Short-Term Natural Gas Consumption Forecast

Dejan Ivezić
Assistant Professor
University of Belgrade
Faculty of Mining and Geology

The results of investigation of an artificial neural network (ANN) model for short term natural gas consumption forecasting are presented. The proposed methodology uses multilayer artificial neural networks to incorporate historical weather and consumption data. Parameters of ANN are obtained from the historical data using a Levenberg-Marquardt training algorithm. Qualities of proposed networks are tested with real data for specific urban consumption area. It was shown that ANN application presented reliable and efficient solution for proposed problem.

Keywords: *artificial neural network, natural gas, consumption.*

1. INTRODUCTION

Accurate forecasting of natural gas consumption for specific distributive area is of great importance for economical and reliable operation of distributive network. According to lack of underground storage in Serbia, forecasting of needs on daily and weekly basis is particularly interesting in the cases of high demands, when the accumulation ability of network itself is decreased. Actual practice is based on experience of operators in companies required for natural gas distribution.

Artificial neural networks (ANN) appeared as a "new" solution for modeling and solution of different classes of problems because of variety of advantages compared to conventional numeric methods. Advantages of ANN are in ability of complex, high non linear functions synthesis, which are very hard and sometimes impossible for explicit mathematic expressions, higher speed of evaluation, robustness and adaptability to changing, etc. Main fields of ANN applications are classifications problems, pattern recognition and functions approximation.

Problem of natural gas consumption forecasting with complex and non linear dependence of different parameters is the typical functions approximation problem. The ability of ANN to learn and construct a complex nonlinear mapping through a set of input/output examples is used in this paper for efficient and accurate prediction of natural gas consumption. Multilayer ANN is designed for modeling dependence of future natural gas consumption to weather conditions, previous consumption, and other parameters of interest. Natural gas consumption forecasting problem is not treated, to the author′s knowledge, in modern references related to neural networks, fuzzy systems, expert systems, etc. However, well known analogy between electricity and natural gas consumption allowed the author to use references related to electric load forecasting problem [1-4]. Consumption covered by selected main measurement and control station is considered. Consumption analysis was done and the training set for ANN was assembled. Quality of proposed ANN is tested with real data.

Received: September 2006, Accepted: November 2006
Correspondence to: Dejan Ivezić
Faculty of Mining and Geology,
Djušina 7, 11000 Belgrade, Serbia
E-mail: ivezic@rgf.bg.ac.yu

2. MULTILAYER ARTIFICIAL NEURAL NETWORKS

Artificial neural networks could be understood as systems for data collecting and processing, composed of large number of mutually related elements – neurons. Neuron (Figure 1.) is the basic unit of the network. This is a computation unit, which produces its output by taking a linear combination of the input signals and by transforming this by a function called activity function. The output of the neuron as a function of the input signals can thus be written:

$$y = f(\sum_{i}^{n} w_i x_i - b) \tag{1}$$

where

y is the output of a neuron;

x_i are the input signals;

w_i are adjustable weights that represent the connection strength;

b is the bias (adjustable, scalar parameter of the neuron);

f is the activity function, which is differentiable and non decreasing, usually represented using a sigmoid function, such as a logistic sigmoid, a tangent sigmoid, etc.

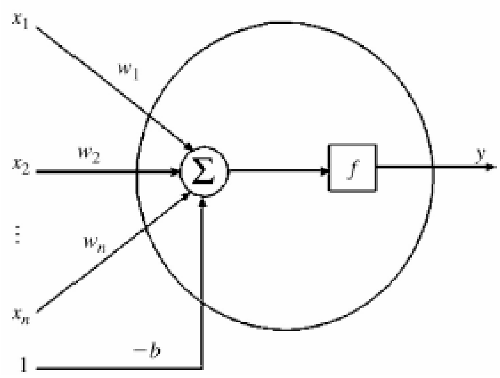

Figure 1. Model of an ANN neuron

Multilayer ANN is characterized by organization of neurons in several layers and a large number of connections between neurons in different layers. The first layer is the input layer, and the last layer is the output layer. Between them are one or more hidden layers. The main question here is: how many hidden layers are necessary for getting a good solution? It was shown [1] that, at least theoretically, a single hidden layer is always used quite safely for the majority of practical problems. Such three-layer ANN is shown in Figure 2. In addition, more than two hidden layers should not be needed in practice.

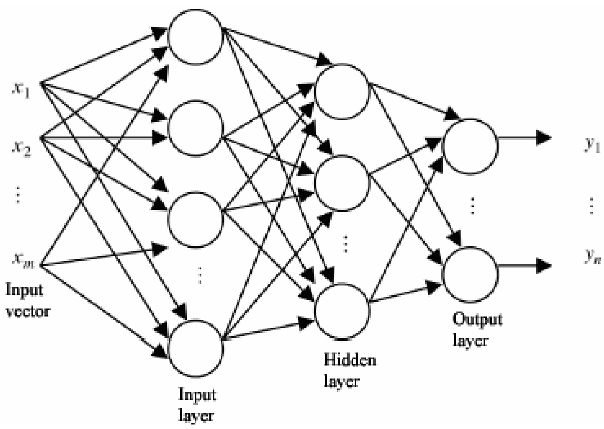

Figure 2. A three-layer ANN

The network weights are adjusted by training the network (Figure 3.). It is said that the network learns through examples. The idea is to give the network input signals and desired outputs. To each input signal the network produces an output signal, and the learning aims at minimizing the sum of squares of the differences between desired and actual outputs. The learning is carried out by repeatedly feeding the input-output patterns to the network. One complete presentation of the entire training set is called an epoch. The learning process is usually performed on an epoch-by-epoch basis until the weights stabilize and the sum of squared errors converges to some minimum value.

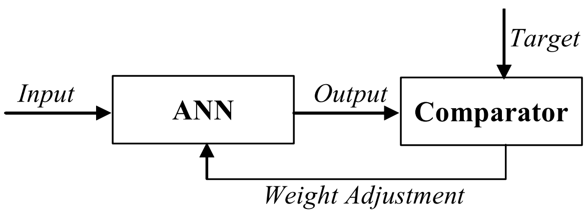

Figure 3. Block diagram of the ANN training

The most often used learning algorithm for the ANN is the backpropagation algorithm. A more powerful algorithm is obtained by using an approximation of Newton's method called Levenberg-Marquardt. In applying the algorithm to the network training, the derivatives of each sum of squared error (i.e. with each training case) to each network weight are approximated and collected in a matrix. This matrix represents the Jacobian of the minimized function. Levenberg-Marquardt algorithm is particularly suitable in solution of approximation problem [2]. In essence, the learning of the network is nothing but estimating the model parameters. In the case of the ANN model, the dependency of the output on the model parameters is, however, very complicated as opposed to the most commonly used mathematical models (for example regression models). This is the reason why the iterative learning is required on the training set in order to find suitable parameter values. There is no way to be sure of finding the global minimum of the sum of squared error. On the other hand, the complicated nonlinear nature of the input-output dependency makes it possible for a single network to adapt to a much larger scale of different relations than, for example, regression models. That is why the term learning is used in connection with neural network models of this kind.

The training aims at minimizing the errors of the network outputs with regard to the input-output patterns of the training set. The success in this does not, however, prove anything about the performance of the network after the training. More important is the success in generalization. A network is said to generalize well, when the output is correct (or close enough) for an input, which has not been included in the training set.

Generalization is influenced by three factors: the size and efficiency of the training set, the model structure (architecture of the network), and the physical complexity of the problem at hand. The latter of these can not be controlled, so the means to prevent overfitting are limited to affecting the first two factors. The larger the training set, the less likely the overfitting is. However, the training set should only include input-output patterns that correctly reflect the real process being modeled. Therefore, all invalid and irrelevant data should be excluded. The effect of the model structure in the generalization can be seen in two ways. First, the selection of the input variables is essential. The input space should be reduced to a reasonable size compared to the size of the training set. If the dimension of the input space is large, then the set of observations can be too sparse for a proper generalization. Therefore, no unnecessary input variables should be included, because the network can learn dependencies on them that do not really exist in the real process. On the other hand, all factors having a clear effect on the output should be included.

One of the problems that occurs during neural network training is called overfitting. The error on the training set are driven to a very small value, but when new data is presented to the network the error is large. The network has memorized the training examples, but it has not learned to generalize to new situations. Early stopping is one of the methods for improving generalization. In this technique the available data is divided into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The validation error will normally decrease during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned. The test set error is not used during the training, but it is used to compare different models. It is also useful to plot the test set error during the training.

Neural network training can be made more efficient if certain preprocessing steps are performed on the network inputs and targets. Before training, it is often useful to scale the inputs and targets to fall within a specified range and to normalize their mean and standard deviation. Also, in some situations, the dimension of the input vector is large, but the components of the vectors are highly correlated (redundant). It is useful in this situation to reduce the dimension of the input vectors. An effective procedure for performing this operation is principal component analysis.

3. CONSUMPTION ANALYSIS AND THE TRAINING SET ASSEMBLING

Increasing of natural gas share in fulfillment of energy demand is strategic target in Serbian energy policy for a long time. Natural gas consumption structure in Serbia and EU is shown in Table 1, [5]. It is evident that industry plays dominant role in Serbian gas consumption. This is because only northern part of the country has complete gas infrastructure and full ability to use natural gas in household sector. Also, negligible share of electricity production with natural gas as a fuel is evident. Comparison with structure of consumption in EU shows that significant growth in the household and electricity production sector could be expected.

Table 1. Structure of natural gas consumption in 2000

Sector	Serbia	EU
Industry, %	65.63	37.60
Heating, %	31.57	41.70
Electricity production, %	2.80	20.70

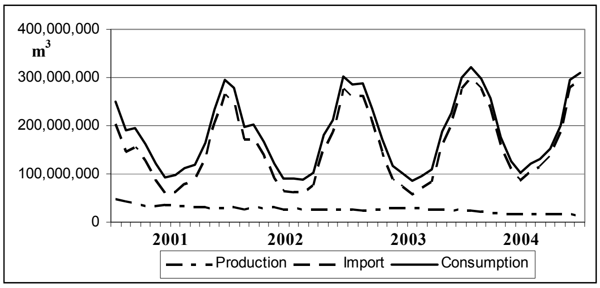

Figure 4. Balance of natural gas in Serbia (2001 – 2004)

Natural gas consumption in the household sector is considered in this paper. (Industry consumption, except for offices and workshops heating is excluded because it is determined by adequate business plans, so that operators of natural gas network have that precise information in advance.) Consumption in the household sector covers natural gas utilization for heating, cooling, cooking and hot water preparation. In Serbia, according to local climate conditions and consumers' practice dominant share in overall consumption is for heating purposes. Illustrative verification of that fact is shown in Figure 4. [5] and it is evident that the summer months' consumption is almost triple less than in winter months. Seasonal inequality, shown in Figure 4, also means that even in industry sector a greater part of natural gas is used for heating purposes (heating, cooking, sanitary water preparation,etc.). As Serbia has no underground storage, management of gas network is a complex task, especially in the period of higher consumption during the winter season. For that reason, consumption forecast is interesting for research in the period from October to April. Presently, experience of operators in companies required for natural gas distribution plays dominant role.

The first idea was to create ANN for prediction of hourly consumption, as it was practice in electricity load forecasting. In natural gas network system such information is not of crucial importance for system operation because network itself has to some degree accumulation ability and can overcome short duration problems. As present operational and system's management practice implies daily order of natural gas, ANN was designed for daily consumption forecast.

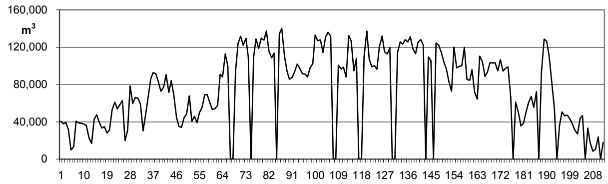

Figure 5. Daily consumption, October 2002 – April 2003

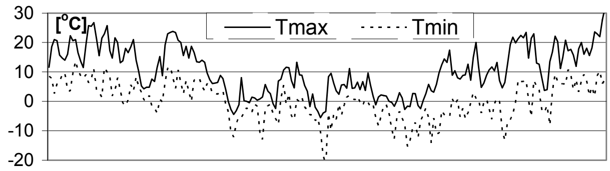

Figure 6. Minimal and maximal temperature, October 2002 – April 2003

Information about daily consumption covered by single gas pressure reduction and metering station (GPRMS) was available for the period from 01/10/2001 till 31/12/2004. Also, correspondent information about daily minimal and maximal temperatures was available. As an example, dates for heating season 2002/2003 are presented in Figures 5-6 (Zero consumption in Fig. 5 means that the information about consumption for specific days are not valid). Consumers covered by this GPRMS are mostly householders, but also, some small industry and public buildings are included in. Authors have no evidence about exact structure of consumption, i.e. exact share of industrial and public consumption in total sum.

Analysis of available data showed:

- Data for October and April are inappropriate for including in training set because of significant variation in consumption without visible reasons in temperature or any other reason. As these months are connected to the begin and the end of the heating seasons, it is reasonable to assume that the tests of greater consumers and irregular operations during the day are reasons for that. Also, absolute consumptions in these two months (Figure 7.) are significantly below the rest of heating season and so are not a valid represent of heating period.

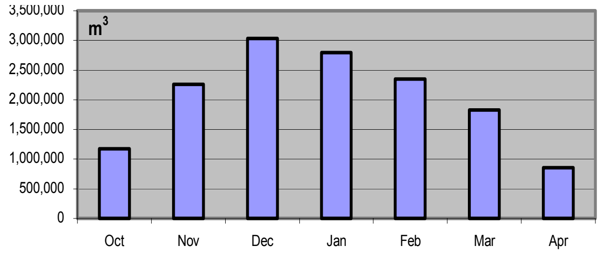

Figure 7. Average consumption in the period October - April for selected consumers

- Weekly consumption variation is significant and particularly visible as a decrease in weekend consumption for the same weather conditions. This is the proof of industry share in heating demands. As a consequence, it is necessary to include information about day in week in the training set.
- Presence of inertia in consumption is evident. Regardless of significant changing in weather conditions, significant decrease or increase of temperature is the clear parameter for that, level of consumption stays the same for a while. Therefore, consumption for previous day is a valuable indication for forecasting and so it is included in the training set.

4. DESIGN, TRAINING AND TEST OF ANN

Previous analysis and the experience in load forecast [1-4] were the basic for ANN selection. Network with one input, one output and one hidden layer is formed, and its training was done by the application of MATLAB Neural Network Toolbox [6]. The number of neurons used for the input layer is determined by the amount and type of the available data:

1 input – neuron for forecasted maximal temperature of the forecasting day
1 input – neuron for forecasted minimal temperature of the forecasting day
1 input – neuron for the consumption in previous day
7 inputs – neurons for the identification of the day in the week. Corresponding day was identified with 1 to appropriate input neuron, while on the other inputs are 0. Only Sunday and Saturday are identified with 0,5 to appropriate input neurons.

Neurons in input layer only distribute input signals to the rest of network. For neurons in hidden layer sigmoid function is used as activity function (2) and linear function is activity function in output layer (3).

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

$$f(x) = k \cdot x \tag{3}$$

Number of neurons in input (10) and output (1) layer is defined by correspondent vectors size. Variations of neurons number in hidden layer showed that 20 is the optimal number of neurons in this layer, i.e. ANN speed and precision performances are optimal with such network structure. Network training is realized by Levenberg-Marquardt learning algorithm and early stopping technique is used for overfitting prevention and improving generalization. Half of available data was used as a training set, one quarter for verification (validation set) and another quarter for the test (test set). Also, preprocessing procedure, including inputs scaling and principal component analysis, was done. Training process is shown in Figure 8. Equality of error changing for all sets show that the selection of training, validation and test set was done properly. Average percentage forecasting errors by months, for the period covered by aggregated training set, are given in Table 2. This measure of the ANN performance is defined as:

$$E = \frac{1}{N}\sum_{i=1}^{N}\frac{|y - y_p|}{y_p} \cdot 100\% \tag{4}$$

where
N is the number of days for consumption forecast;
y_p is the real consumption in selected day;
y is the forecasted consumption in selected day .

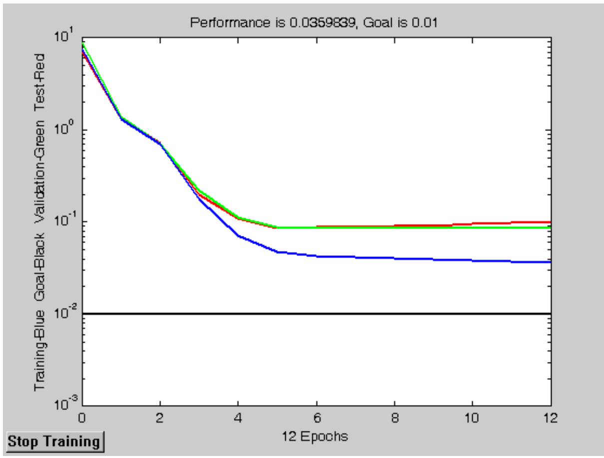

Figure 8. ANN training process

Obtained results on training sets are relatively acceptable. Detailed analysis of results shows that they are better on lower temperature and therefore higher level of consumption. Explanation used for elimination data for October and April from the training set is acceptable as a reason for somewhat higher value of average forecasting error in November and March. For example, average temperature in March 2002 was above 15°C so greater consumers were in intermittent operation, but ANN wasn't trained for such conditions. Also, January is the month with a lot of holidays but information about them is not included in training process.

Table 2. Average percentage forecast errors for training set

	Jan	Feb	Mar	Nov	Dec
2001	-	-	-	5.9	4.06
2002	8.95	9.78	21.26	10.83	5.95
2003	7.22	4.99	9.51	10.47	4.18
2004	5.42	6.26	8.14	11.22	8.51
Total	**7.56**	**7.28**	**14.01**	**10.09**	**5.68**

Designed ANN is tested with real data in period from 02.01.2005 to 07.02.2005 Variation of absolute error in specified period is shown in Figure 9. Average percentage forecasting errors in this period was 5.07% and only in 3 days forecasting errors were over 10%. As this period included few holidays, these days were presented to ANN as weekend days regardless of the day itself. Real and forecasted consumption is presented in Figure 10. and it is visible that forecast values quite well follow the trends of real consumption.

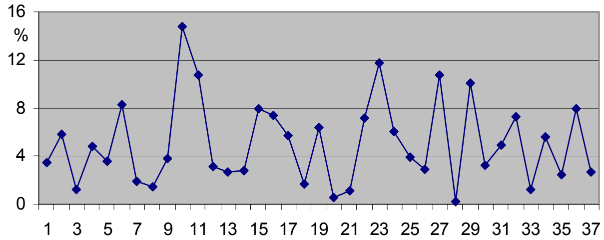

Figure 9. Absolute percentage forecast errors in period 02.01.- 07.02.2005

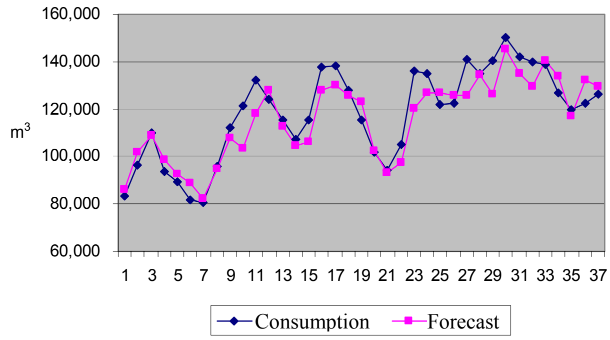

Figure 10. Real and forecasted gas consumption in period 02.01.-07.02.2005

Obtained results are quite satisfactory from the network operators' point of view because natural gas network itself has accumulation ability and can adjust the errors of that range. Hence, proposed ANN could be recognized as a support tool for natural gas network operators. Accumulation ability is also the main difference between electricity and natural gas consumption forecast problem.

If, for some reason, more accurate forecast would be necessary, more careful selection of training data and eventually training of different ANN for separate temperature ranges or separate months could be done.

5. CONCLUSION

Possibility of artificial neural network utilization for forecasting of natural gas consumption is shown in this paper. Specified consumption area is analyzed and appropriate training set, which includes historical weather and consumption data, is defined. Parameters of ANN are obtained using a Levenberg-Marquardt training algorithm. Analyses of results obtained for training and test sets show that designed artificial neural network could be useful for natural gas consumption forecast problem

REFERENCES

[1] Yalcinoz, T. and Eminoglu, U.: Short term and medium term power distribution load forecasting by neural networks, Energy Conversion and Management, Vol. 46, No. 9-10, pp. 1393-1405, 2005.

[2] Beccali, M., Cellura, M., Lo Brano, V., Marvuglia, A.: Forecasting daily urban electric load profiles using artificial neural networks, Energy Conversion and Management, Vol. 45, No. 18-19, pp. 2879–2900, 2004.

[3] Alsayegh, O.A.: Short-term load forecasting using seasonal artificial neural networks. Int. Journal of Power and Energy Systems, Vol. 23, No. 3, pp. 137-142, 2003.

[4] Satish, B., Swarup, K.S., Srinivas, S., Hanumantha Raoa, A.: Effect of temperature on short term load forecasting using an integrated ANN, Electric Power Systems Research, No. 72, pp. 95–101, 2004.

[5] Đajić, N., Mesarović, M.: Cogeneration of heat and power – Chance for increase of natural gas consumption in Serbia, Gas, Vol. 7, No. 4, pp. 9-16, 2004.

[6] Demuth, H., Beale, M.: *Neural Network Toolbox - User's Guide*, The MathWorks, 2005.

КРАТКОРОЧНА ПРОГНОЗА ПОТРОШЊЕ ПРИРОДНОГ ГАСА

Дејан Ивезић

У раду су презентовани резултати истраживања примене вештачких неуронских мрежа на прогнозу потрошње природног гаса. Предложена методологија као улазне податке користи потрошњу гаса и временске услове (минимална и максимална дневна температура). Параметри неуронске мреже су добијени на бази архивских података о потрошњи и температури, коришћењем Levenberg-Marquardt алгоритма обуке. Квалитет предложене неуронске мреже је проверен реалним подацима о потрошњи са одређеног урбаног подручја. Добијени резултати указују на значајне могућности коришћења вештачких неуронских мрежа за решавање постављеног проблема, односно њиховог коришћења при управљању радом гасоводног система.