# Comparing performance of MLP and RBF neural network models for predicting South Africa's energy consumption

**Olanrewaju A Oludolapo**
*Department of Industrial Engineering, Tshwane University of Technology, South Africa*

**Adisa A Jimoh**
*Department of Electrical Engineering, Tshwane University of Technology, South Africa*

**Pule A Kholopane**
*Department of Industrial Engineering, University of Johannesburg, South Africa*

***Abstract***
*In view of the close association between energy and economic growth, South Africa's aspirations for higher growth, more energy is required; formulating a long-term economic development plan and implementing an energy strategy for a country /industry necessitates establishing the correct relationship between energy and the economy. As insufficient energy or a lack thereof is reported to be a major cause of social and economic poverty, it is very important to select a model to forecast the consumption of energy reasonably accurately. This study presents techniques based on the development of multilayer perceptron (MLP) and radial basis function (RBF) of artificial neural network (ANN) models, for calculating the energy consumption of South Africa's industrial sector between 1993 and 2000. The approach examines the energy consumption in relation to the gross domestic product. The results indicate a strong agreement between model predictions and observed values, since the mean absolute percentage error is below 5%. When performance indices are compared, the RBF-based model is a more accurate predictor than the MLP model.*

*Keywords: multilayer perceptron, radial basis function, energy consumption, gross domestic product*

## 1. Introduction

In 1995, South Africa found itself among the top 50 countries (developed and developing countries) in the world, and first in Africa among countries south of the equator in terms of the per capita commercial energy consumption in 1995 (2 405 kg of oil equivalent per capita in 1995). However, where energy efficiency is measured as the ratio of real GDP to energy use (1 US$ GDP at 1987 values per kg of oil equivalent), South Africa's ranking was very low (among the last 50 out of a possible 150 countries) (World Bank, 1998). Long-term development can only take place when there is access to affordable energy. Insufficient energy or a lack thereof has been reported to be a cause of social and economic poverty, so it is very important to select a model to forecast consumption of energy reasonably accurately.

Accurate forecasting is important to both government and industry who need to provide viable estimates on future revenue, cost, demands (Wedding and Cios, 1996) and energy consumption. A lack or a shortage of energy is perceived to have a detrimental effect on the economy and gross domestic growth (Anonymous, 2003). Energy is the basis for sustainable development and the means to ensure a healthy economy (Wang, 2009). Implementing a long-term economic development plan and an energy strategy for a country/industry requires establishing the optimal relationship between energy supply and the economy.

It is worth noting that economic events and regime changes in the economic environment, in energy policy and fluctuations in energy prices can lead to structural changes in the pattern of energy consumption in a period under study(Chiou-Wei *et al.,* 2008), consequently, the relationship between energy consumption and economic growth (Chiou-Wei *et al.,* 2008; Lee and Chang 2005) should be regarded as non-linear.

A sound forecasting technique is crucial to develop an accurate plan and to formulate an energy strategy. To date, the most popular modelling technique used to predict energy consumption has been regression analysis. To predict energy consumption based on adequate data analysis, this study has used data from the gross domestic product (GDP) of South Africa's industrial sector, as well as total energy consumption in neural network analysis, to achieve more reliable results.

Using an artificial neural network method rather than a traditional classification method derives from the success in estimating the non-linear function (Mabel and Fernandez 2008). However, apart from estimating the non-linear function in a shorter period of time, the advantage of the artificial neural network (ANN) approach is that energy applications are more viable, making them more attractive to potential users such as energy engineers *(Kaukal et al.,* 2011).

ANNs are computer programs that are designed to recognize both linear and non-linear relationships between the input and the output variables in a given data set (Al-Alawi *et al.*, 2003). ANNs are able to process information and provide models even when the information and data are complex, noise-contaminated, non-linear or incomplete. The goal of an ANN is to map a set of input patterns against a corresponding set of output patterns. The network accomplishes this mapping by learning from a series of examples and defining the input and output sets for a given system (Amir Heydari *et al.*, 2006). The network then applies what it has learnt to a new input pattern to predict the appropriate output (Amir Heydari *et al.,* 2006; Zuptan and Gasteiger 1999).

Many studies on energy demand and consumption forecasting exist in the literature. Among these studies, linear and non-linear statistical models, including ANN programs, have been used by Pao (2006) to determine the influence of four economic factors on the electricity consumption in Taiwan and to develop an economic forecasting model. An ANN model that has four independent variables, namely GDP, population, and import and export costs has been used by Geem and Roper (2009) to estimate the energy demand in South Korea accurately, and Bianco *et al.*, (2009) took into account the influence of several economic and demographic variables related to the annual electricity consumption in Italy, to develop a long-term consumption forecasting model.

The industrial sector is at the core of developing projects because it is the most important end-user in developing countries, to ensure economic growth (Lee and Chang, 2007).This study aims to determine the empirical factors affecting estimation of energy consumption in the industrial sector of South Africa by using the multilayer perceptron (MLP) and radial basis function (RBF) of artificial neural network (ANN) models, and comparing the prediction capabilities of the models.

It was found from the comparison of performance indices, based on the statistical measures namely, mean absolute percentage error (MAPE), coefficient of correlations and visual inspection, that prediction performance of the RBF model was superior to that of the MLP function.

## 2. Data

In this paper, the industrial sector of South Africa, a developing country was assessed. Annual data from 1993 to 2000 reflected in Table 1 on total energy consumption and real GDP were used. These were obtained from the Integrated Energy Plan for the Republic of South Africa, and the Department of Minerals and Energy (Anonymous 2003). Total energy consumption is measured in Peta Joules (PJ) with renewable and waste excluded and for GDP, 1995 was used as the base year.

## 3. MLP neural network and RBF neural network

### *3.1. MLP structure and design*

Since their inception in the 1940s, different neural network models have been developed, but the MLP is still the most widely used (Mata, 2011). This network consists of three layers namely, input layer, hidden layer and output layer, with each layer having one or more neurons. In addition, bias neurons are connected to the hidden and output layers as shown in Figure 1.

The computational procedure of the network is described below(Hsu and Chen, 2003):

$$Y_j = f(\Sigma_i w_{ij} X_{ij}) \qquad (1)$$

where $Y_j$ is the output of node j, $f(.)$the transfer function, $w_{ij}$ the connection weight between node j and node i in the lower layer and $X_i$ the input signal from the node *i* in the lower layer. The backpropagation is based on the steepest descent technique with a momentum weight (bias function) which calculates the weight change for a given neuron. It is expressed as follows (Hsu and Chen, 2003; Huang *et al.,* 2002): let $\Delta w^p_{ij}(n)$ denote the synaptic weight connecting the output of neuron *i* to the input of neuron *j* in the *p*th layer at iteration *n*.

The adjustment $\Delta w^p_{ij}(n)$ to $w^p_{ij}(n)$ is given by

$$\Delta w^p_{ij}(n) = \eta(n)\frac{\delta E(n)}{\delta^p_{ij}} \qquad (2)$$

where $\eta(n)$ is the learning rate parameter. By using the chain rule of differentiation, the weight of the network with the backpropagation learning rule is updated using the following formulae:

$$\Delta w^p_{ij}(n) = \eta(n)\delta^p(n)\, X^{p-1}{}_i(n)m(n)\Delta w^p_{ij}(n-1) \qquad (3)$$

$$\Delta w^p_{ij}(n+1) = w^p_{ij}(n) + \Delta w^p_{ij}(n) \qquad (4)$$

**Table 1: Industrial sector data**

| | *1993* | *1994* | *1995* | *1996* | *1997* | *1998* | *1999* | *2000* |
|---|---|---|---|---|---|---|---|---|
| GDP - All industries | 472 | 486 | 500 | 521 | 534 | 538 | 549 | 571 |
| Total final energy consumption (PJ) | 1766 | 1789 | 2016 | 1996 | 2071 | 2098 | 2026 | 2003 |

*All industries are listed at basic prices Rand (R) - billion (constant 1995 prices), Renewable and Waste are excluded.*

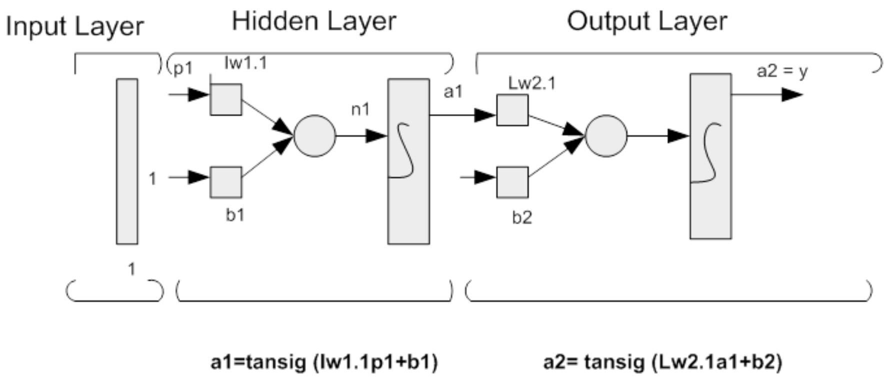


**Figure 1: MLP Single hidden layer neural network structure**

where $\delta^p{}_i(n)$ is the nth error signal at the jth neuron in the pth layer, $X^{p-1}{}_i(n)$ is the output signal of neuron *i* in the hidden layer, and m is the momentum factor.

Newff is a Matlab code which creates a feed-forward backpropagation network. This was used to calculate a precise function of the MLP neural network. The number of hidden neurons was determined by comparing the performance of different cross-validated networks, with 1–15 hidden neurons, and choosing the number that produced the greatest network performance. This resulted in a network with single input neuron (GDP), five hidden neurons and a single output neuron (energy consumption). In the analyses, network parameters of learning rate and momentum were set at 0.05 and 0.7, respectively. A variable learning rate with momentum (trainlm) as the network's training function, and tansig as activation functions for all layers was used. The data used by the network must be scaled for the network to be effectual. In theory, the inputs to the network can be any value. However, scaling values to the same order of magnitude (generally in the range 0 to 1 or -1 to 1) enables the network to learn relationships more quickly (Hart, 1992). In this paper, the data was scaled to the range -1 to 1 to ensure a consistent scaling regime for input and output. The Matlab code for the design is as follows:

```
p=[472 486 500 521 534 538 549 571];
t=[1766 1789 2016 1996 2071 2098 2026
2003];
[pn,minp,maxp,tn,mint,maxt] = premnmx(p,t);
iitst = 2:4:Q;
iival = 4:4:Q;
iitr = [1:4:Q 3:4:Q];
val.P = pn(:,iival); val.T = tn(:,iival);
test.P = pn(:,iitst); test.T = tn(:,iitst);
ptr =pn(:,iitr); ttr = tn(:,iitr);
net = newff(minmax(ptr),[5 1],{'tansig' 'tan-
sig'},'trainlm');
net.trainParam.show=50;
net.trainParam.lr=0.05;
net.trainParam.mc=0.7;
[net,tr]=train(net,ptr,ttr,[],[],val,test);
an=sim(net,pn);
a = postmnmx(an,mint,maxt);
error=(t-a)
```

### *3.2. RBF structure and design*

Due to their better approximation capabilities, simpler network structures and faster learning algorithms, RBF networks have been widely used in many science and engineering applications (Benghanem and Mellit, 2010; Mellit and Kalogirou, 2008).

The RBF neural network is a kind of feed-forward neural network (Tong *et al.*, 2009). The RBF neural network (Bishop, 1991; Wedding and Cios, 1996) comprises three layers: input layer, hidden layer and output layer. Between the input and output layers is a layer of processing units known as hidden units. Each of these implements a radial basis function (Tong *et al.*, 2009). The distance between hidden-layer neurons is connected with the input of the weight and the vector $a_p$, multiplied by the threshold as their own input, as shown in Figure 2.

A hidden layer of *i* input:

$$t_i^q = \sqrt{\sum_j (w1_{ji} - a_j^p)^2} \text{X d1}_i \quad (5)$$

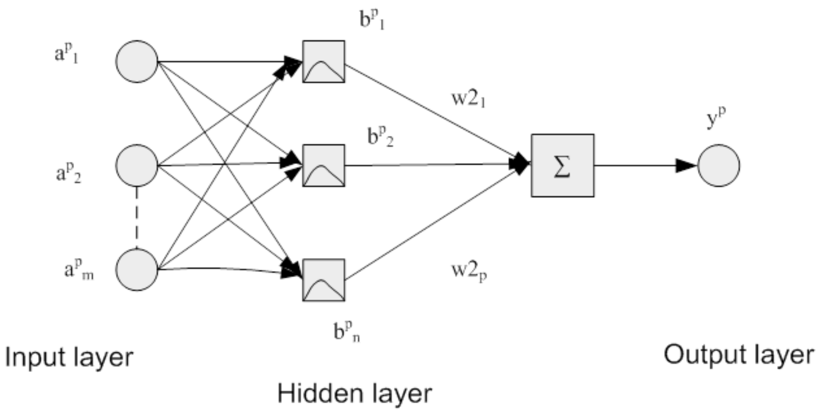


**Figure 2: RBF Neural network structure**

The output:

$$b_i^p = \exp\left[-t_i^{q^2}\right] = \exp\left(\sqrt{\textstyle\sum_j (w1_{ji} - a_j^p)^2}\ \mathrm{X}\ d1_i\right)$$
$$= \exp\{\left(\left\|w1_i - a_p\right\|\ \mathrm{X}\ d1_i\right)^2\} \tag{6}$$

Although RBF threshold d1 can adjust the sensitivity function, in this work, another parameter C (expansion constant) was used. To determine the neural network, the relationship between C and d1 in the Matlab Toolbox is:

$$d1_i = 0.8326/C_i \tag{7}$$

The output of the hidden-layer neurons at this point can be represented by the following equation:

$$S_j^p = \exp\left\{\frac{0.8326\ X\sqrt{\sum_j (w1_{ji} - a_j^p)^2}}{c_i}\right\} \tag{8}$$
$$= exp(0.8326^2 \mathrm{X} \left\{\frac{\{\|w1_i - a_p\|\}^2}{c_i}\right\})$$

The weighted sum of the hidden layer neurons output serves as input data for the output:

$$y^p = S^n_{i=1} b_i\ \mathrm{X}\ w2_i \tag{9}$$

Training the RBF network entails two steps: the first step is to learn without been taught, determining weight w1 between input layer and hidden layer; the second step is to identify weight w2 between the hidden layer and output layer (Caiqing et al., 2008).

Newrbe is a Matlab code which designs a radial basis network with zero error in the design vectors. The code was used to create a precise function for the RBF neural network, which automatically chooses the number of the hidden layer, making predictions more accurate (Caiqing et al., 2008). In the analysis, the network parameter SPREAD which is the distribution density of RBF was set to 2.3 for the network. The Matlab code for the design is as follows:

```
p=[472 486 500 521 534 538 549 571];
t=[1766 1789 2016 1996 2071 2098 2026
2003];
p=mapminmax(p,0,1);
[t,ts]=mapminmax(t,0,1);
spread=2.3;
net=newrbe(p,t,spread);
yn=sim(net,p)
```

## 4. Prediction performance comparisons

Two different statistical measures were employed to evaluate the energy consumption prediction capability of each of the ANN models.

### *4.1. Coefficient of correlation R2 and mean absolute percentage error (MAPE)*

Table 2 reflects the $R^2$ and MAPE for the energy consumption for each model. The total sum of squared deviations in Y (energy consumption) can be decomposed into two qualities, the first, SSR, measures the quality of x (GDP) as a predictor of Y, and the second, SSE, measures the error in the prediction. Thus, the square of the correlation coefficient between the response variable Y and the predictor x is:

$$R^2 = 1 - \frac{SSE}{SST} = \frac{SSR}{SST} \tag{10}$$

where SST = SSR + SSE (11)
*SST* = Total sum of squares, *SSR* = Sum of squares due to regression and *SSE* = Sum of squares due to error.

It should be noted that $0 \le R^2 \le 1$ because SSE $\le$ SST. If $R^2$ is near 1, then x (GDP) accounts for a large part of the variation in Y (energy consumption) (Chatterjee and Hadi, 2006).

The mean absolute percentage error (MAPE) which is a measure of accuracy in a fitted series value in statistics, is expressed in per cent

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Ai - Pi}{Ai}\right| \mathrm{x}\ 100 \tag{12}$$

where $A_i$ is the actual value, $P_i$ is the predicted value and N is the number of data. A MAPE below 5% is the measure of a highly accurate prediction.

## 5. Results and discussion

Computer codes for MLP and RBF-models were developed in Matlab Software (version 2010a). The models were trained until the best performance was obtained. The optimal parameters (weights and bias) of the networks were saved and used for testing and validating operation of the models.

In order to test and validate the different models, two statistical tests (the correlation coefficient and the mean absolute percentage error 'MAPE') between the measured and the estimated annual energy consumption data using the MLP and RBF network were carried out. The results obtained are summarised in Table 2. From the simulations carried out, it was found that better performance was delivered by the RBF-model according to the correlation coefficient between both sets of data (measured and estimated).The obtained is 0.9998, is higher than the corresponding one in the MLP-model, while the MAPE is lower than that of the MLP-model. In order to demonstrate the efficiency of the proposed RBF-model, a comparison was done between the developed RBF, and MLP models in Figures 3 to 6. Figure 7 illustrates the deviation from actual data of the two models. The RBF model easily learnt to capture the industrial sectors' energy consumption (with the least forecasting errors).

**Table 2: Performance indices for models**

| Model | MAPE | $R^2$ |
|---|---|---|
| MLP | 3.3 X $10^{-2}$ | 0.9959 |
| RBF | 2.07 X $10^{-2}$ | 0.9998 |

*MAPE = mean absolute percentage error, $R^2$ = correlation coefficient*

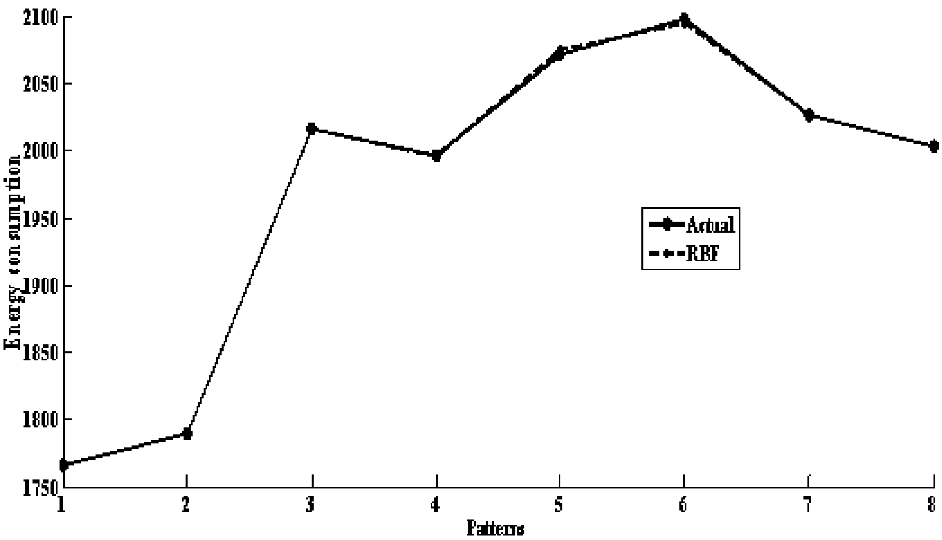


**Figure 3: Comparison between the actual and RBF predictions**

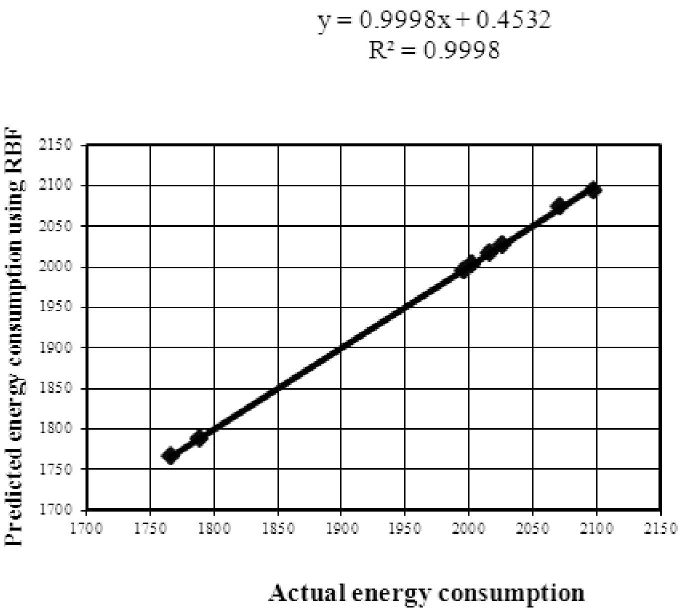


**Figure 4: Regression analysis between the actual and RBF predictions**

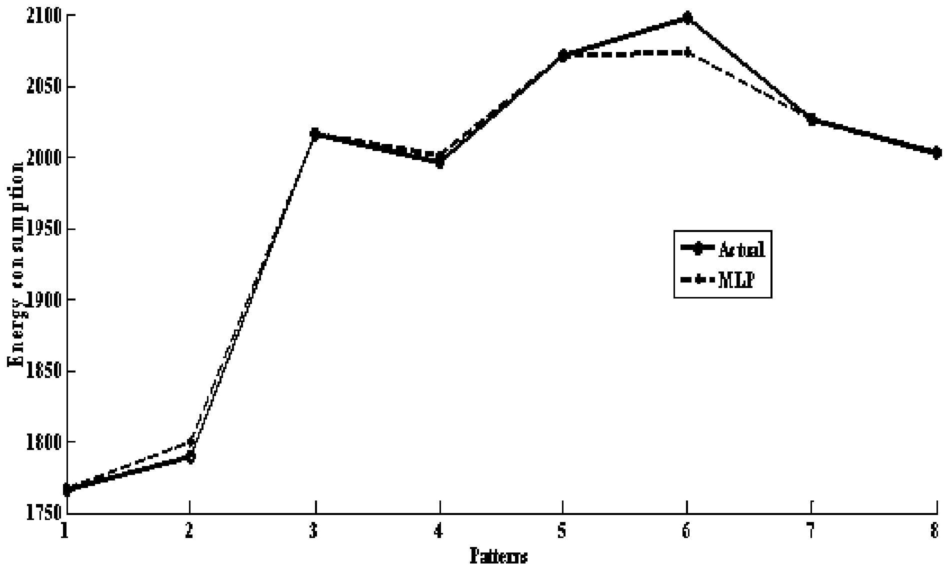


**Figure 5: Comparison between the actual and MLP predictions**

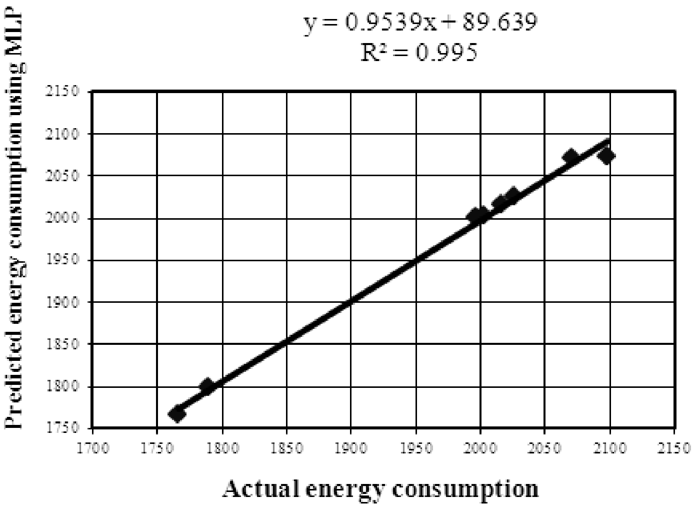


**Figure 6: Regression analysis between the actual and MLP predictions**

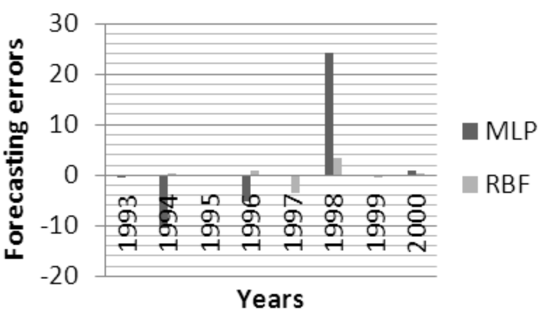


**Figure 7: Deviation from actual data of the two models**

### 6. Conclusion

Forecasting energy consumption is essential for the long-term development of South Africa, especially for the industrial sector which plays a pivoted role in the country's economic growth. In order to establish accurately the relationship between energy consumption and the economy of an industrial sector for estimating the energy consumption, ANN models were employed. In this paper, MLP and RBF models for estimating South Africa's industrial annual energy consumption were adapted. The measured annual energy consumption was compared with that estimated using MLP and RBF models. The predictive performance of each model was assessed using two statistical measures: $R^2$, MAPE and a study of the graphs were used. The results of the statistical measures suggest that RBF-model provides more accurate results than the MLP-model.

It has been demonstrated that the RBF-model gives more accurate results when compared with those obtained using the MLP network model. Both models deliver highly accurate predictions, since the MAPE values are below 5%. The developed model is suitable for South Africa's industrial sector. It is concluded that the predicted data generated by the RBF network is evidently suitable for estimating the energy consumption of South Africa's industrial sector to formulate an accurate development plan and to implement a viable energy strategy.

## References


Al-Alawi, S., Al-Badi, A., and Ellithy, K. (2003). An artificial neural network model for predicting gas pipeline induced voltage caused by power lines under fault conditions. *An artificial neural network model*, 69.

Amir Heydari, Shayesteh, K., and Kamalzadeh, L. (2006). Prediction of Hydrate formation temperature for natural gas using artificial neural network. *Oil and Gas Business*.

Anonymous. (2003). Integrated Energy Plan for the Republic of South Africa. South Africa.

Benghanem, M., and Mellit, A. (2010). Radial Basis Function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia. *Energy*, 35, 3751-3762.

Bianco, V., Manca, O., and Nardni, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, 34, 1413-1421.

Bishop, C. M. (1991). Improving the generalization properties of radial basis function neural networks. *Neural computation*, 3, 579-588.

Caiqing, Z., Ruonan, Q., and Zhiwen, Q. (2008). Comparing BP and RBF Neural Network for Forecasting the Resident Consumer Level by MATLAB. International Conference on Computer and Electrical Engineering.

Chatterjee, S., and Hadi, A. S., eds. (2006). *Regression Analysis by Example*, A John Wiley and Sons inc.

Chiou-Wei, S. Z., Chen, C.-F., and Zhu, Z. (2008). Economic growth and energy consumption revisited - Evidence from linear and non-linear Granger causality. *Energy Economics*, 30, 3063-3076.

Geem, Z. W., and Roper, W. E. (2009). Energy demand estimation of South Korea using artificial neural network. *Energy Policy*, 37, 4049-4054.

Hart, A. (1992). Using neural networks for classification tasks - some experiments on datasets and practical advice. *J.Opl. Res. Soc.*, 43, 215 - 226.

Hsu, C.-C., and Chen, C.-Y. (2003). Regional load forecasting in Taiwan-applications of artificial neural networks. *Energy conversion and Management*, 44, 1941-1949.

Huang, H., Hwang, R., and Hsieh, J. (2002). A new artificial intelligent peak power load forecaster based on non-fixed neural networks. *Electric Power Energy Syst*, 24, 245-250.

Kaukal, M., Akpinar, A., Komurcu, M. I., and Ozsahin, T. S. (2011). Modelling and Forecasting of Turkey's energy consumption using socio-economic and demographic variables. *Applied Energy*, 88, 1927-1939.

Lee, C.-C., and Chang, C.-P. (2007). The impact of energy consumption on economic growth: Evidence from linear and nonlinear models in Taiwan. *Energy*, 32, 2282-2294.

Lee, C. C., and Chang, C. P. (2005). Structural breaks, energy consumption, and economic growth revisited: evidence from Taiwan. *Energy Economics*, 27, 857-872.

Mabel, M. C., and Fernandez, E. (2008). Analysis of wind power generation and prediction using ANN: A case study. *Renewable Energy*, 33, 986-992.

Mata, J. (2011). Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, 33, 903-910.

Mellit, A., and Kalogirou, S. (2008). Artificial intelligence techniques for photovoltaic applications: a review. *Prog Energy Comb Sci*, 34, 574-632.

Pao, H. T. (2006). Comparing linear and non-linear forecasts for Taiwan's electricity consumption. *Energy* (31), 2129-2141.

Tong, X., Wang, Z., and Yu, H. (2009). A research using hybrid RBF/Elman neural networks for intrusion detection system secure model. *Computer Physics Communication*, 180, 1795 - 1801.

Wang, L. (2009). Grey Incidence Analysis between Energy Consumption Structure and Chinese GDP Growth. *Proceedings of 2009 IEEE International Conference on Grey Systems and Intelligent Services, November 10-12, 2009*, Nanjing, China.

Wedding, D. K., and Cios, K. J. (1996). Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. *Neurocomputing*, 10, 149 - 168.

World Bank. (1998). World Development Indicators. The International Bank for Reconstruction and Development/the World Bank Washington DC.

Zuptan, J., and Gasteiger, J., eds. (1999). *Neural Networks in Chemistry and Drug Design*, Wiley/VCH, New York.


*Received 15 June 2011; revised 7 May 2012*