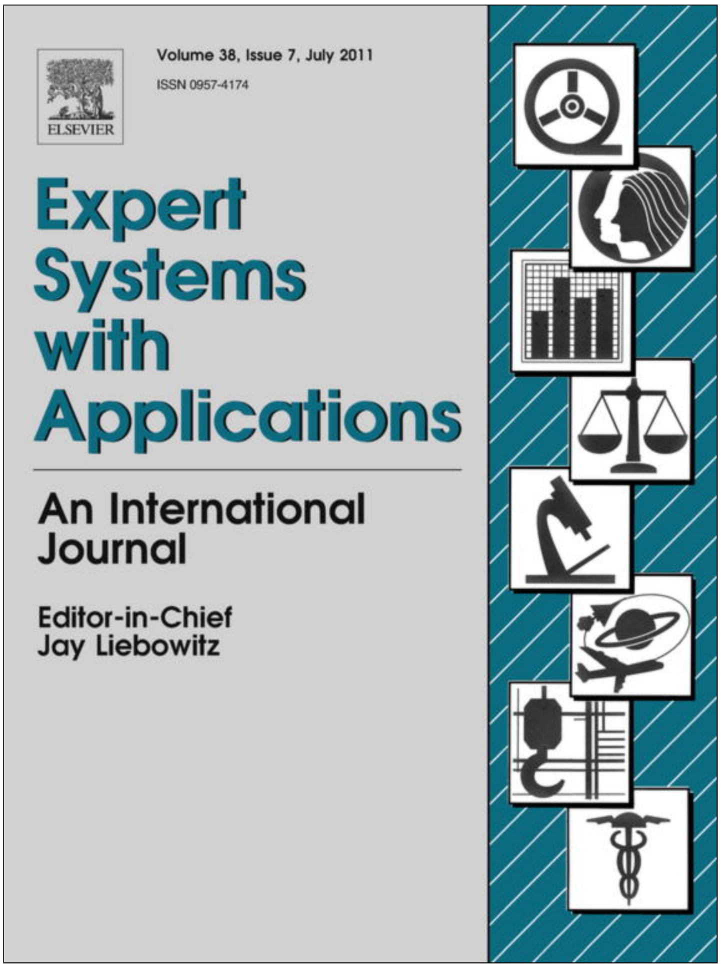
Volume 38, Issue 7, July 2011
ISSN 0957-4174
ELSEVIER
Expert Systems with Applications
An International Journal
Editor-in-Chief
Jay Liebowitz

Expert Systems with Applications 38 (2011) 8999–9007

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Prediction of tractor repair and maintenance costs using Artificial Neural Network

Abbas Rohani [a,*], Mohammad Hossein Abbaspour-Fard [b], Shamsolla Abdolahpour [c]

[a] Dept. of Farm Machinery Engineering, College of Agriculture, Shahrood University of Technology, Shahrood, Iran
[b] Dept. of Farm Machinery Engineering, College of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran
[c] Dept. of Farm Machinery Engineering, University of Tabriz, Tabriz, Iran

ARTICLE INFO

Keywords:
ANN
BDLRF algorithm
Basic Back-propagation
Repair and maintenance cost

ABSTRACT

The prediction of repair and maintenance costs has significant impacts on proper economical decisions making of machinery managers, such as machine's replacement and substitution. In this article the potential of Artificial Neural Network (ANN) technique has evaluated as an alternative method for the prediction of machinery (specifically tractor) repair and maintenance costs. The study was conducted using empirical data on 60 two-wheel drive tractors from Astan Ghodse Razavi agro-industry in Iran. Optimal parameters for the network were selected via a trial and error procedure on the available data. In this paper, the performance of Basic Back-propagation (BB) training algorithm was also compared with Back-propagation with Declining Learning-Rate Factor algorithm (BDLRF). It was found that BDLRF has a better performance for the prediction of tractor's costs. The prediction of repair and maintenance cost components of tractors with a single network produced a better result than using separate networks for prediction of each cost component. It has been concluded that ANN represents a promising tool for predicting repair and maintenance costs.

1. Introduction

The greatest returns in farming are gained from 'brain activity' rather than 'brawn activity'. In other words, although a large part of farming is the implementation of production practices, i.e. planting and harvesting, the greatest income of farming are coming from activities such as decisions making (Jensen, 1977). Forecast is simply a prediction of what will happen (Makridakis & Wheelwright, 1989). In this regard, economic forecasting is the study of historic data to discover their underlying tendencies and patterns (Hanke & Reitsh, 1995). Forecast can be the key of success or failure of a plan. To understand why it is important to have an accurate method of predicting tractor's repair and maintenance costs, it is necessary to first have an understanding of what these forecast costs can be used for. Repair and maintenance costs are important component of total tractor ownership costs and they are important to consider, both in their timing and magnitude. When a tractor is approaching the end of its economical life, the farm manager must make a replacement decision. The decrease in ownership costs with the concurrent increase in operating costs gives rise to the notion of economic life. There is theoretically, an optimum age to replace a tractor. To properly analyze economical life, one must be armed with detailed knowledge of the elements and behavior of owning and operating costs. Ownership costs are not too difficult to understand and quantitate, instead operating costs are complex and highly intensive in data. If the operating costs stream is properly tracked and analyzed, it can be a reliable input into the economic modeling process.

The mathematical models proposed in the literature are very simplistic, old, and broad in scope. Moreover, these models are seldom, if ever, used in practice. Regression models were firstly employed for prediction of repair and maintenance costs of farm machinery by American Society of Agricultural Engineers (ASAE) (Bowers & Hunt, 1970), since then it has been continued by others (Fuls, 1999; Morris, 1988; Rotz, 1987).

Owing to natural uncertainties associated with repair and maintenance costs of different operations of tractor, exact mathematical relationships between these parameters and tractor age are difficult to be derived. Hence, recourse is normally made to the statistical technique of non-linear regression. Despite this, the resulting equations suffer from approximation and unreliability. An attempt is therefore made in this paper to provide an alternative to these conventional statistics-based methods, by adapting Artificial Neural Network (ANN).

Although the concept of ANN analysis was almost discovered 50 years ago, it is only in the last two decades that its application software has been developed to handle practical problems. ANN basically provides a non-deterministic mapping between sets of random input–output vectors. Absence of any preliminary assumed relationship beforehand between input–output quantities, in-built dynamism and robustness towards data errors, are some advantages of these networks over statistical methods. The ANN approach has been successfully employed in farm engineering.

* Corresponding author. Tel.: +98 2733332204.
E-mail address: abassrohani@yahoo.com (A. Rohani).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.01.118

However, its applicability to many problems in farm machinery management analysis is, yet to be properly established. The present analysis is an effort in that direction. The ANN mimics somewhat the learning process of a human brain. Neural networks are composed of simple elements operating in parallel. These elements are inspired by biologic nervous systems. As in nature, the network function is determined largely via connections between its elements. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between the elements. In addition, inherently noisy data do not seem to create a problem, as ANNs are tolerant to noise variations. Neural networks are being used in a wide variety of applications as an effective decision making tool. This can be attributed to the fact that these networks are attempted to model the capabilities of human brain. They are being used in the areas of prediction and classification; the areas where statistical prediction and classification techniques have traditionally been used.

The main advantages of using neural networks are learning directly from examples without attempting to estimate the statistical parameters. More generally, there is no need for firm assumptions about the statistical distributions of the inputs and generating any continuous nonlinear function of input (universal approximating). ANNs are highly parallel which makes them especially amenable to high-performance parallel architectures (Gupta, Jin, & Homma, 2003; Vakil-Baghmisheh, 2002). Because of these unique characteristics, it also can be employed for prediction of repair and maintenance costs of tractor.

In spite of different structures and training paradigms, all ANNs essentially perform the same function, vector mapping, i.e. accepting an input vector and producing an output vector. Likewise, all ANN applications are special cases of vector mapping (Vakil-Baghmisheh, 2002). In this study, the most common type of network, namely, Multilayer Perceptron (MLP) is used. It supposedly has the ability to approximate any continuous function. The input nodes receive the input values and pass them to the hidden nodes, which multiply the input by connection weights. Subsequently, adding up such product and attaching a bias, the result is then transformed through a transfer function. Before it is put into actual operation, the network weight and bias values should be fixed. This can be established, by employing a training algorithm and using a set of known input–output patterns, until the error between the network-generated and the actual output reaches a minimum. Among different available algorithms, Basic Back-propagation (BB) and Back-propagation with Declining Learning-Rate Factor (BDLRF) were used in this study. This is mainly because, the former is most common and the latter is more efficient. Use of both training schemes insured that the desired training was correctly established. The details could be obtained elsewhere (e.g. Vakil-Baghmisheh & Pavešic, 2001).

The main objective of this study was to develop tractor's repair and maintenance costs prediction models with readily available data that could be easily applied by a farm manager. The specific objectives were: (1) to investigate the effectiveness of ANN for predicting repair and maintenance costs for typical conditions using field-specific costs and historic costs data; (2) study the variation of model performance with different ANN model parameters; (3) select optimum ANN parameters for accurate prediction of repair and maintenance costs.

2. Materials and methods

2.1. Data recording

First of all, it has to be assumed that the data were completely and accurately collected by the company (Astan Ghodse Razavi). It is not possible to go back and verify all expenditures; hence the existing records have to be trusted. However, the trustworthiness of these records was verified by visiting the company and method of its recording data. Historical data (1986–2003) of repair and maintenance costs of tractors was obtained from Astan Ghodse Razavi agro-industry Company in Iran. Records of the repair and maintenance costs, including parts, labor, fuel and oil, were available for 60 two-wheel drive (2WD) tractors, over 18 years. These tractors were used for various operations such as tillage, planting and harvesting as well as transportation.

The available data contain: monthly usage, monthly repair costs (including parts and labor), monthly maintenance costs (including fuel, oil, fuel filter and oil filter), year of purchase and tractor make and model. The data were shuffled and split into two subsets: a training set and a test set. The splitting of samples plays an important role in the evaluation of an ANN performance. The training set is used to estimate model parameters and the test set is used to check the generalization ability of the model. The training set should be a representative of the whole population of input samples. In this study, the training set and the test set includes 130 patterns (60% of total patterns) and 86 patterns (40% of total patterns), respectively. There is no acceptable generalized rule to determine the size of training data for a suitable training; however, the training sample should cover all spectrums of the data available (NeuroDimensions Inc, 2002). The training set can be modified if the performance of the model does not meet the expectations (Zhang & Fuh, 1998). However, by adding new data to the training samples, the network then can be retrained.

2.2. Tractors differences

In this study, we are aiming to provide an effective tool for accurately forecasting repair and maintenance costs of tractors. Repair and maintenance costs, as well as initial purchase price, can differ considerably among different models of tractor. Despite this variation, it is essential to create the accumulative repair and maintenance costs of different models of tractor relatively. The convenient way of comparing the repair and maintenance costs of dissimilar tractors is to index them to their initial price. Subsequently, the repair and maintenance costs of different tractors can be compared by means of cumulative cost index (Mitchell, 1998). In this regard, the cumulative cost index (CCI) can be calculated as

$$\mathrm{CCI}_t = \frac{\sum_{o}^{t}(P_t + L_t + O_t)}{\mathrm{PP_o}} \tag{1}$$

where CCI_t is the cumulative cost index at time "t", P_t and L_t are costs of tractor parts and labor at time "t" respectively; O_t is the other miscellaneous maintenance costs including fuel, oil, fuel filter and oil filter at time "i" and $\mathrm{PP_o}$ is the initial purchase price of tractor. CCI is the output of network and bearing in mind that it should not decrease with increasing tractor age, but it may increase or remain constant with tractor age.

2.3. Tractor age

The tractor age is considered as input of network and may be defined in various terms. These are the calendar age of tractor, tractor age as units of production, and tractor age as cumulative hours of usage (Mitchell et al., 1998). The calendar age is conveniently obtained by subtracting the original purchase date from the current date. Because of natural uncertainties associated with tractor repair and maintenance costs, they do not accrue as a result of elapsed calendar time. Tractor age as units of production is the measure of amount of work a tractor has actually accomplished. Defining tractor age as units of production is difficult and may be defined in a number of ways. It could be in terms of working area, working hours, traveling distance, etc. The actual quantitation of production units can also be a difficult task. Tractor age as cumulative hours of use is a measure of how many hours the tractor physically operated. It dampens many of the cyclical variations in repair and maintenance costs. Considering the characteristics of three types of tractor age, the cumulative hours of use was chosen. The data for tractor age in cumulative hours of use is not always easy to obtain but it can be available. The company under study follows up oil-changing program. The times of oil-change in machines' life are usually recorded in terms of a calendar date. Considering the calendar date of the engine oil change and the associated monthly cost data, the cumulative costs for a given number of cumulative hours is determined.

2.4. Inflation effect

The impact of inflation can be a major concern when trying to make a conscious business decision regarding cash flows that takes place over any appreciable length of time. The company under study keeps the tractors for at least 12 years. During this time, the economy could be subjected to any number of twists and turns. If the original purchase year is used as the base year, then the inflation-adjusted cost per month is c_{ki} which is calculated as

$$c_{ki} = c_k \times (1 + I_{ga})^{n-k}, \quad k = 1, 2, \ldots, n \tag{2}$$

where c_k, n and I_{ga} are the monthly cost, total tractor life in month, and the average yearly inflation rate, respectively. The cumulative cost per month (cc_{ki}) is then calculated as

$$cc_{ki} = c_{k-1i} + c_{ki} \tag{3}$$

Consequently, the cumulative cost index per month (CCI_k) can be calculated as

$$\mathrm{CCI}_k = \frac{cc_{ki}}{\mathrm{PP_o}} \times 100 \tag{4}$$

2.5. Data preprocessing

Based on these available data, the cumulative hour of usage as a percentage of 100 h (CHU) was selected as variable input. The cumulative repair cost index ($\mathrm{CCI_{repair}}$), the cumulative oil cost index ($\mathrm{CCI_{oil}}$), the cumulative fuel cost index ($\mathrm{CCI_{fuel}}$), and the cumulative repair and maintenance cost index ($\mathrm{CCI_{rm}}$) were selected as variable outputs. Prior to any ANN training process with the trend free data, the data must be normalized over the range of [0,1]. This is necessary for the neurons' transfer functions, because a sigmoid function is calculated and consequently these can only be performed over a limited range of values. If the data used with an ANN are not scaled to an appropriate range, the network will not converge on training or it will not produce meaningful results. The most commonly employed method of normalization involves mapping the data linearly over a specified range, whereby each value of a variable x is transformed as follows

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \times (r_{\max} - r_{\min}) + r_{\min} \tag{5}$$

where x is the original data, x_n the normalized input or output values, $x_{\max}$ and $x_{\min}$, are the maximum and minimum values of the concerned variable, respectively. $r_{\max}$ and $r_{\min}$ correspond to the desired values of the transformed variable range. A range of 0.1–0.9 is appropriate for the transformation of the variable onto the sensitive range of the sigmoid transfer function.

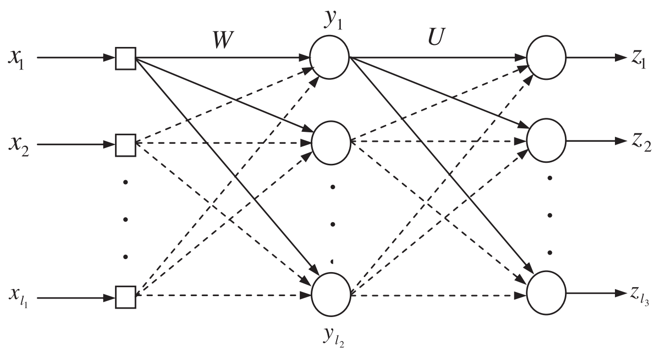

Fig. 1. Configuration of the MLP with one hidden layer (Vakil-Baghmisheh, 2002).

2.6. The multilayer perceptron neural network

Among various ANN models, Multilayer Perceptron (MLP) has maximum practical importance. MLP is a feed-forward layered network with one input layer, one output layer, and some hidden layers. Fig. 1 shows a MLP with one hidden layer. Every node computes a weighted sum of its inputs and passes the sum through a soft nonlinearity. The soft nonlinearity or activity function of neurons should be non-decreasing and differentiable. The most popular function is unipolar sigmoid:

$$f(\theta) = \frac{1}{1 + e^{-\theta}} \tag{6}$$

The network is in charge of vector mapping, i.e. by inserting the input vector, x^q the network will answer through the vector z^q in its output (for $q = 1,\ldots,Q$). The aim is to adapt the parameters of the network in order to bring the actual output z^q close to corresponding desired output d^q (for $q = 1,\ldots,Q$). The most popular method of MLP training is the back-propagation algorithm, and in literatures there exist many variants of this algorithm.

This algorithm is based on minimization of a suitable error cost function. In this study, two variants of MLP training algorithm, i.e. Basic Back-propagation (BB) and Back-Propagation with Declining Learning-Rate Factor (BDLRF) where employed. A computer code was also developed in MATLAB software to implement these ANN models.

2.6.1. BB algorithm

In this algorithm the total sum-squared error (TSSE) is considered as the cost function and can be calculated as

$$\mathrm{TSSE} = \sum_q E_q \tag{7}$$

$$E_q = \sum_k (d_k^q - z_k^q)^2 \quad \text{for } (q = 1, \ldots, Q) \tag{8}$$

where d_k^q and z_k^q are the kth components of desired and actual output vectors of the qth input, respectively. Network learning happens in two phases: forward pass and backward pass. In forward pass an input vector is inserted to the network and the network outputs are computed by proceeding forward through the network, layer by layer:

$$\begin{cases} \mathrm{net}_j = \sum_i x_i w_{ij} \\ y_j = \frac{1}{1+e^{-\mathrm{net}_j}} \end{cases}, \ j = 1, \ldots, l_2 \tag{9}$$

$$\begin{cases} \mathrm{net}_k = \sum_j y_j u_{jk} \\ z_k = \frac{1}{1+e^{-\mathrm{net}_k}} \end{cases}, \ k = 1, \ldots, l_3 \tag{10}$$

Table 1
Performance variation of a three-layer BB-MLP with different number of neurons in the hidden layer.

Parameters	Criterion	Number of neurons in the hidden layer								
		2	3	4	5	6	7	8	9	10
CCI_{repair}	MAPE (%)	24.81	15.94	13.52	**12.80**	15.63	16.20	14.93	14.16	13.33
	RMSE	0.8204	0.7097	0.6913	**0.5905**	0.7280	0.6843	0.6913	0.6630	0.6555
	TSSE[a]	0.0034	0.0025	0.0024	**0.0022**	0.0026	0.0024	0.0023	0.0024	0.0022
	R	0.9995	0.9995	0.9996	0.9998	0.9998	0.9998	0.9999	0.9999	0.9999
CCI_{oil}	MAPE (%)	6.78	3.71	4.48	6.14	**4.18**	4.24	4.71	4.20	3.80
	RMSE	0.2035	0.1875	0.1866	0.1860	**0.1832**	0.1825	0.1820	0.1905	0.1870
	TSSE	0.0093	0.0080	0.0079	0.0072	**0.0063**	0.0076	0.0073	0.0069	0.0069
	R	0.9994	0.9994	0.9995	0.9995	**0.9997**	0.9997	0.9999	0.9999	0.9999
CCI_{fuel}	MAPE (%)	1.99	1.90	2.65	2.07	**1.83**	2.50	2.29	1.85	2.10
	RMSE	0.9076	0.9350	0.0987	0.0952	**0.0913**	0.0998	0.0993	0.0955	0.0970
	TSSE	0.0009	0.0011	0.0011	0.0011	**0.0009**	0.0011	0.0011	0.0010	0.0011
	R	0.9997	0.9997	0.9998	0.9998	**0.9999**	0.9999	0.9999	0.9999	0.9999
CCI_{rm}[b]	MAPE (%)	16.04	14.01	11.87	11.59	**10.00**	14.22	16.02	13.02	13.25
	RMSE	0.8000	0.7870	0.7666	0.7555	**0.7078**	0.7321	0.7493	0.7683	0.7644
	TSSE	0.0017	0.0016	0.0021	0.0020	**0.0017**	0.0015	0.0015	0.0018	0.0018
	R	0.9998	0.9998	0.9998	0.9997	**0.9999**	0.9999	0.9999	0.9999	0.9999

[a] TSSE is estimated in the training phase.
[b] $CCI_{rm} = CCI_{repair} + CCI_{oil} + CCI_{fuel}$.

where w_{ij} is the connection weight between nodes i and j, and u_{jk} is the connection weight between nodes j and k; w_{ij} and u_{jk} are set to small random values [−0.25, 0.25]; l_2 and l_3 are the number of neurons in the hidden and output layers. In backward pass the error gradients versus weight values, i.e. $\frac{\partial E}{\partial w_{ij}}$ (for $i = 1,\ldots,l_1$, $j = 1,\ldots,l_2$) and $\frac{\partial E}{\partial u_{jk}}$ (for $j = 1,\ldots l_2$, $k = 1,\ldots,l_3$) are computed layer by layer starting from the output layer and proceeding backwards. The connection weights between nodes of different layers are updated using the following equations:

$$u_{jk}(n+1) = u_{jk}(n) - \eta \times \frac{\partial E}{\partial u_{jk}} + \alpha(u_{jk}(n) - u_{jk}(n-1)) \tag{11}$$

$$w_{ij}(n+1) = w_{ij}(n) - \eta \times \frac{\partial E}{\partial w_{ij}} + \alpha(w_{ij}(n) - w_{ij}(n-1)) \tag{12}$$

where η is the learning rate adjusted between 0 and 1, α is the momentum factor at interval (Bowers & Hunt, 1970). Momentum factor is used to speed up the convergence. The decision to stop training is based on some test results of the network, which is carried out every N epoch after TSSE becomes smaller than a threshold value. The details could be seen in Vakil-Baghmisheh and Pavešic (2003). The number of input and output nodes is determined by functional requirements of the ANN.

2.6.2. BDLRF algorithm

We have also used a modified version of BB algorithm which is Back-Propagation with Declining Learning-Rate Factor (BDLRF) algorithm (Vakil-Baghmisheh & Pavešic, 2001). This training algorithm is started with a relatively constant large step size of learning rate η and momentum term α. Before destabilizing the network or when the convergence is slowed down, for every T epoch ($3 \leqslant T \leqslant 5$) these values are decreased monotonically by means of arithmetic progression, until they reach to x% (equals to 5) of their initial values. η (and similarly α) was decreased using the following equations:

$$m = \frac{Q - n_1}{T} \tag{13}$$

$$\eta_n = \eta_o + n\eta_o \frac{x-1}{m} \tag{14}$$

where m, n_1, η_n and η_o are the total number of arithmetic progression terms, the start point of BDLRF, the learning rate in nth term of arithmetic progression, and the initial learning rate, respectively.

2.7. Performance evaluation criteria

To evaluate the performance of a model some criteria have been defined in the literature. These criteria include: total sum of squared error (TSSE), mean square error (MSE), mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), root mean square percentage error (RMSPE), coefficient of determination (R^2), etc. Among of them, MAPE, RMSE, TSSE and R^2 (the coefficient of determination of the linear regression line between the predicted values from the NN model and the actual output) are the most widely used performance evaluation criteria and may be used to compare the predicted and actual values which will be used in this study. They are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}\sum_{i=1}^{m}(d_{ji} - p_{ji})^2}{nm}} \tag{15}$$

$$R^2 = \frac{\left(\sum_{j=1}^{n}(d_j - \bar{d})(p_j - \bar{p})\right)^2}{\sum_{j=1}^{n}(d_j - \bar{d})^2 \cdot \sum_{j=1}^{n}(p_j - \bar{p})^2} \tag{16}$$

$$\mathrm{TSSE} = \sum_{j=1}^{n}(d_j - p_j)^2 \tag{17}$$

$$\mathrm{MAPE} = \frac{1}{nm}\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\frac{d_{ji} - p_{ji}}{d_{ji}}\right| \times 100 \tag{18}$$

where d_{ji} is the ith component of the desired (actual) output for the jth pattern; p_{ji} is the ith component of the predicted (fitted) output produced by the network for the jth pattern; $\bar{d}$ and $\bar{p}$ are the average of the desired output and predicted output, respectively; n and m are the number of patterns and the number of variable outputs, respectively. A model with the smallest RMSE, TSSE, MAPE and the largest R^2 is considered to be the best.

3. Results and discussion

Individual networks were developed in order to establish the relationships between (i) CCI_{repair} and CHU; (ii) CCI_{oil} and CHU; (iii) CCI_{fuel} and CHU; (iv) CCI_{rm} and CHU. Subsequently, a single network was developed for simultaneous prediction of CCI_{repair}, CCI_{oil} and CCI_{fuel}.

All networks were three-layered feed forward type, trained using both BB and BDLRF training algorithms.

3.1. MLP topology (number of neurons in the hidden layer)

Based on universal approximation theorem, a neural network with a single hidden layer and sufficiently a large number of neurons can well approximate any arbitrary continuous function (Haykin, 1994). Therefore, the ANNs designed in this study are equipped with a single hidden layer. Determination of the number of neurons in the hidden layer is rather an art than science, because it may vary depending on the specific problem under study. In this study, the optimal number of neurons in the hidden layer was selected using a trial-and-error method and keeping the learning rate, momentum term and epoch size constant ($\eta = 0.4$, $\alpha = 0.8$ and epoch = 10,000). The process was repeated several times, one for each set of data. Table 1 shows the effect of number of neurons in the hidden layer on the performance of BB-MLP model. It is observed that the performance of BB-MLP is improved as the number of hidden neurons increased. However, too many neurons in the hidden layer may cause over-fitting problems, which results in good network learning and data memorization, but lack of ability to generalize. On the other hand, if the number of neurons in the hidden layer is not enough, the network may not be able to learn. Considering Table 1, a BB-MLP model with five neurons in the hidden layer seems to be appropriate for modeling CCI_{repair} and with six neurons for modeling CCI_{oil}, CCI_{fuel} and CCI_{rm}.

These topologies can be more versatile for future applications of repair and maintenance costs prediction.

3.2. Learning rate and momentum term

In order to speed up convergence, an extra term called momentum (α) is used to the weights update (Gupta et al., 2003; Vakil-Baghmisheh, 2002). The learning rate and momentum factors are only used in the learning process, so the criteria used to optimize them are based on the learning error and the iteration number. When the optimal topology of the neural network was found, the learning rate (η) and momentum term (α) was also optimized throughout a trial–error method. For the selected topology, several learning processes were performed with different coefficients, ranged from 0 to 0.99 and 0.1 to 0.99 for learning rate and momentum term, respectively.

Figs. 2–5 show the mean absolute percentage error values versus the learning rate and momentum term. The learning rate and

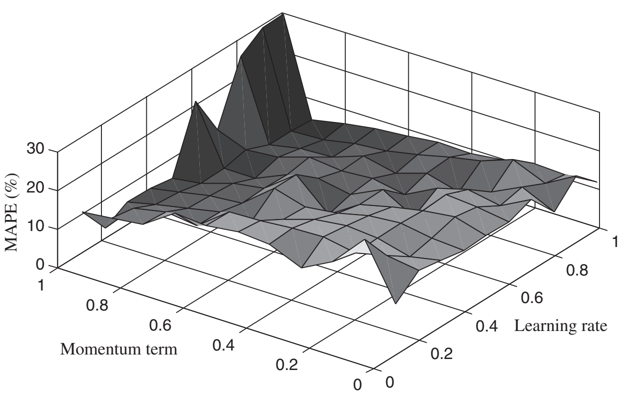

Fig. 2. MAPE (%) profile as a function of learning rate and momentum term for modeling CCI_{repair}.

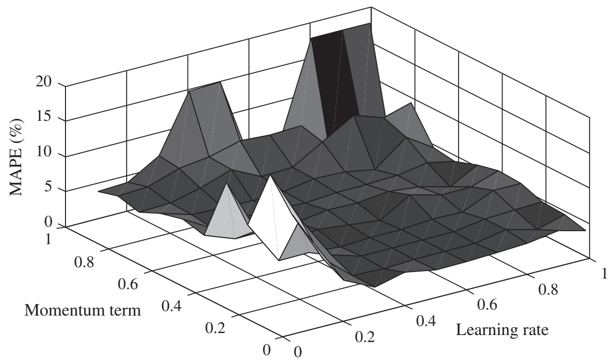

Fig. 3. MAPE (%) profile as a function of learning rate and momentum term for modeling CCI_{oil}.

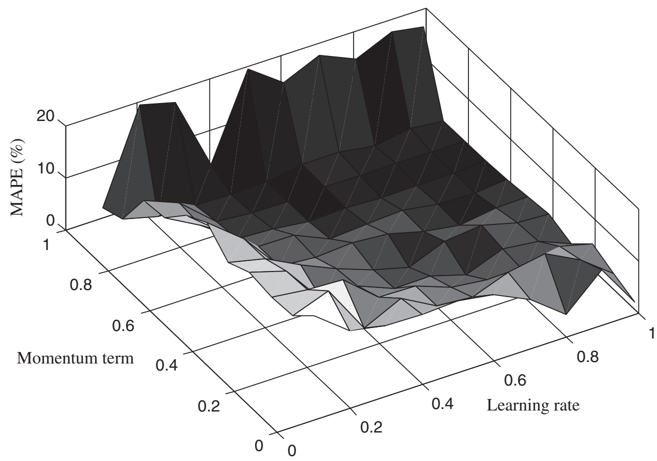

Fig. 4. MAPE (%) profile as a function of learning rate and momentum term for modeling CCI_{fuel}.

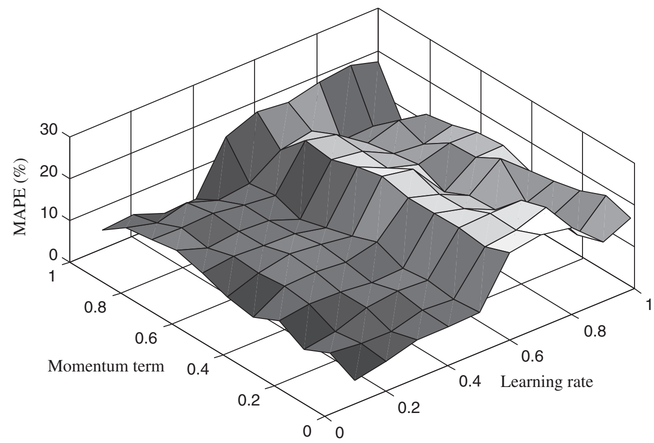

Fig. 5. MAPE (%) profile as a function of learning rate and momentum term for modeling CCI_{rm}.

momentum factors have interactive impacts on network training. This makes parameter tuning a difficult task where momentum term is added. It is observed that the error value is increased and the convergence speed of the learning process is decreased when the momentum term is zero. The results also revealed that the convergence could be faster with a relatively larger learning rate (close to 1). However, with a very high learning rate, the neural network will not converge to its true optimum and the learning process will be instable. It is also evident that, the convergence speed of the learning process was improved through an appropriate choice of parameters η and α. Table 2 shows a relative minimum and maximum learning rate and momentum term and minimum epoch associated with BB-MLP. The optimum factors must reach the minimum error in the lower iteration number.

According to Vakil-Baghmisheh and Pavešic (2001), in order to improve the behavior of MLP during training, and due to simplicity of adjusting process of network parameters, we also used BDLRF algorithm. The results obtained by BDLRF have shown that the best performance of MLP was obtained via a constant momentum term equals to 0.95. Therefore, when the convergence was slowed down, a point was chosen and η was only decreased using Eq. (14). The initial and final values of η were 0.6 and 0.03, respectively. In this study, 1000, 1500, 1500 and 2000 epochs were selected as start points of the BDLRF (n_1) for CCI_{rm}, CCI_{oil}, CCI_{fuel} and CCI_{repair}, respectively. After these points, for every 5 epochs parameter η was calculated using Eq. (14). The epochs were the same in both BB and BDLRF algorithms (Table 2).

3.3. Statistical analysis

3.3.1. Training phase

During training phase the network used the training set. Training was continued until a steady state was reached. The BB and BDLRF algorithms were utilized for model training. Some statistical properties of the sample data used for training process and the prediction values associated with different training algorithms are shown in Table 3. Considering the average values of standard deviation and variance, it can be deduced that the values and the distribution of real and predicted data are analogous. However, the differences of minimum and maximum values are remarkable. This is probably due to the fact that the extreme values were not well represented in the training data set, because these were only one or two points. Accordingly, the neural networks have been learned the training set very well, hence the training phase has been completed.

3.3.2. Test phase

In test phase, we used the selected topology with the previously adjusted weights. The objective of this step was to test the network generalization property and to evaluate the competence of the trained network. Therefore, the network was evaluated by data, outside the training set.

Table 4 shows some statistical properties of the data used in test phase and the corresponding prediction values associated with different training algorithms. It can be seen that the differences of statistical values between the desired and predicted data is less than 0.8% and 0.6% for BB and BDLRF, respectively. While in training phase these values were less than 0.01% for both of training algorithms (Table 3). This fact can be justified since these data are completely new for the MLP. On the other hand, the kurtosis, sum and the average values are similar, hence it can be deduced that both series are similar. The predicted values were very close to the desired values and were evenly distributed throughout the entire range. Although the results of training phase were generally better than the test phase, the latter reveals the capability of neural network to predict the repair and maintenance costs with new data.

From statistical point of view, both desired and predicted test data have been analyzed to determine whether there are statistically significant differences between them. The null hypothesis assumes that statistical parameters of both series are equal. p value was used to check each hypothesis. Its threshold value was 0.05. If p value is greater than the threshold, the null hypothesis is then fulfilled. To check the differences between the data series, different tests were performed and p value was calculated for each case. The results are shown in Table 5. The so called t-test was used to compare the means of both series. It was also assumed that the variance of both samples could be considered equal. The obtained p values were greater than the threshold, hence the null hypothesis cannot be rejected in all cases ($p > 0.99$). The variance was analyzed using the F-test. Here, a normal distribution of samples was assumed. Again, the p values confirm the null hypothesis in all cases ($p > 0.97$). The analysis of medians by Wilcoxon rank-sum test also verified that there is no statistically difference be-

Table 2
Optimum parameters of neural network (BB-MLP).

Parameters of modeling	Parameters of neural network					
	Range of learning rate	Learning rate	Range of momentum term	Momentum term	Epoch	Topology
CCI_{repair}	0.1–0.3	0.2	0.85–0.95	0.95	50,000	1-5-1
CCI_{oil}	0.5–0.7	0.6	0.85–0.95	0.95	10,000	1-6-1
CCI_{fuel}	0.3–0.5	0.4	0.85–0.95	0.95	5000	1-6-1
CCI_{rm}	0.2–0.4	0.3	0.85–0.95	0.95	45,000	1-6-1

Table 3
Statistical variables of desired and predicted values in training phase (MLP).

Parameter of cost		Training algorithm	Statistical values							
			Average	Variance	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
CCI_{repair}	Desired values	BB and BDLRF	49.8932	1593.0	39.9125	0.2082	122.7284	1.7403	0.3502	7184.6
	Predicted values	BB	49.8932	1592.8	39.9095	0.1175	122.1067	1.7393	0.3501	7184.6
		BDLRF	49.8929	1592.8	39.9103	0.1275	122.3605	1.7397	0.3501	7184.6
CCI_{oil}	Desired values	BB and BDLRF	8.1640	35.9620	5.9968	0.0425	19.9007	1.8904	0.3824	1175.6
	Predicted values	BB	8.1640	35.9607	5.9967	0.0521	19.9008	1.8904	0.3823	1175.6
		BDLRF	8.1639	35.9597	5.9966	0.0551	19.8834	1.8903	0.3823	1175.6
CCI_{fuel}	Desired values	BB and BDLRF	9.1214	77.6226	8.8104	0.0514	27.9555	2.3228	0.8508	1313.5
	Predicted values	BB	9.1212	77.6225	8.8104	0.0334	27.6789	2.3224	0.8506	1313.5
		BDLRF	9.1218	77.6088	8.8096	0.0847	27.7326	2.3221	0.8509	1313.5
CCI_{rm}	Desired values	BB and BDLRF	68.3836	3075.2	55.4541	0.1827	170.0639	1.7082	0.3690	9847.2
	Predicted values	BB	68.3833	3075.0	55.4525	0.1733	168.9779	1.7077	0.3689	9847.2
		BDLRF	68.3834	3074.9	55.4520	0.1882	169.1109	1.7077	0.3689	9847.2

Table 4
Statistical variables of desired and predicted values (test phase).

Parameters of cost		Training algorithm	Statistical values							
			Average	Variance	Standard deviation	Minimum	Maximum	Kurtosis	Skewness	Sum
CCI_{repair}	Desired values	BB and BDLRF	50.9539	1430.1	37.9673	0.0887	117.3654	1.7131	0.2107	3668.7
	Predicted values	BB	50.9314	1441.5	37.8164	0.0352	117.6715	1.7139	0.2190	3667.1
		BDLRF	51.0109	1439.1	37.9350	0.0272	117.630	1.7087	0.2145	3672.8
CCI_{oil}	Desired values	BB and BDLRF	7.6400	33.2837	5.7661	0.1672	19.0021	1.8551	0.4264	550.078
	Predicted values	BB	7.6719	33.2483	5.7692	0.1651	19.1077	1.8532	0.4249	552.3781
		BDLRF	7.6379	33.2575	5.7669	0.1521	19.0968	1.8596	0.4302	549.929
CCI_{fuel}	Desired values	BB and BDLRF	9.9962	76.3916	8.7402	0.2624	28.0593	2.1819	0.7894	719.7232
	Predicted values	BB	9.9834	76.2279	8.7309	0.2359	27.8349	2.1714	0.7850	718.8035
		BDLRF	9.9907	76.3238	8.7364	0.2634	27.8317	2.1699	0.7837	719.3335
CCI_{rm}	Desired values	BB and BDLRF	66.1773	2514.3	50.1433	1.2877	170.6872	2.1023	0.4761	4764.8
	Predicted values	BB	66.1565	2530.3	50.3016	1.1273	170.0113	2.1112	0.4840	4763.3
		BDLRF	66.2529	2520.8	50.2079	1.1951	169.9842	2.1080	0.4809	4770.2

Table 5
Statistical comparisons of desired and predicted test data and the corresponding p values.

Parameters of cost	Training algorithm	Analysis types			
		Comparisons of means	Comparisons of variances	Comparisons of medians	Comparisons of distribution
CCI_{repair}	BB	0.9972	0.9733	0.9952	1.000
	BDLRF	0.9928	0.9790	0.9761	1.000
CCI_{oil}	BB	0.9735	0.9964	0.9347	1.000
	BDLRF	0.9983	0.9974	1.000	1.000
CCI_{fuel}	BB	0.9930	0.9928	0.9697	1.000
	BDLRF	0.9970	0.9970	0.9793	1.000
CCI_{rm}	BB	0.9980	0.9789	0.9888	1.000
	BDLRF	0.9928	0.9914	0.9920	1.000

tween medians at 95% confidence level in all cases ($p > 0.93$). Finally, the Kolmogorov–Smirnov test also confirmed the null hypothesis. From statistical point of view, both desired and predicted test data have a similar distribution for both of training algorithms ($p = 1.000$).

Figs. 6–9 show the actual cumulative cost indices versus the predicted ones. It is clear that the regression coefficients of determination between actual and predicted data ($R^2 = 0.999$) are high for the test data sets. Since excellent estimation performances were obtained using the trained network, it demonstrates that the trained network was reliable, accurate and hence could be employed for tractor repair and maintenance costs prediction. These figures reveal that the cumulative cost indices predictions from BB training algorithm were not as good as fit to actual cumulative cost indices in comparison to BDLRF cumulative cost indices prediction. Comparisons of actual versus predicted cumulative cost indices for BB training algorithm resulted in a least squares linear regression lines with slopes equal to BDLRF, while the BDLRF training algorithm resulted in a lines with y-intercepts lower than BB.

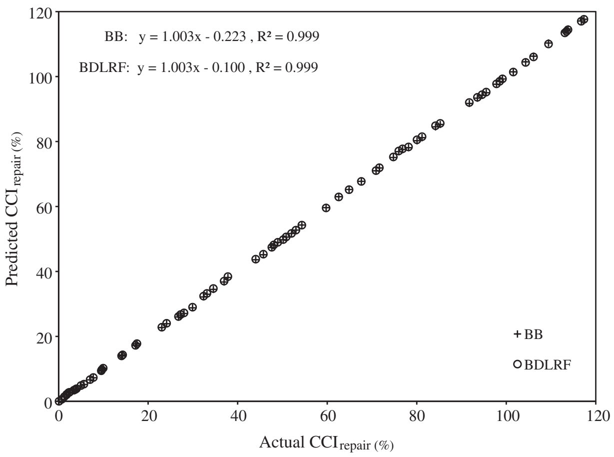

Fig. 6. Predicted values of Artificial Neural Networks versus actual values of CCI_{repair} for BB and BDLRF training algorithms.

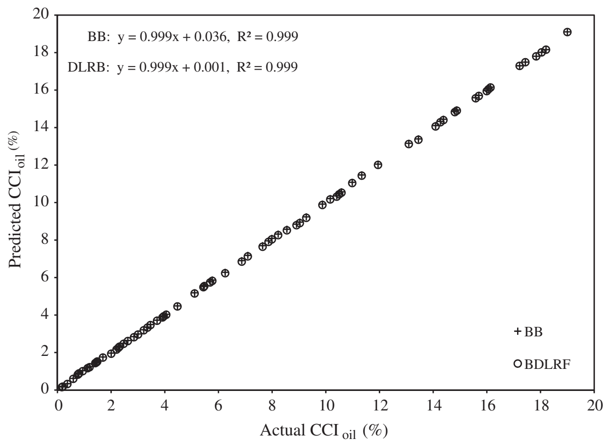

Fig. 7. Predicted values of Artificial Neural Network versus actual values of CCI_{oil} for BB and BDLRF training algorithms.

3.4. Comparison of training algorithms

The modeling of all components of repair and maintenance costs was investigated with some separate networks. For prediction of each component, several networks with different settings and training algorithms were trained. The performances of the two training algorithm are shown in Table 6. For this specific case study, the comparison of results reveals that both algorithms are

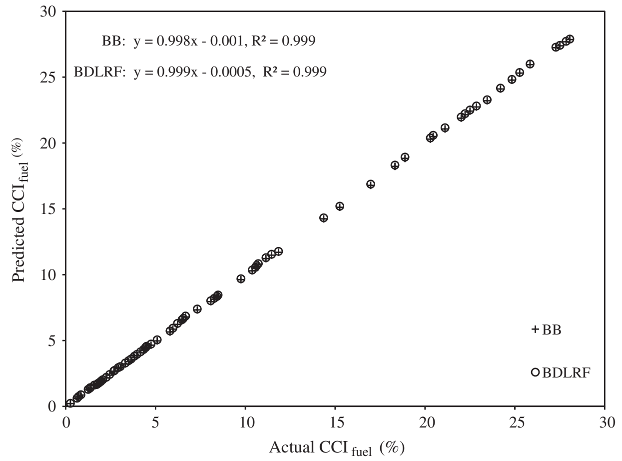

Fig. 8. Predicted values of Artificial Neural Network versus actual values of CCI_{fuel} for BB and BDLRF training algorithms.

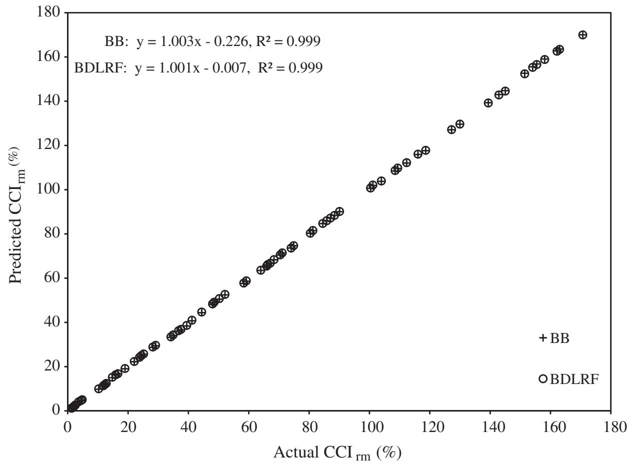

Fig. 9. Predicted values of Artificial Neural Network versus actual values of CCI_{rm} for BB and BDLRF training algorithms.

capable of generating accurate estimates within the preset range. However, it was noticed that BDLRF algorithm had a higher decrease of MAPE, RMSE and TSSE for training phase and also a higher decrease of TSSE in the test phase in comparison to BB algorithm. It was quite clear that the BDLRF training algorithm achieved a much better performance than the BB training algorithm. Bearing all the results obtained by this study in mind, the advantages of the BDLRF training algorithm over BB are: faster convergence, lower training time and also it eases the process of parameter adjusting by decreasing the sensitivity to the parameters' values. The results also conforms the findings of Vakil-Baghmisheh (2002).

3.5. Simultaneous prediction of CCI_{repair}, CCI_{oil}, CCI_{fuel}

Any relationship, linear or nonlinear, can be learned and approximated by an ANN such as a three-layer MLP with sufficiently large number of neurons in the hidden layer. Another remarkable advantage of ANN is its capability of modeling the data of multiple inputs and multiple outputs. ANN has no restriction on the number of outputs. No matter how many outputs and how many inputs a problem has, their relationships can all be learned simultaneously by an ANN with multiple inputs and multiple outputs. In contrast, the conventional regression techniques can only be used to learn the relationship between a single output and one or more inputs but cannot be used to model the data of multiple inputs and multiple outputs.

The feed forward network structure was used for simultaneous prediction of three cost indices as shown in Fig. 10. The input layer had one neuron representing the cumulative hours of usage (CHU), while the output layer consisted of three neurons which corresponded to cumulative repair cost index (CCI_{repair}), cumulative oil cost index (CCI_{oil}) and cumulative fuel cost index (CCI_{fuel}). Based on findings described in Section 3.4, only BDLRF training algorithm was used. Different number of neurons (5–15) in the hidden layer was tested. However, the best results were obtained with ten neurons. The same initial values (0.9) were used for both η and α. The final values of η and α were 0.045. In our study, 10,000 epochs was selected as a point of switching to BDLRF algorithm and after this point, for every 5 epochs α and η was calculated using Eq. (14). The algorithm was terminated at epoch 1,20,000.

The performances associated with MLP network employing the BDLRF training algorithm for prediction of repair and maintenance cost indices are presented in Table 7. The results reveal a very good agreement between the predicted and the desired values of repair and maintenance cost indices ($R^2 = 0.999$). Comparing the results generated using this network with that of generated by separate

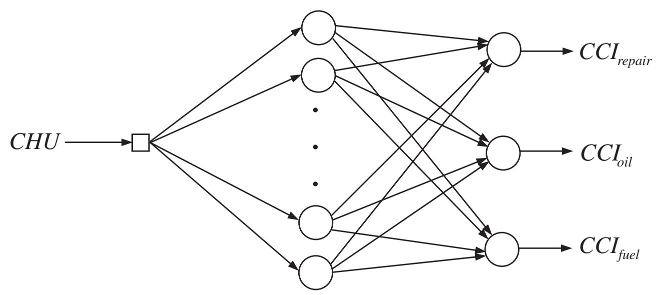

Fig. 10. Multilayer neural network used in the simultaneous prediction of repair and maintenance cost indices.

Table 6
Performances of two training algorithm in prediction of tractor repair and maintenance costs indices.

Parameters of cost	Training algorithm	Performance criterion			
		MAPE (%)	RMSE	TSSE (train phase)	TSSE (test phase)
CCI_{repair}	BB	2.83	0.4609	0.00148	0.00046
	BDLRF	**2.74**	**0.3840**	**0.00095**	**0.00039**
CCI_{oil}	BB	1.83	0.0521	0.00054	0.00040
	BDLRF	**1.68**	**0.0457**	**0.00045**	**0.00028**
CCI_{fuel}	BB	2.63	0.1008	0.00122	0.00056
	BDLRF	**1.78**	**0.0780**	**0.00077**	**0.00029**
CCI_{rm}	BB	1.66	0.4766	0.00066	0.00041
	BDLRF	**1.50**	**0.4259**	**0.00052**	**0.00034**

Table 7
Performances of BDLRF training algorithm for prediction of tractor repair and maintenance costs indices.

Parameters of cost	Performance criterion		
	MAPE (%)	RMSE	R^2
CCI_{repair}	2.29	0.3674	0.999
CCI_{oil}	1.80	0.0415	0.999
CCI_{fuel}	1.59	0.0746	0.999

networks for individual repair and maintenance cost indices (Tables 6 and 7), it can be concluded that both approaches are capable of producing accurate predictions. However, the prediction of repair and maintenance cost indices utilizing single network approach with 1-10-3 topology was more accurate than prediction of separate network approach. Also, these results confirm that the trained network can be used as a global neural model for approximation of tractor repair and maintenance costs.

4. Conclusions

This article focused on the application of ANN to predict tractor repair and maintenance costs. To show the applicability and superiority of the proposed approach, the actual data of tractor repair and maintenance costs from Astan Ghodse Razavi agro-industry (in the north east of Iran) were used. To improve the output, the data were first preprocessed. MLP network was used and applied with the past 18 years tractor repair and maintenance costs as variable inputs. The network trained by both BB and BDLRF learning algorithms. Statistical comparisons of desired and predicted test data were applied to the selected ANN. From statistical analysis, it was found that at 95% confidence level (with p-values greater than 0.9) both actual and predicted test data are similar. The results also revealed that, using BDLRF algorithm yields a better performance than BB algorithm. After testing all possible networks with the test data sets, it has been demonstrated that MLP network with 1-10-3 instruction and BDLRF algorithm had the best output. It is also found that neural network is particularly suitable for learning nonlinear functional relationships and multiple output and input relationships which are not known or cannot be specified.

Because the ANN does not assume any fixed form of dependency in between the output and input values, unlike the regression methods, it seems to be more successful in the application under consideration. It could be said that the neural network provides a practical solution to the problem of estimating repair and maintenance costs in a fast, inexpensive, yet accurate and objective way. It is hoped that the analysis conducted in this article can provide reference for the choice of ANN in such area. Additional research on ANNs is required to make use of these networks more appealing and user-friendly to farm management applications.

Acknowledgements

The authors would like to thank Astan Ghodse Razavi agro-industry in Iran for providing the data and other support to this study.

References

Bowers, W., & Hunt, D. R. (1970). Application of mathematical formula to repair cost data. *Transactions of the ASAE, 13*, 806–809.

Fuls, J. (1999). *The correlation of repair and maintenance costs of agricultural machinery with operating hours management policy and operator skills for South Africa*. <http://www.arc.agric.za> accessed July 2006.

Gupta, M. M., Jin, J., & Homma, N. (2003). *Static and dynamic neural networks: From fundamentals to advanced theory*. Hoboken, New Jersey: John Wiley & Sons, Inc..

Hanke, J. E., & Reitsh, A. G. (1995). *Business forecasting*. Englewood Cliffs, NJ: Prentice-Hall.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. New York: McMillan College Publishing Company.

Jensen, H. R. (1977). Farm management and production economics. In L. R. Martin (Ed.). *A survey of agriculture, economic and literature* (Vol. 1). Minneapolis: Univ. of Minnesota Press.

Makridakis, S., & Wheelwright, S. C. (1989). *Forecasting methods for management*. New York, NY: John Wiley and Sons.

Mitchell, Z. W. (1998). *A statistical analysis of construction equipment repair costs using field data & the cumulative cost model*. PhD Thesis, Faculty of the Virginia Polytechnic Institute and State University.

Morris, J. (1988). Estimation of tractor repair and maintenance costs. *Journal of Agricultural Engineering Research., 41*, 191–200.

NeuroDimensions Inc. (2002). *NeuroSolutions tool for excel*.

Rotz, C. A. (1987). A standard model for repair costs of agricultural machinery. *Applied Engineering in Agriculture., 3*(1), 3–9.

Vakil-Baghmisheh, M. T. (2002). *Farsi character recognition using Artificial Neural Networks*. PhD Thesis, Faculty of Electrical Engineering, University of Ljubljana.

Vakil-Baghmisheh, M. T., & Pavešic, N. (2001). Back-propagation with declining learning rate. *Proceeding of the 10th electrotechnical and computer science conference* (Vol. B, pp. 297–300). Slovenia: Portorož.

Vakil-Baghmisheh, M. T., & Pavešic, N. (2003). A fast simplified fuzzy ARTMAP network. *Neural processing letters, 17*, 273–301.

Zhang, Y. F., & Fuh, J. Y. H. (1998). A neural network approach for early cost estimation of packaging products. *Computers & Industrial Engineering, 34*, 433–450.