, more purification to the 172 samples is conducted to become 156 samples and trained again for better network parameter adaption; the division of these 156 samples is 100 samples for training, 36 samples for validation and 20 samples for testing; the MAPE became 11.8%. Fig. 4 shows the actual and predicted data for the average energy consumed by 20 samples in different months for 156 samples of data.

$$MAPE = \frac{1}{N}\sum_{1}^{N}\left( \frac{|X_i - Y_i|}{X_i} \right) * 100(\%) \qquad (3)$$

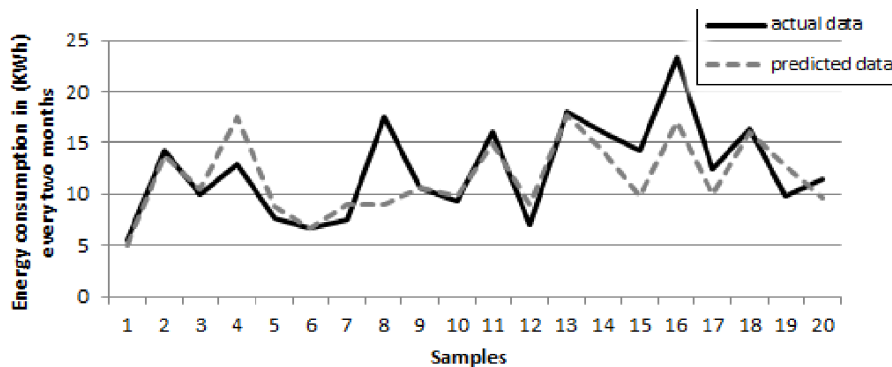where N is number of samples (20 samples). Xi is the actual load and Yi is the predicted load [13].



Fig. 2 Actual and predicted average energy consumed for different 20 samples of 283 samples

World Academy of Science, Engineering and Technology
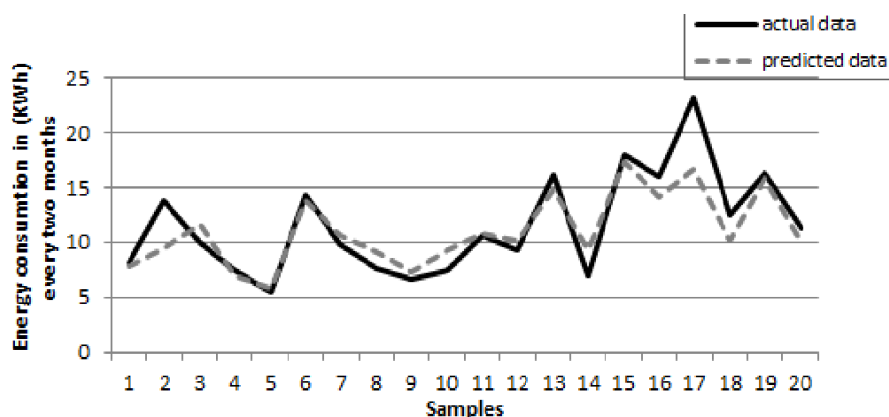International Journal of Computer and Information Engineering
Vol:12, No:2, 2018

Fig. 3 Actual and predicted average energy consumed for different 20 samples of 172 samples
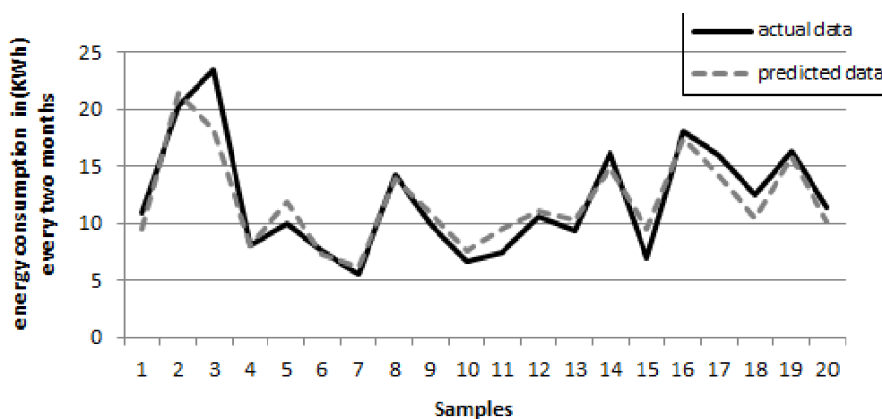


Fig. 4 Actual and predicted average energy consumed for different 20 samples of 156 samples

TABLE II
THE APE% FOR ACTUAL AND PRACTICAL ENERGY THAT HAVE BEEN CALCULATED BY 20 SAMPLES FOR EACH CASE OF 283, 172 AND 156

| Test samples | Total actual energy consumption in (KWh) for 20 samples | Total predicted energy consumption in (KWh) for 20 samples | Absolute percent error (APE%) |
|---|---|---|---|
| The 20 test samples of 283 | 245.9572 | 231.3885 | 5.923287 |
| The 20 test samples of 172 | 230.6293 | 220.8834 | 4.22579 |
| The 20 test samples of 156 | 240.7605 | 237.8069 | 1.226782 |

For each the three cases (283, 172 and 156), test data of 20 samples are used that the network have not seen before, and APE% is used to measure the percentage error for the summed data, as shown in Table II.

When the 283 samples are purified from data that induce large error and trained for better network parameter adaption, the total percentage of error will reduce from 5.9% to 4.2%; moreover, when 172 samples are further purified to become 156 samples, the total error consumption for 20 samples will reduce to 1.22%.

## IX. CONCLUSION

This paper proposed the influence of number of electrical appliances, number of rooms and number of persons living a house, on the average daily energy consumption throughout the different months of a year for every single consumer in a village, using ANN application. The result shows an acceptable percentage of error, while the purpose is to gain an impression of the energy consumption in an electrified village and reflect for the designing and planning for non-electrified similar areas.

## X. RECOMMENDATION

Obtaining detailed and accurate data on energy systems is challenging, particularly for complex systems such as those describing the household energy needs of rural villages, particularly. In the Kurdistan region of Iraq, power is not readily available 24 hours, and the power supply in different regions and villages fluctuates over the time period; for example, one village may have available power for 9 hours a day, while another village may have power for 12 hours. Hence, meeting the need for accurate data is, in fact, one of the greatest obstacles for the modeling process. It is clear that a more detailed representation of reality is needed for modeling purposes, such as the average specific period that an appliance consumes energy, period of occupation in a house and many other factors that affect load changing.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:2, 2018

APPENDIX

TABLE III
A SAMPLE OF DATA FROM REGIONAL SURVEY

| Inputs | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| No. of residents | 13 | 3 | 9 | 4 | 5 |
| No. of rooms | 5 | 4 | 4 | 4 | 3 |
| Split | 1 | 0 | 1 | 0 | 0 |
| Water boiler | 1 | 1 | 1 | 1 | 1 |
| Heater | 1 | 1 | 1 | 1 | 1 |
| Oven | 1 | 0 | 1 | 1 | 0 |
| Iron | 1 | 1 | 1 | 1 | 1 |
| Vacuum cleaner | 1 | 1 | 1 | 1 | 0 |
| Hair dryer | 1 | 1 | 0 | 1 | 0 |
| Water pumping motor | 2 | 0 | 1 | 1 | 1 |
| Cloth washing | 1 | 1 | 1 | 1 | 1 |
| machine | 1 | 1 | 1 | 1 | 1 |
| Refrigerator | 5 | 2 | 3 | 2 | 2 |
| Fan | 1 | 0 | 1 | 0 | 0 |
| Freezer | 2 | 1 | 3 | 2 | 1 |
| TV | 1 | 1 | 0 | 1 | 1 |
| Satellite | 1 | 1 | 1 | 2 | 1 |
| Air conditioner | 30 | 6 | 8 | 4 | 4 |
| Fluorescent lamp | 1 | 1 | 0 | 8 | 6 |
| Incandescent lamp | 1 | 1 | 1 | 1 | 1 |
| Audio/Video | 9 | 9 | 24 | 12 | 12 |
| Hours of electricity available | 0 | 0 | 0 | 0 | 1 |
| Code for different months | 0 | 0 | 1 | 1 | 0 |
| | 0 | 1 | 0 | 1 | 0 |

REFERENCES

[1] N. Cross and C. Gaunt, "Application of rural residential hourly load curves in energy modelling," 2003 IEEE Bologna Power Tech Conference Proceedings.
[2] M. Yi, K. Linn, and M. Kyaw, "Implementation of Neural Network Based Electricity Load Forecasting," *World Academy of Science, Engineering and Technology*, 2008.
[3] P. Mandal, T. Senjyu, and T. Funabashi, "Neural Network Models to Predict Short-Term Electricity Prices and Loads," *2005 IEEE International Conference on Industrial Technology*.
[4] S. Fan, L. Chen, and W.-J. Lee, "Short-Term Load Forecasting Using Comprehensive Combination Based on Multimeteorological Information," *IEEE Transactions on Industry Applications*, vol. 45, no. 4, pp. 1460–1466, 2009.
[5] T. Kakkonda, T. Ishihara, and E. Tsukada, "Electric Load Forecasting by Neural Networks Considering Various Load Types," *IEEE Intelligent System Applications to Power Systems*, ISAP 2003.
[6] A. A. Rasool, A. A. Fttah, and I. Adik, "Short Term Load Forecasting for Erbil Distribution System Using ANN and Wavelet Transform," *International Journal of Computer and Electrical Engineering*, pp. 334–339, 2009.
[7] H. Salama, A. El-Gawad, S. Sakr, E. Mohamed, and H. Mahmoud, "Applications on medium-term forecasting for loads and energy scales by using artificial neural network," *IET Conference Publications*, 2009.
[8] Y. Kareem, "Load Forecasting of the Regional Power System," dissertation, 2007.
[9] J. Yi-feng and W. Shu-wen, "Village Electrical Load Prediction by Genetic Algorithm and SVR." 978-1-4244-5540- 9/10/$26.00 ©2010 IEEE.
[10] C. A. L. Bailer-Jones, R. Gupta, and H. P. Singh, "An Introduction to Artificial Neural Networks," in *Automated data analysis in astronomy*, New Delhi: Narosa Publishing House, 2001, pp. 51–68.
[11] W. Charytoniuk, E. Box, W. Lee, M. Chen, P. Kotas, and P. V. Olinda, "Neural network based demand forecasting in a deregulated environment," *1999 IEEE Industrial and Commercial Power Systems Technical Conference (Cat. No.99CH36371)*.
[12] B. Yegnanarayana, *Artificial neural networks*. New Delhi: Prentice-Hall of India, 2005.
[13] A. El-moudi and S. Al-Alawi, "Using ANN for Forecasting Daily Peak Load of the main Interconnected System of Oman," in *International Conference On Communication, Computer And Power (ICCCP'09) MUSCAT*, 2009 SQU-2009 ISSN: 1813-419X.
[14] R. Zucker, S. Gajjar, V. Rodriguez, and M. Termoul, "Input Techniques for Neural Networks in Stock Market Prediction Ensembles," in *Proceedings of Student-Faculty Research Day, CSIS*, May 7th, 2010.
[15] M. N. Yahya and T. Otsuru, "'Investigation the Capability of Neural Network in Predicting Reverberation Time on Classroom," *International Journal of Sustainable Construction Engineering & Technology*.