

Comparative Application of Radial Basis Function and Multilayer Perceptron Neural Networks to Predict Traffic Noise Pollution in Tehran Roads

Ali Mansourkhaki¹, Mohammadjavad Berangi¹, Majid Haghiri^{1*}

¹ School of Civil Engineering, Iran University of Science & Technology, Narmak, Tehran, Iran

* Corresponding author's e-mail: majid.haghiri@gmail.com

ABSTRACT

Noise pollution is a level of environmental noise which is considered as a disturbing and annoying phenomenon for human and wildlife. It is one of the environmental problems which has not been considered as harmful as the air and water pollution. Compared with other pollutants, the attempts to control noise pollution have largely been unsuccessful due to the inadequate knowledge of its effects on humans, as well as the lack of clear standards in previous years. However, with an increase of traveling vehicles, the adverse impact of increasing noise pollution on human health is progressively emerging. Hence, investigators all around the world are seeking to find new approaches for predicting, estimating and controlling this problem and various models have been proposed. Recently, developing learning algorithms such as neural network has led to novel solutions for this challenge. These algorithms provide intelligent performance based on the situations and input data, enabling to obtain the best result for predicting noise level. In this study, two types of neural networks – multilayer perceptron and radial basis function – were developed for predicting equivalent continuous sound level (LA_{eq}) by measuring the traffic volume, average speed and percentage of heavy vehicles in some roads in west and northwest of Tehran. Then, their prediction results were compared based on the coefficient of determination (R^2) and the Mean Squared Error (MSE). Although both networks are of high accuracy in prediction of noise level, multilayer perceptron neural network based on selected criteria had a better performance.

Keywords MLP; RBF; neural network; noise prediction; road traffic noise

INTRODUCTION

Noise pollution is one of the environmental phenomena which can threaten the human physical and mental health. World Health Organization (WHO) put the noise pollution in the third place of dangerous pollution for human health, right after air and water pollution (WHO 2005). Among the noise producers in cities such as construction, commercial and industrial activities, road traffic noise is one of the most important factors that plays an important role in increasing abnormal noise in cities. Heterogeneity of the urban environment with environmental noise characteristics and interfering factors such as time and spatial diversity make the simulation and prediction of noise a complex and nonlin-

ear problem. In this situation, the application of learning algorithms such as artificial neural networks that mimic some functions of the brain including as learning and modeling to analyze the data, could be very useful. Extensive research proved the predicting power of these algorithms in traffic noise pollution.

Literature Review

Cammaratta et al. (1995) measured the traffic parameters in Sicily, Italy, provided a neural network consisting of two steps. In the first step, the Learning Vector Quantization (LVQ) filtered the data and in the second step, Back Propagation Network (BPN) predicted the sound pressure level. Caponetto et al. (1997) used a genetic al-

gorithm to optimize a fuzzy logic model for predicting the environmental noise. The results indicated that his method was successful. Parabat and Nagamail (2007) provided a comprehensive study to assess the neural network model to predict the noise levels due to continuous and non-continuous traffic flow in India. It was observed that there is no significant difference between the observed and predicted the output parameters. Homoda (2008) used a back-propagation neural network (BPN) and a regression model to predict the noise of construction in Kuwait. Although the regression model was more accurate than the neural network in predicting the noise level, the results showed that the neural networks can be used to predict the noise level of construction in the environmental studies. Givargis and Karimi (2010) provided a neural network for predicting the sound pressure level of Tehran roads by taking into account the average traffic speed lower than 75 kilometers per hour. The results showed that the neural network can be used to predict the level of traffic noise in Tehran. Gennaro et al. (2009) developed an artificial neural network for the prediction of urban noise, based on 25 environmental features. The results obtained by the neural network were compared with classic models and statistical tests and confirmed that the results of the neural network were better than in all the classic models. Kumar, Nigam and Kumar (2014) provided a back-propagation neural network (BPN) to predict (L_{10}) and (LA_{eq}) by including parameters such as traffic volume, average speed, and the percentage of heavy vehicles in Delhi. The results

were compared with the regression models built with the same parameters and proved the superiority of the neural network model in estimation and prediction of the traffic noise.

Some researchers have compared the performance of multilayer perceptron and radial basis function networks in different fields. Memarian and Balasundram (2012) compared the radial and multilayer neural network to estimate the sediment load in a tropical watershed in Malaysia and Capila et al. (2015) compared the radial and multilayer neural network to predict air pollution. The results of the two studies indicated that multilayer neural network model was more accurate than the radial network model. In this study, we will compare the multilayer neural network and radial neural network in terms of their efficiency in predicting the noise pollution caused by traffic.

MATERIALS AND METHODS

Noise measurements

Fifty-one samples were taken from 34 points, from the west and northwest areas of Tehran considering the traffic volume, the average speed of vehicles and percentage of heavy vehicles. Field measurements were conducted from 7 a.m. to 8 p.m. on working days in June. A-weighted Noise levels were measured using a sound level meter (Lutron SL-4023SD) placed at the height of 1.2 meters from the road surface and at a distance of 2 meters from the edge of the carriageway (Fig. 1).



Fig. 1. Placement of sound level meter near the edge of carriageway

Noise measurement period was 5 minutes for each point in fast time weighting mode. Calculation of the equivalent continuous sound level for the period (0 to T) is as follows (Eq. 1): (Management and Planning Organization of Iran, 2006).

$$LA_{eq} = 10 \log \left(\frac{1}{T} \int_0^T \frac{P^2(t)}{P_0^2} dt \right) \quad dB \quad (1)$$

where: T is the measurement duration in seconds, $P(t)$ is instantaneous sound pressure in Newton per square meter (N/m^2) and P_0 is the reference sound pressure, equals to 2×10^{-5} Newton per square meter (N/m^2).

Traffic parameters

In order to collect traffic data, a video camera was placed on a pedestrian bridge for recording the traffic flow and conducting the sound measurement at the same time. The total traffic flow (Q) was measured by counting the vehicles in the videos for the duration of one hour; the percentage of heavy vehicles (P) including buses and trucks

was calculated by dividing their volume by the total traffic volume. The average speed of vehicles (V) obtained by considering a certain distance on the road and measuring the vehicle crossing time and dividing the considered distance by the crossing time. Statistical descriptions of data are shown in Table 1.

Multilayer neural networks (MLP)

The multilayer perceptron is one of the most popular types of artificial neural networks consisting of one or more layers (Rumelhart et al., 1986). MLP network is widely used with three layers including one input layer, one hidden layer and one output layer. The architecture of a typical MLP Network is shown in Figure 2.

The neurons in the hidden and output layers contain an activation function. The activation function of the output layer neurons is linear, but in the neurons of hidden layer, it is nonlinear (usually a sigmoid or hyperbolic tangent). According to Figure 3, the data fed through the input layer is scaled by an initial weighting through the connec-

Table 1. Variables statistical information

Parameter	Q	V	P	LAeq
Mean	3285.99	59.19455	1.886133	71.56608
Median	2536.5	63.835	1.710786	72.26
Standard Deviation	2109.433	17.39559	1.01165	3.754818
Sample Variance	4449708	302.6066	1.023436	14.09866
Kurtosis	1.781876	-0.75243	-0.6548	-1.1684
Skewness	1.563322	0.024895	0.479708	-0.07878
Range	8343.5	70.452	3.971599	13.93
Minimum	1205	24.435	0.313637	64.59
Maximum	9548.5	94.887	4.285236	78.52
Count	51	51	51	51

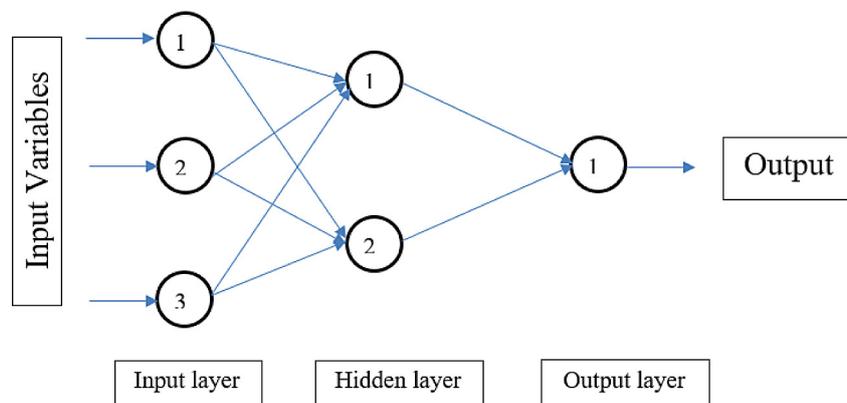


Fig. 2. Architecture of a typical Multilayer Perceptron (MLP) Network

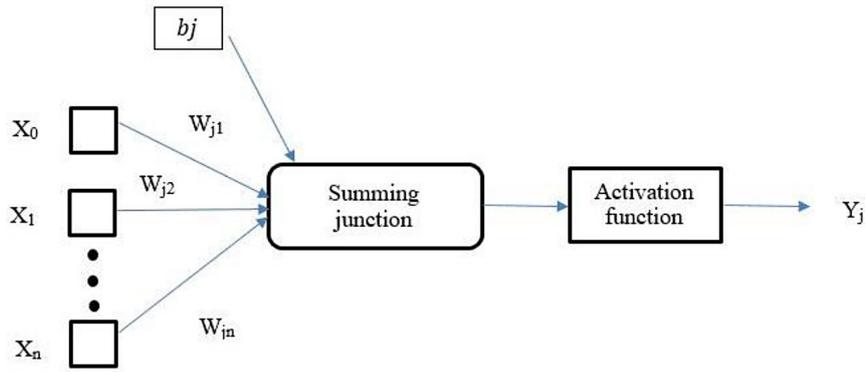


Fig. 3. Computation process of a neuron at hidden layer

tion between neurons. The relationship between inputs (X_i), weights (W_{ji}) and outputs (Y_j) is calculated according to Eqs. (2, 3) (Haykin, 1999; Demuth and Beale, 1998).

$$H_j = \sum_{i=0}^n W_{ji} X_i + b_j \quad (2)$$

$$Y_j = f(H_j) \quad (3)$$

where: b_j is the bias and f is the activation function which is considered as a sigmoid function in this research and calculated by the Eq. (4)

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

Out of the different training methods such as Levenberg-Marquardt, gradient descent and Gauss-Newton, the Levenberg – Marquardt (Levenberg, 1944; Marquardt, 1963) based on back propagation method has been selected in this study due to faster convergence in training medium-sized networks. The Levenberg-Marquardt method is a combination of gradient descent and Gauss-Newton method. The brief description of gradient descent is that it changes network weights and bias values in the direction of the negative of gradient which decreases error function (Rumelhart et al. 1986). Weight adjustment in gradient descent method is as follows (Eq. 5):

$$w_{K+1} = w_K - \alpha_k g_k \quad (5)$$

where: w_k is the vector of weights, α_k is the learning rate and g_k is the gradient of error in the k-th iteration.

The Newton’s method is faster than gradient descent in convergence by using the second order derivative and it is according to Eq. (6)

$$w_{K+1} = w_K - H^{-1} g_k \quad (6)$$

where: H is the Hessian matrix.

Levenberg – Marquardt algorithm uses the following equation to achieve faster network training (Eq. 7):

$$w_{K+1} = w_K - [H + \mu I]^{-1} g_k \quad (7)$$

where: I is the Identity matrix and the mixing level between gradient descent and Newton method is determined by μ

Radial basis function

The radial basis function (RBF) network is another popular type of neural network which is structurally similar to the classical regularization network. It is based on an iterative function approximation and localized basis functions. This type of neural network has a simpler architecture and training process compared to MLP network. Input data are directly transferred to the hidden layer. Using a nonlinear function in the hidden layer, the data are transferred to the hidden space. The process of training in this layer is an unsupervised method. Finally, with a linear transformation function, the network response is achieved in the output layer. RBF activation function in each hidden unit calculates the Euclidean distance between the input vector and the center of that unit by using the Gaussian function. In fact, the output of each hidden unit is the distance between the vector of input data and the center of the hidden unit. The training process of network parameters (weights) between the hidden layers and the output layer is supervised. In addition to the adjustment of weights, modifying the center of the activation function is needed in training the network.

The weights and the center of activation functions are adjusted using the gradient descent method to minimize the Sum of Squared Error (SSE). The hidden layer function is calculated by the Eq. (8):

$$F(x) = \sum_{i=1}^N w_i \phi(\|x - u_i\|) \quad (8)$$

where: N is the number of neurons in the hidden layer,
 u_i is the center vector for neuron i , and
 w_i is the weight of neuron i in the linear output neuron.

According to Eq. (8) N number of radial functions ($\{\phi(\frac{\|x - u_i\|}{i} = 1, 2, \dots, N)\}$) are used for approximating the function $F(x)$. The norm is usually the Euclidean distance in real coordinate space (R^n). (Yu et al., 2011; Dayhoff, 1990; Broomhead and Lowe, 1988).

The Activation function (ϕ) is in the form of Gaussian function according to the Eq. (9):

$$\phi(x) = a \exp\left(-\frac{(x - \mu)^2}{2\sigma_i^2}\right) \quad (9)$$

The graph of a Gaussian is a characteristic symmetric “bell curve” shape. The parameter (a) is the height of the curve’s peak, μ is the position of the center of the peak and σ (the standard deviation) is the opening quantity of bell

shaped function. Figure 4 shows different shapes of Gaussian function for different value of μ and σ (Wikipedia, “Gaussian function”, 2016)

A typical radial network is shown in Figure 5. The input signals are fed directly into the cells of the hidden layer. Unlike MLP network with global activation function, in RBF network the activation function is local. The number of hidden layer neurons is obtained through trial and error. The output layer is only a collector, so that its inputs are the outputs of the hidden layer neurons.

Model development

The two models were implemented in MATLAB version R2014b. In order to develop the considered models, it is necessary to utilize a data set. In this study, a set of 51 data collected from 34 points in different locations of west and north-west of Tehran was utilized.

Application of MLP

Datasets were randomly divided into three subsets; training, validation and test. The dataset was divided into three subsets of 80%, 10% and 10% respectively for training, validation and testing. The training subset is used to train the model, while the validation subset, which is separated from the training subset, is used to check the network validity. The error on the validation set is observed during the training process to prevent

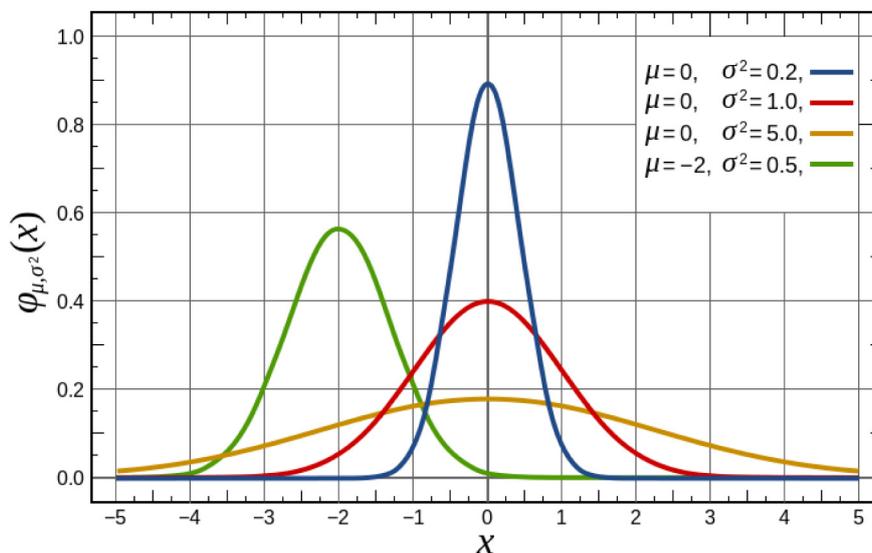


Fig. 4. Normalized Gaussian curves with expected value μ and variance σ^2 (Wikipedia, “Gaussian function”, 2016)

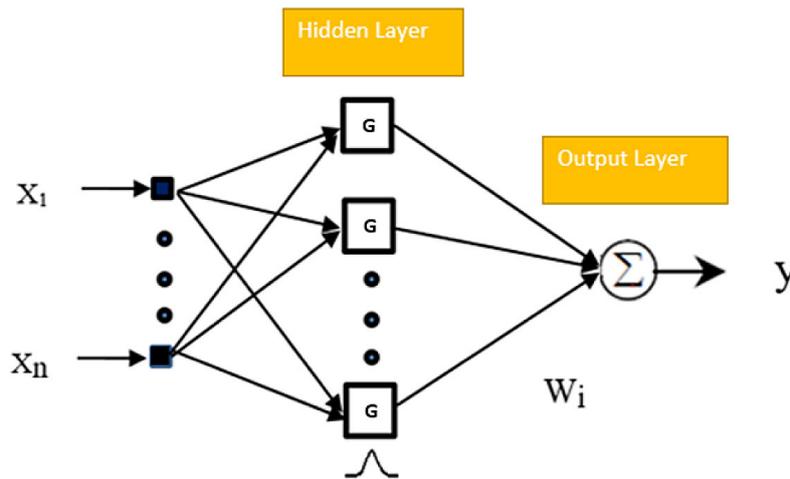


Fig. 5. A typical RBF architecture

the overtraining. The validation and training errors decrease during the initial phase of training. However, the error on the validation set typically begins to increase when the network starts to overfit the data. When the validation error raises for a specified number of iterations, the training is stopped, and the weights and biases giving the minimum of the validation error are returned (Levenberg, 1944). A network test was conducted using the remaining 10% of the total data.

Application of RBF

Network training and testing were performed using the same datasets applied to the MLP network. In order to increase the accuracy of RBF network, two parameters should be optimized: Spread and Maximum Number of Neurons (MNN). In this study, for training RBF network with appropriate generalization ability, the target error was first considered zero, then the spread value and the maximum number of neurons in the hidden layer were chosen by trial and error in order to minimize the Mean Squared Error (MSE) in the training and validation sets. Finally, the performances of the two neural networks were compared based on MSE (Eq.10) and R^2 .

$$MSE = \frac{1}{N} \sum_{i=1}^N ((LA_{eq})_m - (LA_{eq})_p)^2 \quad (10)$$

where: N is the number of observations, $(LA_{eq})_m$ is the measured A-weighted continuous sound level and $(LA_{eq})_p$ is the predicted A-weighted continuous sound level.

RESULTS AND DISCUSSION

Multi-layer perceptron neural network results

Various neural network architectures with a different number of hidden neurons and hidden layers were trained, validated and tested for developing the MLP network. An ANN architecture with 7 neurons in the hidden layer delivered the best results and was subsequently selected (Fig. 6). Then the network was trained and tested 100 times (100 iterations) and among them, the best network performance based on the mean squared error (MSE) and the correlation coefficient was selected. The results showed that the error of the model was in the range of -1.99 to +1.92 dB and correlation coefficient was 0.973 (Fig. 7). The side by side comparison of measured data and MLP prediction results is shown in Fig. 8. The results of the MLP network model demonstrate good agreement with the measured values.

Radial basis function network results

The characteristics of the radial network are shown in Table 2. The error of the RBF network with 41 neurons in the hidden layer reached the lowest level (0.009) in the training process (Fig. 9).

After training and testing the network, the results showed that the network error was between -5.62 dB and +1.19 dB and the correlation coefficient was equal to 0.928 (Fig. 10). The side by side comparison of measured data and RBF prediction results is shown in Fig. 11. It is obvious that the overlapping of predicted data and the measured data in both networks is acceptable.

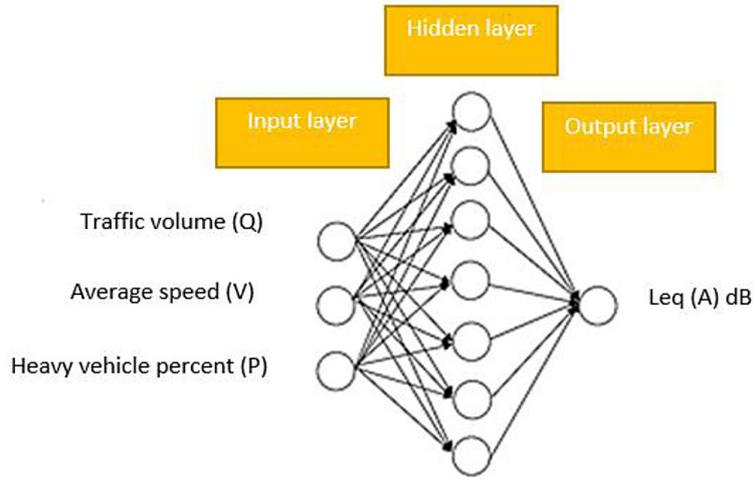


Fig. 6. Proposed MLP architecture for traffic noise modeling (3–8-1)

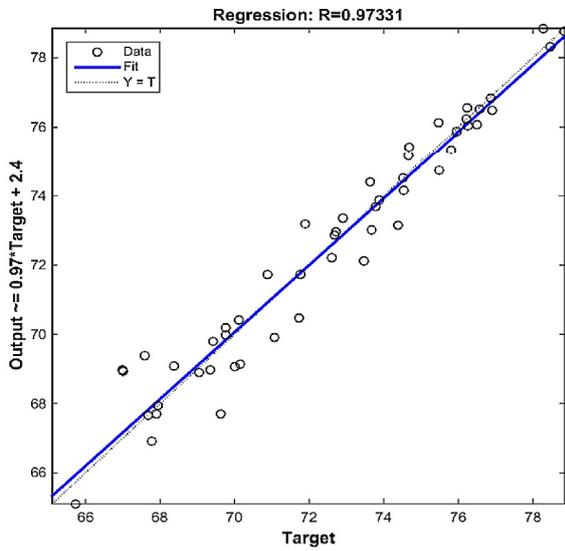


Fig. 7. Comparison of the observed LAeq (dB(A)) and predicted LAeq (dB(A)) by MLP neural network

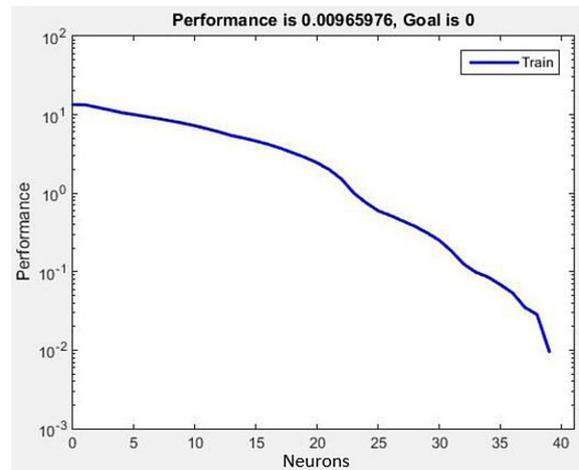


Fig. 9. Error variation of radial basis function neural network in the training phase

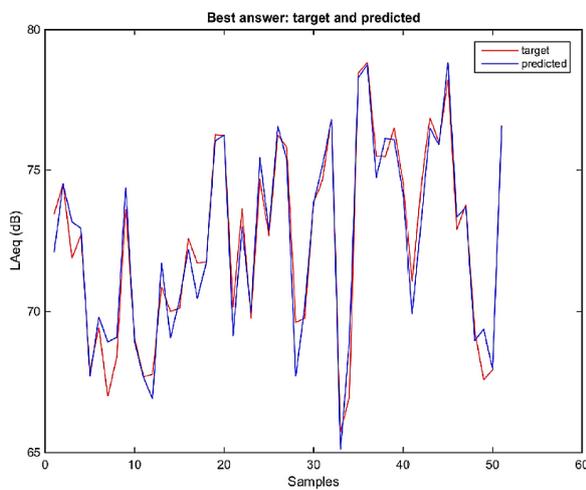


Fig. 8. Side by side comparison of the measured data and MLP prediction results

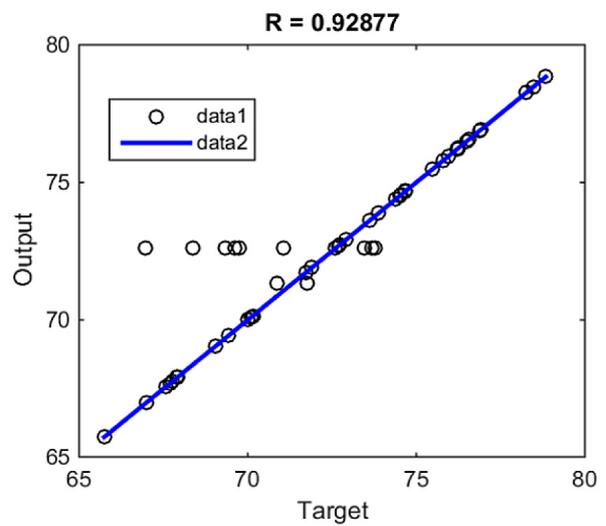


Fig. 10. Comparison of Observed LAeq (dB(A)) and predicted LAeq (dB(A)) by RBF neural network

Table 2. Characteristic of RBF network model

Number of inputs	Number of outputs	Maximum No. of neurons in hidden layer(MNN)	Spread	Activation function
3	1	51	1	Gaussian

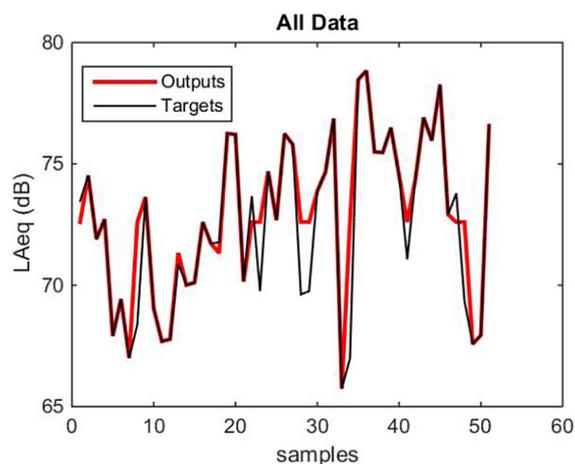


Fig. 11. Side by side comparison of measured data and RBF prediction results

The results of the mean squared error and correlation coefficient for the two models are specified in (Table 3). Multilayer perceptron neural network had a better performance with lower MSE (0.6292) and higher R^2 (0.947) in predicting the level of noise pollution caused by traffic.

CONCLUSIONS

In this study, the data were collected from 34 different points in the west and northwest areas of Tehran. Two types of neural network including multi-layer perceptron network (MLP) and radial network (RBF) were developed to predict the equivalent continuous sound level caused by traffic, by taking into account parameters such as traffic volume, average speed and the percentage of heavy vehicles. The collected data were divided randomly into three data subsets by percentage of 80, 10 and 10 for training, validation and testing the networks, respectively. The best architectures for both networks were determined

by trial and error. Finally, the performances of the two networks were compared based on two criteria, Mean Squared Error (MSE) and coefficient of determination (R^2). The results showed that despite a good overlap of predicted data on the measured data in both networks, MLP network with the MSE equal to 0.6292 and the R^2 equal to 0.947 had better a performance than the RBF network with MSE=1.786 and the $R^2=0.8626$ in the prediction of sound level.

REFERENCES

1. Broomhead D.A. and Lowe D. 1988. Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal Signals and Radar Establishment, Memorandum, GB.
2. Cammarata, G., Cavalieri, S., Fichera, A., 1995. A neural network architecture for noise prediction. *Neural Networks*, 8(6), 963–973.
3. Capilla, C. 2015. Application of radial basis functions compared to neural networks to predict air pollution. *Air Pollution XXIII*.
4. Caponetto, R.; Lavorgna, M.; Martinez, A.; Occhipinti, L. 1997. GAS for fuzzy modeling of noise pollution. *Proceedings of the First International Conference on Knowledge-Based Intelligent Electronic Systems*, Adelaide (Australia), 219–223.
5. Dayhoff, 1990. *Neural network architectures*. New York: Van Nostrand Reinhold.
6. Demuth, H. and Beale, M. 1998. *Neural network toolbox for use with MATLAB*. Natick, Mass.: MathWorks, Inc.
7. Genaro, N., Torija, A., Ramos, A., Requena, I., Ruiz, D.P., Zamorano, M., 2009. Modeling Environmental Noise Using Artificial Neural Networks. *Ninth International Conference on Intelligent Systems Design and Application*, 215–219.
8. Givargis Sh., and H. Karimi, 2010. A basic neural traffic noise prediction model for Tehran’s roads. *Journal of Environmental Management*, 9(1), 2529–2534,

Table 3. Comparison of RBF and MLP Model.

Model	Range of Errors	Mean Squared Error (MSE)	Correlation coefficient (R)	Determination coefficient R^2
MLP	(-1.99 , +1.92)	0.6292	0.973	0.947
RBF	(-5.62 , +1.92)	1.786	0.928	0.8626

9. Hamoda, Mohamed F., 2008. Modeling of construction noise for environmental impact assessment. *Journal of Construction in Developing Countries* 13, 79–89.
10. Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, NJ, USA.
11. Kumar, P., Nigam, S. and Kumar, N. 2014. Vehicular traffic noise modeling using artificial neural network approach. *Transportation Research Part C: Emerging Technologies*, 40, 111–122.
12. Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *J. Appl. Math.* 2, 164–168.
13. Management and Planning Organization of Iran, (2006). Issue No. 342: Acoustical guidelines for Reduction of Traffic Noise for Buildings near Highways. Tehran.
14. Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11(2), 431–441.
15. Memarian, H. and Balasundram, S. 2012. Comparison between Multi-Layer Perceptron and Radial Basis Function Networks for Sediment Load Estimation in a Tropical Watershed. *JWARP*, 4(10), 870–876.
16. Parabat D.K. and P.B. Nagamail, 2007. Assessment and ANN Modelling of Noise Levels at Major Road intersections in an Indian intermediate City. *Journal of Research in Science, Computing, and Engineering*, 4(3), 39–49,
17. Rumelhart E., McClelland J.L. and the PDP Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1: Foundations. MIT Press, Cambridge.
18. Wikipedia, “Gaussian function”, (2016). [Online]. Available: https://en.wikipedia.org/wiki/Gaussian_function. (Accessed: 04-Jan-2016).
19. Yu H., Xie T., Paszczynski S. and Wilamowski B., 2011. Advantages of Radial Basis Function Networks for Dynamic System Design, *IEEE Trans. Ind. Electron.*, 58(12), 5438–5450.