HAUL TRUCK RELIABILITY ANALYSIS APPLYING A META-HEURISTIC-BASED ARTIFICIAL NEURAL NETWORK MODEL: A CASE STUDY FROM A BAUXITE MINE IN INDIA

Snehamoy Chatterjee, snehamoy@gmail.com, Assistant Professor, Department of Mining Engineering, National Institute of Technology Rourkela, Rourkela – 769008, Orissa, India

Sukumar Bandopadhyay*, sbandopadhyay@alaska.edu, Professor of Mining Engineering, Department of Mining and Geological Engineering, University of Alaska Fairbanks, Fairbanks, AK, 99775, USA*Corresponding Authors.

ABSTRACT

A reliability analysis of a dumper machine is performed using the neural network model. A metaheuristic algorithm is applied for proper parameter selection for the neural network. A time series approach is applied in the reliability analysis to forecast the dumper's future failure time. The time series analysis is performed using the neural network model, where the numbers of input variables for the model are selected by calculating the entropy value. The case study data were collected from one open-pit bauxite mine in India. The results of this analysis demonstrate that the developed model can forecast quite accurately the dumper's future failure time.

Keywords: Dumper reliability, metaheuristic, neural network, entropy

INTRODUCTION

Systems reliability plays an important role in performance assessment. Machine reliability modeling has been a critical issue for preventive maintenance in the mining industry. The widely applied system reliability models are based on lifetime distribution models, fault tree analysis, and Markov models (Lu *et al*. 2001). These traditional reliability models, however, fail to provide satisfactory results when the reliabilities are of a non-linear nature.

Researchers have shown that reliability can be modeled by considering it as a time series process (Xie 1998; Ho *et al*. 2002), and they have applied linear time-series models for predicting reliability. However, interest in the use of artificial neural networks (ANNs) for time series forecasting rather than traditional time series models is growing because ANNs are capable of handling both linearity and non-linearity (, and Olatunbosun 2008; Xu *et al*., 2003). The applications of ANNs are very limited, though. Liu *et al*. (1995) established that the ANN method significantly improves accuracy in reliability forecasting when compared with traditional methods. Xu *et al*. (2003), who predicted engine reliability by applying multilayer perceptron (MLP) neural networks and radial basis function (RBF) neural networks, demonstrated that the ANN-based model performed better than the autoregressive moving average (ARMA) model. A hybrid neural-fuzzy model proposed by Chang *et al*. (2004) for predicting engine reliability demonstrates that the model achieved more-accurate prediction than the ARMA model and the generalized regression neural network model.

The drawback of the neural network (NN) model is data over-fitting (Bishop, 1998). Due to over-fitting, the developed NN model may perform nicely with the training data (the data that are used for model development); however, the same model may perform poorly with data that are not used for model development. Over-fitting problems are generally encountered because of the poor selection of NN parameters. The parameters that influence the performance of NNs are learning rate, momentum, hidden node size, and number of hidden layers. The challenge for any modeler is to select suitable learning parameters. Generally, a trial-and-error approach is used for selection of these parameters, but this approach does not guarantee the optimum model solution.

In this paper, an automatic parameter-selection algorithm is applied. A metaheuristic algorithm is applied in selecting the NN parameters for the dumper reliability analysis; the genetic algorithm (GA), a popular metaheuristic algorithm, is applied; and pre-established NN architecture with full connection is used.

METHOD

Reliability Estimation Model

To model the reliability of a machine considering reliability/failure probability data are time series data, a NN model is trained to understand the relationship between past historical reliability indices and corresponding target values. The developed NN model, therefore, is applied to predict future reliability values.

If $Y = [y_1, y_2, ..., y_t]$ is the time-varying reliability data of any machine and p is number of lagged variables for the time series model, then the data set for the reliability model can be extracted from Y by

$$T_i = \left\{ y_{(1+i)}, y_{(2+i)}, ..., y_{(k+i)}, ..., y_{(p+i)}, y_{(p+i+1)} \right\}, i \in \{0,1,2,...,t-p-1\}, k < p \quad (1)$$

In general terms, the reliability-forecasting model can be represented as

$$y_{(p+i+1)} = f(y_{(1+i)}, y_{(2+i)}, ..., y_{(p+i)}), i \in \{0,1,2,...,t-p-1\} \quad (2)$$

where $\{y_{(1+i)}, y_{(2+i)}, ..., y_{(p+i)}\}$ is a vector of lagged variables, $y_{(p+i+1)}$ is the observation at time $t = (p+i+1), i \in (0,1,...,t-p-1)$, and p represents the number of past observations related to the future value. The goal of this study is to approximate the function f of Eq. (2) using the NN model.

An important step in reliability analysis, considering time series data, is suitable selection of the value of lag number p. The importance of the selection of lag number p is discussed elsewhere (Chatterjee and Bandopadhyay, 2012). In this study, the entropy value was considered the criterion for selection of the optimum lag number. Entropy is a statistical measure of randomness that can be used to characterize the data. The mean entropy of data set T_i with the dimensions of $(p+1)$ can be computed as (Honarkhah and Caers, 2010)

$$H = \frac{1}{K} \sum_{i=1}^{K} q_i \log(q_i) \quad (3)$$

where K equals the number of possible outcomes of the random variable, and q_i represents the probability mass function. High entropy represents more randomness. As the size of lag grows, the entropy of the data set T_i image increases; however, after a certain lag size, the entropy value reaches a steady state.

Neural Network Algorithm

After selection of the optimal lag size, time series data were extracted from the observed reliability data using Eq. (1). The reliability model was developed using a backpropagation NN algorithm. The three-layer ANN used in this study and the details about NN fundamentals are documented elsewhere (Bishop, 1998; Hagan *et al*., 1995; Haykins, 1999). The input variables for the NN model are $\{y_{(1+i)}, y_{(2+i)}, ..., y_{(p+i)}\}$. The output variable is $y_{(p+i+1)}$. The hidden layer was fixed during the training. The gradient descent with a momentum-learning algorithm was applied for NN training. The algorithm is an iterative algorithm, and the error function, which is a scalar function of the weights, is minimized when the network outputs match the desired output. The error function that will be minimized during training is presented as

$$E(w) \equiv \frac{1}{2} \sum_{i=1}^{c} \left| (y_{(p+i+1)} - \hat{y}_{(p+i+1)}) \right|^2 \quad (4)$$

where $y_{(p+i+1)}$ and $\hat{y}_{(p+i+1)}$ are the targets, and the network output vectors of length, *c*, and *w* represent the weights in the network. The weights are initialized with random values and then changed in a direction that reduces the error:

$$\Delta w(m) = -\eta \frac{\partial E(m)}{\partial w} + \mu \Delta w(m-1) \tag{5}$$

where η is the learning rate and μ is momentum.

The iterative algorithm requires taking a weight vector at iteration *m* and updating it as

$$w(m+1) = w(m) + \Delta w(m) \tag{6}$$

The weights are updated iteratively until the error function reaches its threshold value. Observe from Eq (5) that the two parameters—learning rate (η) and momentum (μ)—play an important role in the weight-updated process. Traditionally, η and μ values are selected by trial and error, a selection process that may produce a model that is far from optimum. The GA was used for selection of parameters for ANN learning.

Neural Network Training Using Metaheuristics

The genetic algorithm (GA) is a metaheuristic algorithm used when searching for the best learning parameters (Holland, 1975; Dengiz *et al*., 1997; Wei *et al*., 2003; Li *et al*., 2007). The GA is initialized with a population of chromosomal solutions, and it updates the solution iteratively by three genetic operators—selection, crossover, and mutation—to get an improved solution. Fitness function is used for evaluating the probability of selection of individual chromosomes for the next iteration. Selected chromosomes (the parent solution) are used for the crossover operation. After crossover, a set of parents produces two children solutions. Mutation changes the binary bit value at randomly selected bit locations, from 0 to 1 or vice versa, based on a user-selected mutation rate. The mutation operation helps the algorithm to avoid trapping at local minima. These three genetic operations are applied iteratively, and the worst solutions are eliminated from the population after iteration to help maintain a constant population size. The iterative process continues until it reaches the predefined iteration number or the selected threshold value.

The chromosomes represent both parameters—learning rate (η) and momentum (μ)—with a binary-coded number. An 8-bit binary code number was used for representing 1 parameter. Therefore, each chromosome consisted of 16 bits: 8 for each learning parameter. The chromosome was uniformly initialized. The number of chromosomes in the population, that is, the population size, totaled 50. The selection based on probability of an individual's fitness is implemented to select the better individuals. The normalized geometric ranking scheme was applied in this study. The crossover operation is performed in iteration to generate an improved solution. Details of the crossover operation for reliability analysis are presented in Chatterjee and Bandopadhyay (2012). Mutation is the genetic operator responsible for maintaining diversity in the population. The mutation probability used was 1/*p*, where *p* is the length of each of the two parts of the chromosomes. The normalized mean squared error (NRMSE) of the validation data set was used as the fitness function for NN parameter selection using a genetic

algorithm. The *k*-fold cross-validation method was adopted. The fitness function used is the average of the NRMSE of the *k*-fold cross-validation data and is presented as

$$\text{Fitness function} = \frac{1}{k}\sum_{i=1}^{k}\sqrt{\frac{\sum_{j=1}^{n}(y_{ij}-\hat{y}_{ij})^2}{\sum_{j=1}^{n}y^2{}_{ij}}} \tag{7}$$

where y_{ij} and $\hat{y}_{ij}$ are actual and predicted reliability values, respectively, and *n* is the number of validation data at the k^{th} fold. After performing GA learning, the chromosomes with the best learning parameters (learning rate and momentum) for the ANN model are selected as the optimum parameters for the problem. The single hidden layer network was used. The optimum number of nodes in hidden layer is selected by changing the hidden nodes number in the NN models and by comparing the performance of different NN models of different numbers of hidden node. The number of hidden node is selected as the best-suited hidden nodes number which provides minimum error value in NN model, and the corresponding learning parameters were the optimum parameters for the NN model.

CASE STUDY

The case study for the proposed method involved failure prediction of a *dumper* in the mining industry, an industry in which dumpers play an important role. The dumper is used for carrying extracted material to its destination for further processing. If a dumper stops, the operation in one part of the mine is affected, should a spare dumper not be available for allocation or replacement. Therefore, mine management gives full attention to avoiding dumper failure. If management can predict the next likely occurrence of failure, suitable maintenance can be planned and performed to avoid such failure.

Historical time-to-failure data of a dumper from an open-cast metal mine in India were collected, altogether constituting 49 failure data. The calculated cumulative time-to-failure for the dumper is presented in Table 1. Out of 49 cumulative failure data, the first 43 data were used for training the algorithm and the remaining 6 were used as test data.

Table 1 - Failure data of load-haul-dump (LHD) machine.

Failure order number (*i*)	Time-to-failure (h_i)	Cumulative time-to-failure (H_i)		Failure order number (*i*)	Time-to-failure (h_i)	Cumulative time-to-failure (H_i)
1	29	29		26	25.75	951.5
2	2.5	31.5		27	2.75	954.25
3	17.5	49		28	25.75	980

Failure order number (i)	Time-to-failure (h_i)	Cumulative time-to-failure (H_i)		Failure order number (i)	Time-to-failure (h_i)	Cumulative time-to-failure (H_i)
4	79	128		29	22	1002
5	64.5	192.5		30	12.5	1014.5
6	65.5	258		31	28.5	1043
7	77.5	335.5		32	144	1187
8	136.5	472		33	140	1327
9	22.25	494.25		34	46	1373
10	7	501.25		35	75.75	1448.75
11	1	502.25		36	44.5	1493.25
12	4.5	506.75		37	2.25	1495.5
13	13	519.75		38	40	1535.5
14	25	544.75		39	20.25	1555.75
15	45.5	590.25		40	24.5	1580.25
16	22.75	613		41	22.75	1603
17	46.75	659.75		42	23	1626
18	43	702.75		43	19.25	1645.25
19	5	707.75		44	123	1768.25
20	16.75	724.5		45	15.5	1783.75
21	7.5	732		46	20.5	1804.25
22	47	779		47	16.5	1820.75
23	122.25	901.25		48	121.5	1942.25
24	21.75	923		49	25.5	1967.75
25	2.75	925.75				

Before modeling the cumulative failure time for the dumper, the lag value *p* was optimally selected by monitoring the mean entropy values. It was observed from the mean entropy results that the optimum lag value *p* is 6. The results were cross-checked by developing different NN models for different *p* values and calculating the error (mean squared error) of the cross-validation data. The manual runs of NN models revealed that the minimum error was observed when *p* value is 6. The training and testing data for NN modeling were extracted from Table 1 data, and using Eq. (1) and optimum *p* value of 6. For neural network training, all input and output data were normalized between 0 and 1. The activation functions used in the hidden layer and the output layer are tan-hyperbolic and pure linear, respectively. The learning parameters of NN (η and μ) were selected using the metaheuristic algorithm presented in Section 2.3. The optimum values of η and μ were 0.092 and 0.28, respectively. The optimum hidden node size was calculated, and the value is 9 (Figure 1).

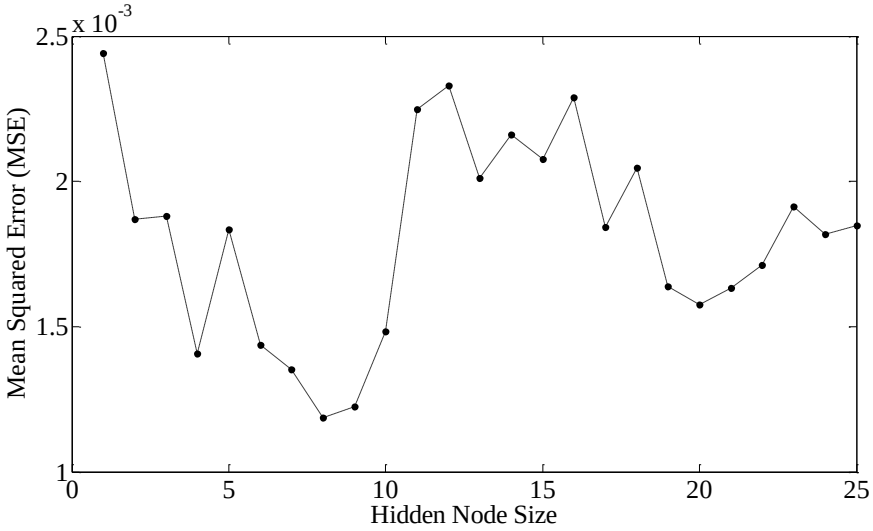

Figure 1 - Error values of 5-fold cross-validation of cumulative dumper failure time data by varying the hidden node number.

The NN model was developed for dumper cumulative-time-to-failure data, using the optimum NN parameters and the optimum hidden node size. The actual values and the predicted values of the training data are presented in Figure 2. Results of Figure 2 show that in some cases the predicted values underestimate the actual values, and in some cases, the predicted values overestimate the actual values. On average, however, the predicted values closely match the actual values.

Generalization of the developed model was performed by comparing the test data actual values with the test data predicted values from the developed model. The error statistics of the actual values and the model-predicted values for the training and test data set are presented in Table 2. The mean squared error values for the training and test data are 4310 and 1067, respectively. The R^2 values for the training and test data are 0.97 and 0.86, respectively. The coefficient of determination (R^2) represents the proportion of actual data variance of a predictor variable that can be explained by the model. High R^2 values for the proposed model of the reliability estimation illustrate that 86% of the cumulative time-to-failure data variances is explained by

the model. The pair sample *t*-statistics were also performed to check the statistical similarities between the actual values and the model-predicted reliability values for both training and test data sets. The *t*-statistics results revealed that there is no significant difference in means for cumulative time-to-failure of training and test data. Figure 3 shows a scatterplot of actual values and predicted values for the test data set. The results show that the scattered data fall close to the bisector line.

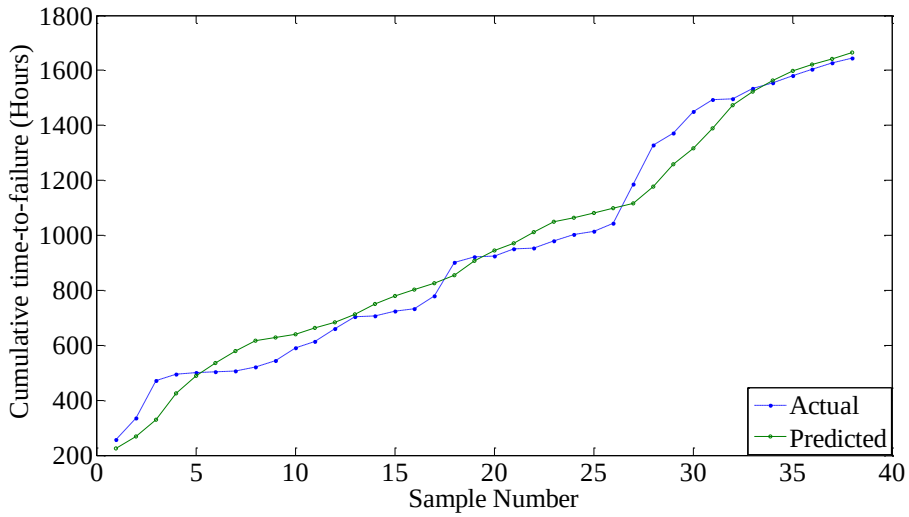

Figure 2 - Actual and predicted cumulative time-to-failure training data from the case study dumper.

To compare our proposed method with other methods, a comparative study was carried out. The comparison was performed by calculating the mean squared errors of the test data set using different methods. Table 3 presents the actual and predicted reliability values of the test data set. Observe from the table that, in terms of minimizing the mean squared error, the proposed method performed better than traditional neural networks, the support vector machine, and autoregressive models.

CONCLUSIONS

An application of the metaheuristic-based neural network model was performed by predicting the reliability of a dumper operated at an open-pit mine in India. The genetic algorithm was applied to search for the best learning parameters of the neural network (NN) model when predicting systems reliability. The input variables for the NN models were selected by maximizing the mean entropy value. The study represented in this paper demonstrates that selected parameters can improve the performance of the model and at the same time significantly reduce computational time by eliminating the trial-and-error exercise for selection of NN parameters. The reliability of the dumper is forecasted indirectly by predicting the cumulative failure time. The comparison with other methods revealed that the proposed method performs significantly better than other methods. The results of this research exhibit the potential of this approach for predicting failure and reliability of any system and its subsystems.

Table 2 - Training and test data error statistics of the proposed model for reliability of the LHD machine.

	Training	Test
Mean error	-1.6973	-14.4604
Mean absolute error	53.897	26.8314
Mean squared error	4310	1067
Error variance	4423	1030
R^2	0.97	0.86

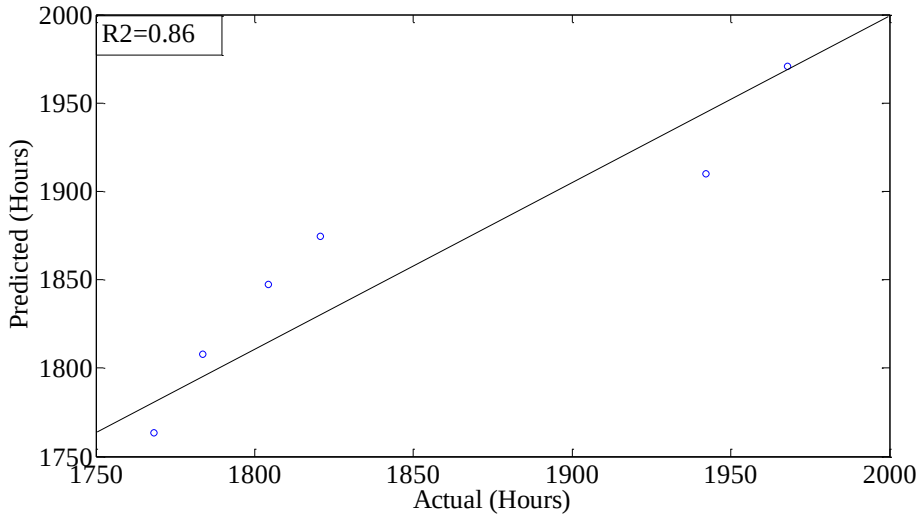

Figure 3 - Scatter plot of the observed vs. predicted value of four attributes using neural regression.

Table 3 - Actual and predicted reliability values of the test data set.

Number	Cumulative	MLP	Support	RBF	AR	Our

	time-to-failure (Actual) (Hour)	(Logistic) (Hour)	vector machine (Hour)	(Gaussian) (Hour)	(Hour)	proposed method (Hour)
44	1768.25	1784.36	1765.62	1763.02	1756.05	1763.60
45	1783.750	1816.97	1811.95	1801.34	1780.87	1807.96
46	1804.25	1822.62	1848.43	1842.78	1847.41	1847.19
47	1820.75	1883.93	1878.01	1888.07	1899.32	1874.51
48	1942.25	1907.90	1917.28	1920.17	1933.68	1909.79
49	1967.75	1945.32	1958.28	1970.96	1965.31	1970.71
NRMSE		1229	1124	1141	1379	1067

REFERENCES

Bishop, M. (1998). Neural Networks for Pattern Recognition. Oxford. Clarendon Press.

Chang, P.T., Lin, K.P., Pai, P.F. (2004). Hybrid learning fuzzy neural models in forecasting engine systems reliability. In Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference. Gold Coast. Australia. pp.2361–2366.

Chatterjee, S., Bandopadhyay, S. (2012). Reliability estimation using a genetic algorithm-based artificial neural network: An application to a load-haul-dump machine. Journal of Expert Systems with Applications 39. pp.10943–10951.

Dengiz, B., Atiparmak, F., Smith, A.E. (1997). Local search genetic algorithm for optimization of highly reliable communications networks. IEEE Transactions on Evolutionary Computation. v.1. n.3. pp. 179–188.

Hagan, M.T., Demuth, H.B., Beale, M. (1995). Neural Network Design, PWS Publishing Company. Boston. MA.

Haykins, S. (1999). Neural Networks: A Comprehensive Foundation. 2nd ed.. Prentice Hall. New Jersey.

Ho, S.L., Xie, M., Goh, T.N. (2002). A comparative study of neural network and Box-Jenkins ARIMA modeling in time series forecasting. Computers and Industrial Engineering. v. 42. pp.371–375.

Holland, J. (1975). Adaptation in Natural and Artificial Systems, The University of Michigan Press, Ann Arbor, MI.

Honarkhah, M., Caers, J. (2010). Stochastic simulation of patterns using distance-based pattern modeling. Mathematical Geosciences. v.42. pp.487–517.

Li, J-P., Balazs, M.E., Parks, G.T. (2007). Engineering design optimization using species-conserving genetic algorithms. Engineering Optimization. v. 39(2). pp.147–161.

Liu, M.C., Kuo, W., Sastri, T. (1995). An exploratory study of a neural network approach for reliability data analysis. Quality and Reliability Engineering International. v. 11. pp. 107–112.

Lolas, S., Olatunbosun, O.A. (2008). Prediction of vehicle reliability performance using artificial neural networks. Expert Systems with Applications. v. 34(4). pp. 2360–2369.

Lu, S., Lu, H., Kloarik, W.J. (2001). Multivariate performance reliability prediction in real-time. Reliability Engineering and System Safety. v. 72. pp. 39–45.

Wei, Z., Ma, D., Hong-Jun, Z., Wang, B.-L., Yun-Tao, C. (2003). An application of multi-population genetic algorithm for optimization of adversaries' tactics and strategies in battlefield simulation. International Conference on Machine Learning and Cybernetics. v. 3(2–5). Nov: pp. 1704–1707.

Xu, K., Xie, M., Tang, L.C. (2003). Application of neural networks in forecasting engine systems reliability. Application of Soft Computation Journal. v. 2 (4). pp. 255–268.