

A comparative study of ANN and Neuro-fuzzy for the prediction of dynamic constant of rockmass

T N SINGH¹, R KANCHAN², A K VERMA² and K SAIGAL²

¹Department of Earth Sciences, Indian Institute of Technology, Powai, Mumbai 400 076, India.
e-mail: tnsingh@iitb.ac.in

²Institute of Technology, Banaras Hindu University, Varanasi 221 005, India.

Physico-mechanical properties of rocks have great significance in all operational parts in mining activities, from exploration to final dispatch of material. Compressional wave velocity (*p*-wave velocity) and anisotropic behaviour of rocks are two such properties which help to understand the rock response under varying stress conditions. They also influence the breakage mechanism of rock. There are different methods to determine the *p*-wave velocity and anisotropy *in situ* and in the laboratory. These methods are cumbersome and time consuming. Fuzzy set theory, Fuzzy logic and Neural Networks techniques seem very well suited for typical geotechnical problems. In conjunction with statistics and conventional mathematical methods, hybrid methods can be developed that may prove to be a step forward in modeling geotechnical problems. Here, we have developed and compared two different models, Neuro-fuzzy systems (combination of fuzzy and artificial neural network systems) and Artificial neural network systems, for the prediction of compressional wave velocity.

1. Introduction

Physico-mechanical properties of rocks are important for the planning and designing of mining and civil excavations. Long term stability can be only achieved when the compressional wave velocity of the rock mass is fully known. The compressional wave velocity depends on the chemical composition, density and hardness of the rock material. It is difficult to determine compressional wave velocity in the field and in the laboratory and this adds to its importance during blasting of the rock mass. When field and laboratory geophysical data are considered together with data of geology and geochemistry, it is possible to obtain much information about the rocks and minerals that are likely to form the deeper layers of the Earth. Hence, it is necessary to study and understand the physical properties of the rocks and minerals of which the layers of the Earth are composed. Laboratory measurements of the elastic constants or the elastic

wave velocities in the rocks are needed for interpretation of seismic velocities (Singh *et al* 1999). A considerable number of measurements of elastic wave velocities in rocks and other materials have been made in laboratories around the world (Balakrishna and Ramana 1968, Inoue and Ohomi 1981, Gaviglio 1989, Diallo *et al* 2003, Singh *et al* 2004).

The relation of the seismic velocities in rocks of the western region of central Asia to density and other physical parameters has been discussed (Yudborovsky and Vilenskaya 1962). The advantages and disadvantages of the static and dynamic methods have also been discussed by Volarovich and Fan (1962), who suggested the use of field seismic observations and the data obtained by static methods in geodynamics. Anisotropy in granite was found in the Sidobre massif (Aveline 1964). Studies both in the laboratory and the field revealed some anisotropy, ascribed to be due to microfissures (Singh and Dubey 2000). These

Keywords. *p*-wave velocity, Neuro-fuzzy; Artificial neural network; compressive strength; density; hardness.

studies revealed the lower velocity in weathered granite as compared to fresh granite. Berezkin and Mikhaylov (1964) found a linear correlation between density and elastic wave velocities in rocks of the central and eastern regions of the Russian platform.

This paper presents an alternative modeling approach to assist the prediction of elastic property of rock (p -wave velocity). The principal constituents of the modeling approach are fuzzy set, fuzzy logic, neural network and data clustering. These are combined into the so-called *hybrid modeling framework* (Neuro-fuzzy). In the present work two different models have been designed using two different systems, ANN and Neuro-fuzzy, and comparison is made between them to know which is the better method. The focus here is not only on how to construct the model but also on how to use this modeling framework to interpret the results and assess the reliability of the model.

2. The dataset

As already stated, the aim of the present investigation is to predict the elastic property of rocks (p -wave velocity) taking physico-mechanical properties and chemical composition of the rocks as inputs. It is, however, uneconomical to obtain all the parameters because this is expensive and time-consuming. On the other hand, all the parameters are not independent and some of them are strongly correlated (Hogstrom 1994). Hence, it is not important to use all the variables as input parameters.

Hence, taking the above discussion and the aim of the investigation under consideration, the following parameters have been taken as input parameters for the network.

Input parameters for the networks

Parameters	Variables
Physico-mechanical properties:	Compressive strength, density, hardness.
Chemical composition:	quartz. (wt. %)

So in all, four parameters have been taken as input parameters for both the Neuro-fuzzy and ANN systems. p -wave velocity has been taken as the output parameter for both systems.

For the ANN system all the input and output parameters were scaled between 0 and 1. This was done to utilize the most sensitive part of neuron and since the output neuron being sigmoid can only

give an output between 0 and 1, the scaling of the output parameter was necessary.

$$\text{Scaled value} = \frac{(\text{max. value} - \text{unscaled value})}{(\text{max. value} - \text{min. value})}.$$

3. Artificial neural network

Neural networks are a branch of “Artificial Intelligence”, besides Case-based Reasoning, Expert Systems, and Genetic Algorithms. Neural networks are able to detect similarities in inputs, even though a particular input may never have been seen previously. This property allows for excellent interpolation capabilities, especially when the input data are noisy (not exact). Neural networks may be used as a direct substitute for autocorrelation, multivariate regression, linear regression, trigonometric and other statistical analysis and techniques.

When a data stream is analyzed using a neural network, it is possible to detect important predictive patterns that were not previously apparent to a non-expert. Thus, the neural network can act as an expert. A particular network can be defined using three fundamental components: transfer function, network architecture and learning law (Simpson 1990). One has to define these components depending upon the problem to be solved.

A network first needs to be trained before interpreting new information. Several different algorithms are available for training of neural networks, but the back-propagation algorithm is the most versatile and robust technique for it provides the most efficient learning procedure for multilayer neural networks. Also, the fact that back-propagation algorithms are especially capable to solve problems of prediction makes them highly popular (Maulenkamp and Grima 1999).

During training of the network, data are processed through the network until they reach the output layer (forward pass). In this layer, the output is compared to the measured values (the “true” output). The difference or error between the two is processed back through the network (backward pass) updating the individual weights of the connections and the biases of the individual neurons. The input and output data are mostly represented as vectors called training pairs. The process as mentioned above is repeated for all the training pairs in the data set, until the network error has converged to a threshold minimum defined by a corresponding cost function, usually the root mean squared error (RMS) or summed squared error (SSE). A simple ANN network is shown in figure 1.

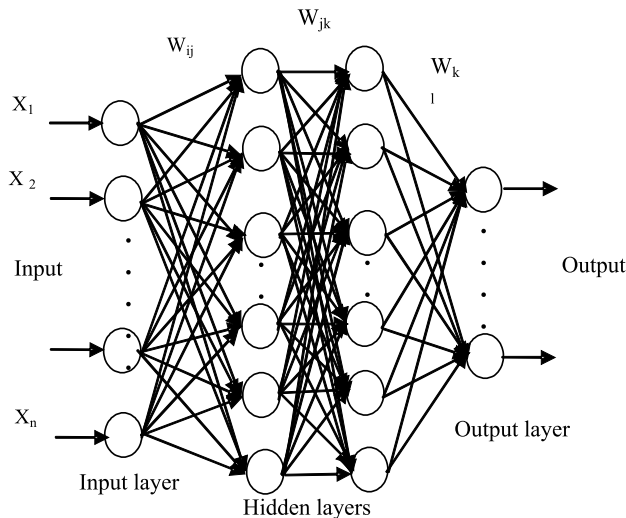


Figure 1. Backpropagation neural network.

4. Fuzzy backpropagation system

Fuzzy systems are currently being used in a wide field of industrial and scientific applications. Since the design and especially the optimization process of fuzzy systems can be very time consuming, it is convenient to have algorithms, which construct and optimize them automatically. One popular approach is to combine fuzzy systems with learning techniques derived from neural networks. Such approaches are usually called Neuro-fuzzy systems (Grima *et al* 2000).

ANN's have been employed in several applications ranging from target recognition to financial forecasting. Their most prominent feature is to learn from examples, and then adapt themselves based on actual solution space training data sets. They are practically powerful in clustering the solution space identifying the important features. Fuzzy systems are based on the idea that all sets are not crisp but some are fuzzy, and these can be modeled in linguistic human terms such as large, small and medium. In a fuzzy system, rules can be formulated that use these linguistic human expressions (Takagi and Sugeno 1985). The combination of ANN's and fuzzy sets offers a powerful method to model human behaviour. The ANN is used to define the clustering in the solution space, which results in creation of fuzzy sets (Jang *et al* 1997).

4.1 Fuzzy system

The classical set is a set with a crisp boundary. The variable either belongs or does not belong to the set. In contrast to a classical crisp set, a fuzzy set is defined as without a crisp boundary, where the transition "belong to a set" is gradual and this transition is characterized by the membership

functions in the range that give fuzzy sets flexibility in modeling (Bezdek 1981).

The membership may be described either in a discrete form as a set of membership values or as a continuous function valid over some range of values of the variable x . To the most popular types of membership functions belong the triangle, trapezoidal, Gaussian or bell functions. We have used here the generalized description of the Gaussian function, given in the form

$$f(x, \sigma, c) = e^{-\frac{(x-c)^2}{2\sigma^2}}. \quad (1)$$

This is the Gaussian function that depends on two parameters σ and c . The parameters for the Gaussian function represents the parameters σ and c listed in order to the vector.

The most popular solution of the fuzzy networks is based on the so-called fuzzy inference system, fuzzy *if-then* rules and fuzzy reasoning. Such a fuzzy inference system implements a nonlinear mapping from input space to output space. This mapping is accomplished by a number of fuzzy *if-then* rules, each of which describes the local behaviour of the mapping, like it is done in radial basis function networks. The antecedent of the rule defines the fuzzy region in the input space, while the consequent specifies the output of the fuzzy region.

There are different solutions of fuzzy inference systems. Two well known fuzzy modeling methods are the Tsukamoto fuzzy model and Takagi–Sugeno–Kang (TSK) model. In the present work, only the TSK model has been considered. A typical fuzzy rule in this model has the form

If x_1 is A_1 and x_2 is $A_2 \dots$

And x_N is A_N then $y = f(x)$ (2)

crisp function in the consequent. The function $y = f(x)$ is a polynomial in the input variables x_1, x_2, \dots, x_N . We will apply here the linear form of this function. The aggregated values of the membership function for the vector x may be assumed either in the form of MIN operator or in a product form. For M fuzzy rules of the equation (2), we have M such membership functions $\mu_1, \mu_2, \dots, \mu_M$. We assume that each antecedent is followed by the consequent of the linear form

$$= p_{io} + \sum_{j=1}^N p_{ij}x_j, \quad (3)$$

where p_{ij} are the adjustable coefficients, for $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$.

Generalized Gaussian membership function have been described by equation (1). The algebraic

product aggregation of the input variables, at the existence of M rules, the Neuro-fuzzy TSK system output signal $y(x)$ upon excitation by the vector x is described by the equation (3).

The adjusted parameters of the system are the nonlinear parameters $(c_j^{(k)}, \sigma_j^{(k)}, b_j^{(k)})$ for $j = 1, 2, \dots, N$ and $k = 1, 2, \dots, M$ of the fuzzifier functions and the linear parameters (weights p_{kj}) of TSK functions. In contrary to the Mamdani fuzzy inference system, the TSK model generates a crisp output value instead of a fuzzy one. The defuzzifier is not necessary.

$$y(x) = \frac{1}{\sum_{r=1}^M [\prod_{J=1}^N \mu_r(x_j)]} \times \sum_{K=1}^M \left(\left[\prod_{J=1}^N \mu_k(x_j) \right] \left(p_{k0} + \sum_{J=1}^N p_{kj} x_j \right) \right) \quad (4)$$

5. Neuro-fuzzy network system

The TSK fuzzy inference systems described by equation (4) can be easily implanted in the form of a so called Neuro-fuzzy network structure. Figure 2 presents the 5-layer structure of a Neuro-fuzzy network, realizing the TSK model of the fuzzy system. It is assumed that the functions $y_i, y_i = f_i(x)$ are linear of the form

$$f_i(x) = p_{i0} + \sum_{j=1}^N p_{ij} x_j. \quad (5)$$

The adaptable parameters of the networks are the variables of the membership functions $c_j^{(k)}, \sigma_j^{(k)}, b_j^{(k)}$ for $j = 1, 2, \dots, N$, $k = 1, 2, \dots, M$ and the coefficients (linear weights) p_{ij} for $i = 1, 2, \dots, M$ and $j = 0, 1, 2, \dots, N$ of the linear Takagi-Sugeno functions.

The network in figure 2 has a multilayer form. The first layer performs the fuzzification according to the membership function $\mu_k(x_j)$, described by equation (1). The second layer aggregates the fuzzified results of the individual scalar functions of every variable and determines the membership function of the whole vector x . This is the product type aggregation. Each node of this layer represents the firing strength of a rule. The third layer calculates the aggregated signal of the fuzzy inference for each inference rule. The output signal of each unit of this layer is the product of the firing strength of the rule and the consequent membership value. The fourth layer determines the output membership function. Layer five calculates only the sum of the signal of the third and the second layers

of the network. The final sixth layer contains only one neuron for output. In the case of multiple outputs, we can add as many output neurons as needed in a fashion similar to the case of one output. The output neuron computes the overall output signal according to the equation (4). Thus, we have constructed the Neuro-fuzzy network that is functionally equivalent to Takagi-Sugeno fuzzy model. Only the first and the third layers are parametric. The parameters of the first layer are associated with the nonlinear functions and the weights p_{ij} of the third layer are linear.

5.1 Hybrid learning algorithm

Learning of the Neuro-fuzzy network, that is adaptation of the parameters of the first $(c_j^{(k)}, \sigma_j^{(k)}, b_j^{(k)})$ and third (p_{ij}) layers of the network, can be done either in supervised or self-organizing mode. For the purpose of approximation, the more efficient and straightforward is the supervised one.

In practical implementation, the hybrid approach has been applied. In this method, we take into account that the network is linear in the parameters p_{ij} , and thus, can identify these linear parameters by a linear least squares method based on singular value decomposition (SVD). At this stage we assume that all nonlinear parameters are fixed. This is the first run of the learning stage. In the second run we fix the linear parameters of the network and apply the gradient steepest descent method for the estimation of the nonlinear parameters of the membership functions.

In hybrid learning each iteration is composed of a forward pass and a backward one. In the forward pass, after the input vector is presented, we calculate the node outputs in the network layers and on the basis of this the linear parameters p_{ij} are adjusted using pseudo inverse based on SVD technique. After the linear parameters are identified we can compute the error for training data pairs. In the backward pass the error signals propagate from the output end toward the input nodes; the gradient vector is calculated and the nonlinear parameters $c_j^{(k)}, \sigma_j^{(k)}, b_j^{(k)}$ updated by steepest descent method. The learning step of the nonlinear parameters update is adjusted using adaptive approach. This process is repeated many times until there is sufficient change of the values of the adapted parameters of the network.

The important advantage of the hybrid algorithm is splitting the learning process into two independent stages: the adaptation of linear weights and adaptation of parameters of the nonlinear membership functions. This algorithm decreases the complexity of the algorithm and at the same time the efficiency of learning increases.

5.2 Rules generation using fuzzy self-organization

One of the most important stages of the Neuro-fuzzy TSK network generation is the establishment of the inference rules (Takagi and Sugeno 1985). Often used is the so-called grid method, in which the rules are defined as the combinations of the membership functions for each input variable. If we split the input variable range into a limited number (say n_i for $i = 1, 2, \dots, N$) of membership functions, the combinations of them lead to many different inference rules. For example for 10 input systems, at 3 membership functions each, the maximum possible number of rules is equal $M = 3^{10} = 59049$. The problem is that these combinations correspond in many cases to the regions of no data, and hence a lot of them may be deleted.

This problem can be solved by using the fuzzy self-organization algorithm. This algorithm splits the data space into a specified number of overlapping clusters. Each cluster may be associated with the specific rule of the center corresponding to the center of the appropriate cluster. In this way all rules correspond to the regions of the space-containing majority of data and the problem of the empty rules can be avoided (Chiu 1994).

The ultimate goal of data clustering is to partition the data into similar subgroups. This is accomplished by employing some similar measures (e.g., the Euclidean distance). In this paper data clustering is used to derive membership functions from measured data, which, in turn, determine the number of If-Then rules in the model (*i.e.*, rules indication).

Several clustering methods have been proposed in the literature (Nauck *et al* 1997). The method employed in this paper is the *subtractive clustering method*.

6. Network architecture for ANN model

A feed-forward network is adopted here as this architecture is reported to be suitable for problems based on pattern identification. Pattern matching is basically an input/output mapping problem. The closer the mapping, the better is the performance of network.

In the small sample, as the data were analyzed here, a cross-validation technique termed leaving-one-out is more appropriate. Out of 55 data sets, 54 were taken to train the network and tested on the remaining one data set. The procedure is repeated 15 times, leaving one observation randomly chosen out at a time. In the experiment the first observation was predicted, using the outcome of an analysis based on the observations 2, 3, ..., 55, and the

55th observation was predicted from the observations 1, 2, 3, ..., 54. This method is advantageous as it uses nearly the entire data sets for training the network. For the ANN networks used for prediction, this cross-validation technique was used. The network has input layer consisting of 4 neurons and output layer consisting of 1 neuron. The number of hidden layers was decided by training and predicting the "training data" and "testing data" by varying the number of hidden layers and neurons in the hidden layer. A suitable configuration has to be chosen for the best performance of the network. Out of the different configurations tested, single hidden layer with 5 hidden neurons produced the best result. Hence, the final configuration chosen for the network is: 4 input neurons, 1 hidden layer with 5 hidden neurons, and 1 output neuron. Suitable numbers of epochs have to be assigned to overcome the problem of overfitting and underfitting of data. In the present paper, to deal with the above mentioned problem, Bayesian regulation (MacKay 1992) was used. Bayesian regulation is an automated regulation, the use of which removes the danger of overfitting, as it never lets the data to suffer from overfitting. This eliminates the guesswork required in determining the optimum number of epochs of the network. In the network the learning rate assigned was 10^{-2} , number of training epochs given were 1500, and error goal was set to be 0.0001.

7. Network architecture for Neuro-fuzzy model

For Neuro-fuzzy model, similar training and testing data sets were used as in ANN model. The membership function of each input was tuned using the hybrid method consisting of backpropagation for the parameters associated with the input membership function and the least square estimation for the parameters associated with the output membership functions (figure 3a-d). The computations of the membership function parameters are facilitated by a gradient vector which provides a measure of how well the FIS system is modeling the input/output data. For a given set of parameters, the numbers of nodes in the training data were found to be 57. The numbers of linear parameters and non-linear parameters were found to be 25 and 40 respectively. The hypothesized initial number of membership functions and the type used for each input were 5 and Gaussian respectively. Now, the hypothesized FIS model is trained to emulate the training data by modifying the membership function parameters according to the chosen error criterion. A suitable configuration has to be chosen for the best performance of the network. Goal for

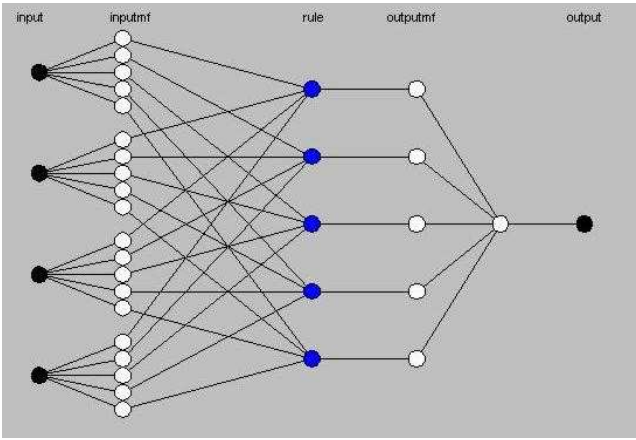


Figure 2. ANFIS structure for the PR model, Neuro-fuzzy, with four input parameters and five rules. Layer-1 represents inputs, layer-2 input membership function, layer-3 rules, layer-4 output membership function, layer-5 weighted sum output and layer-6 output.

the error was set to be 0.01 and number of training epochs was given 50. After the training (with 50 epochs) was complete final configuration for the FIS are

Number of output membership functions = 5

Number of fuzzy rules = 5

Neuro-fuzzy adaptive network for p -wave velocity

1. No. of input = 4
2. No. of membership functions for each input = 5
3. Type of membership functions for each input = Gaussian
4. No. of rules = 5
5. Type of membership functions for each output = Linear
6. No. of output membership function = 5
7. No. of training epochs = 50
8. No. of training data sets = 40
9. No. of testing data sets = 15
10. Error goal = 0.01
11. Error achieved = 5.18597

The clustering method used in this paper is subtractive clustering. The purpose of using clustering method is to identify natural groupings of data from a large set of data to produce a concise representation of a system's behaviour. ANFIS structure for the PR model, Neuro-fuzzy, with four input parameters and five rules is shown in figure 2.

8. Results from ANN system

Observed and predicted values of p -wave velocity along with the percentage error are given in table 1. As the Bayesian regulation (MacKay 1992) is used

in this case, there is no danger of overfitting problems; hence the network trained with 1500 training epochs. The correlation coefficient for the relationship between predicted and observed values is 0.871 (figure 4a). The mean absolute percentage errors (MAPE) for the variables are 2.266% respectively. Figure 4b shows the performance of the network during the training process. The predicted number of parameters on which output is depending is 18, which is more than 6 (input parameters given for this network), which shows that all the parameters were not included.

9. Results from Neuro-fuzzy system

The results are presented in this section to demonstrate the performance of the network. The mean absolute percentage error (MAPE) and coefficient of correlation between the predicted and observed value are taken as the performance measures. The prediction was based on the input data sets discussed above.

Using 40 training data sets and 15 testing data sets and hybrid method consisting of backpropagation for the parameters associated with the input membership function and the least square estimation for the parameters associated with the output membership functions, prediction was made for compressional wave velocity (p -wave velocity) of the rockmasses. Predicted and observed values along with the percentage errors are given in table 2. The correlation coefficients and training performance are shown in figure 5(a-b). Coefficient of correlation between predicted and observed values and MAPE are 0.9901 and 1.692% respectively. From coefficient of correlation and MAPE, it can be said that results obtained were highly encouraging and precise.

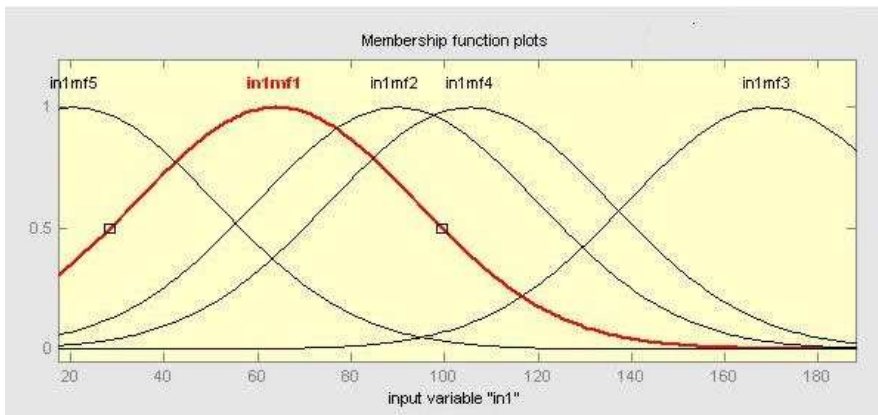
Surface graphs between the inputs and the predicted output are shown in figure 7(a-f). As can be seen, the variation of predicted value (p -wave velocity) with the input parameters is found to be in agreement with the literature. This indicates the excellent identification capability of the ANFIS algorithm.

10. Comparison of the Neuro-fuzzy and ANN models

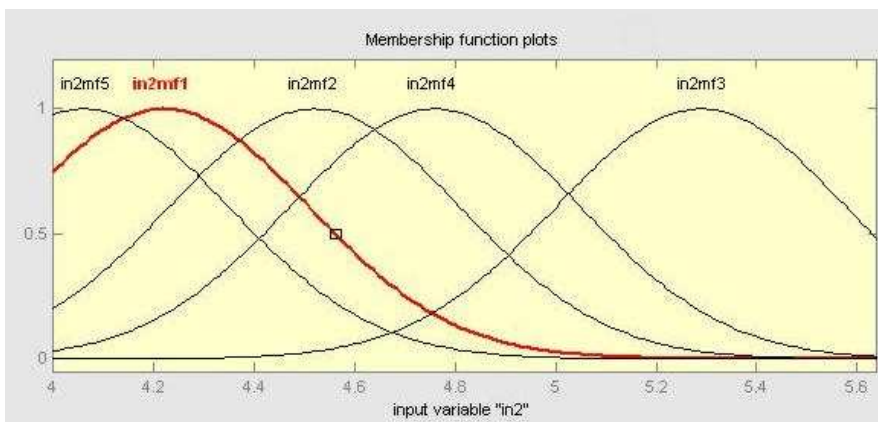
Results from two models are presented in this section to access and compare the degree of prediction accuracy and generalization capabilities of the two networks designed in the present problem. The same training and testing data sets were used to train and test both models to extract more solid conclusions from the comparison results.

Table 1. Observed and predicted values from ANN model along with the percentage error.

Sl. no.	Hardness	Compressive strength (kg/cm ²)	% of quartz	Density (gm/cm ³)	Observed <i>p</i> -wave velocity (m/s)	Predicted <i>p</i> -wave velocity (m/s)	% Error
1	5.69	87.69	2.20	194.87	2217.69	2215	0.12
2	5.68	87.75	2.21	212.14	2230.61	2221.3	0.42
3	5.76	87.84	2.22	213.85	2234.65	2224.4	0.46
4	5.74	87.89	2.22	231.05	2247.58	2230	0.78
5	5.74	87.91	2.27	239.69	2307.35	2246.1	2.66
6	5.85	88.69	2.28	255.26	2324.72	2253.1	3.08
7	5.94	89.65	2.28	256.97	2328.76	2255.4	3.15
8	5.96	89.76	2.38	258.68	2474.15	2278.3	7.92
9	5.98	89.85	2.46	284.68	2741.88	2290.2	16.47
10	4.79	84.98	2.02	115.57	1957.19	1957	0.01
11	4.92	85.56	2.02	127.63	1957.19	1967.7	-0.54
12	4.93	85.69	2.04	125.92	1982.63	1985.9	-0.17
13	4.23	75.94	1.82	68.95	1666.40	1665.8	0.04
14	4.25	76.01	1.82	68.95	1670.84	1670.4	0.024
15	4.26	76.05	1.85	72.45	1747.58	1710.2	2.14

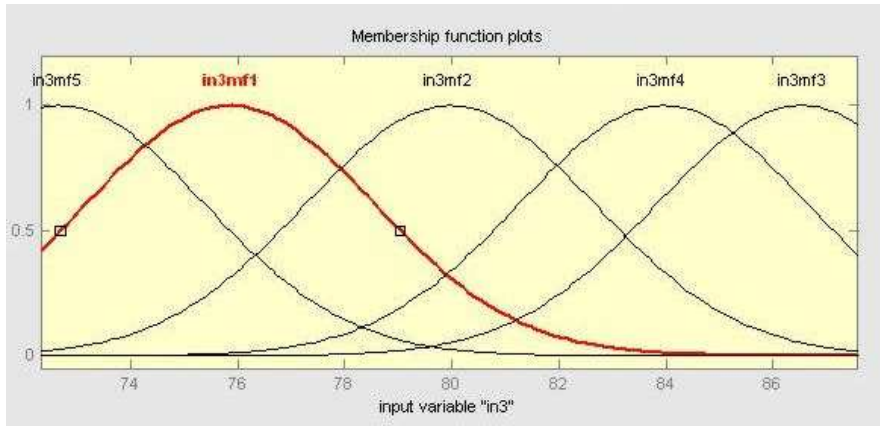


(a)

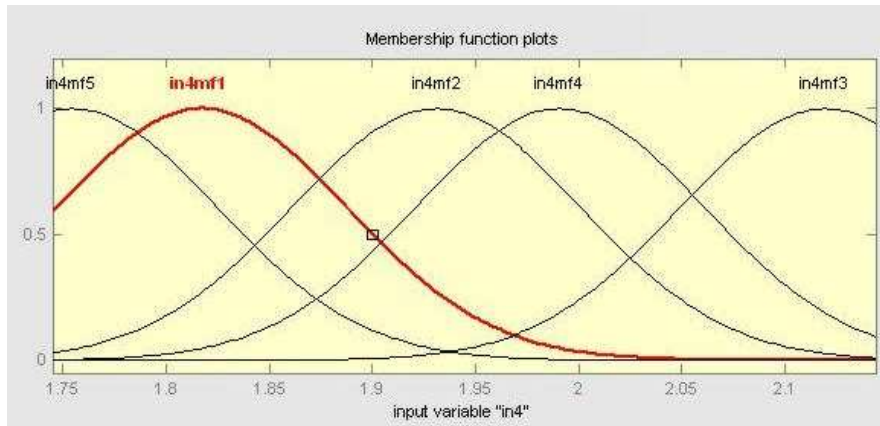


(b)

Figure 3. Membership function plots for (a) input 1 (compressive strength), (b) input 2 (hardness), (c) for input 3 (% quartz) and (d) for input 4 (density).



(c)



(d)

Figure 3. (Continued)

Table 2. Observed and predicted values from Neuro-fuzzy model along with the percentage error.

Sl. no.	Hardness	Compressive strength (kg/cm ²)	% of quartz	Density (gm/cm ³)	Observed <i>p</i> -wave velocity (m/s)	Predicted <i>p</i> -wave velocity (m/s)	% Error
1	5.69	87.69	2.20	194.87	2217.69	2299.52	-3.69
2	5.68	87.75	2.21	212.14	2230.61	2274.98	-1.99
3	5.76	87.84	2.22	213.85	2234.65	2296.42	-2.76
4	5.74	87.89	2.22	231.05	2247.58	2278.56	-1.38
5	5.74	87.91	2.27	239.69	2307.35	2382.13	-3.24
6	5.85	88.69	2.28	255.26	2324.72	2371.42	-2.01
7	5.94	89.65	2.28	256.97	2328.76	2349.99	-0.91
8	5.96	89.76	2.38	258.68	2474.15	2653.55	-7.25
9	5.98	89.85	2.46	284.68	2741.88	2828.55	-3.16
10	4.79	84.98	2.02	115.57	1957.19	1964.29	-0.36
11	4.92	85.56	2.02	127.63	1957.19	1957.14	0.00
12	4.93	85.69	2.04	125.92	1982.63	1989.28	-0.34
13	4.23	75.94	1.82	68.95	1666.40	1671.43	-0.30
14	4.25	76.01	1.82	68.95	1670.84	1671.43	-0.04
15	4.26	76.05	1.85	72.45	1747.58	1710.72	2.11

The predicted values of the *p*-wave velocity from both the networks are plotted against the original experimental values of *p*-wave velocity in the testing dataset (figures 4(a) and 5(a)). Accuracy of

prediction in each case is estimated by drawing the 1:1 line, which represents 100% accuracy, 95% confidence ellipse is also drawn which has its major axis superimposed on the best-fit line. 95%

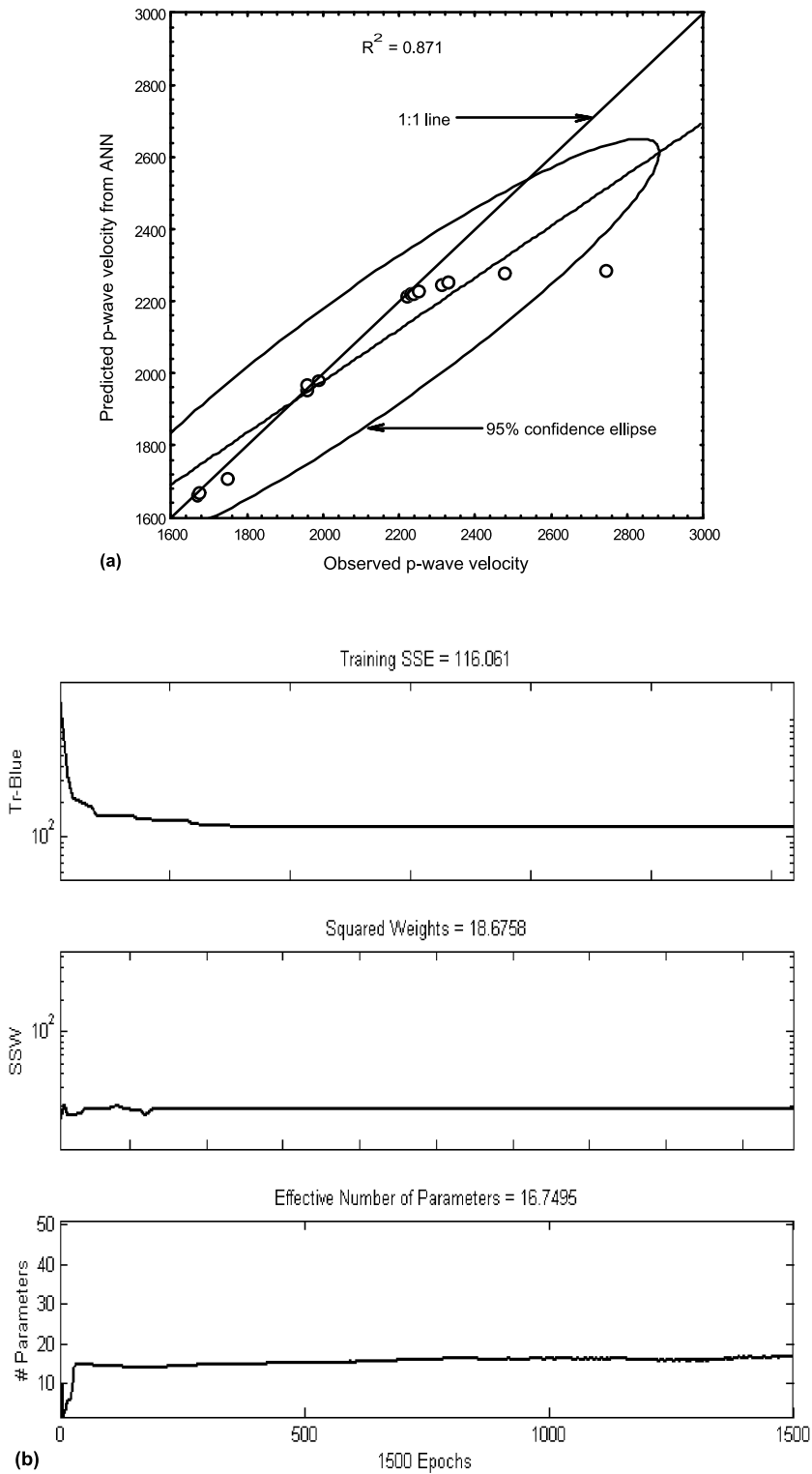
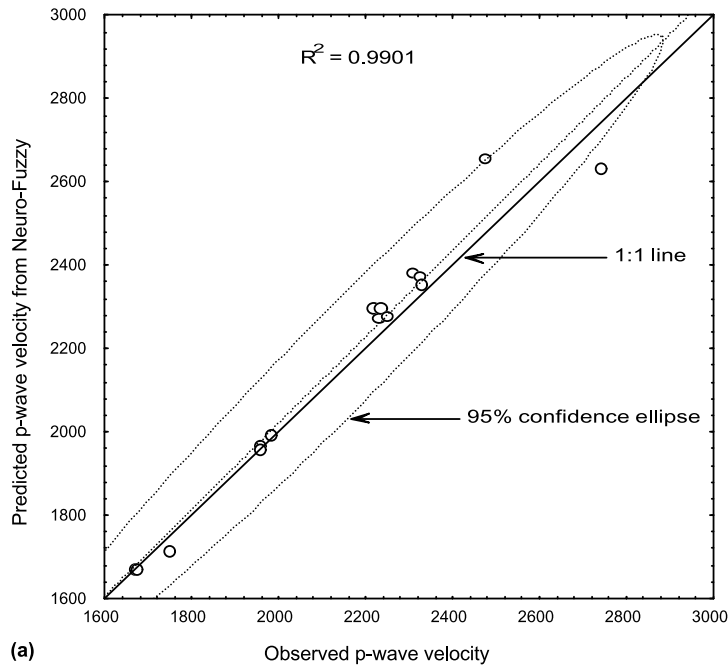


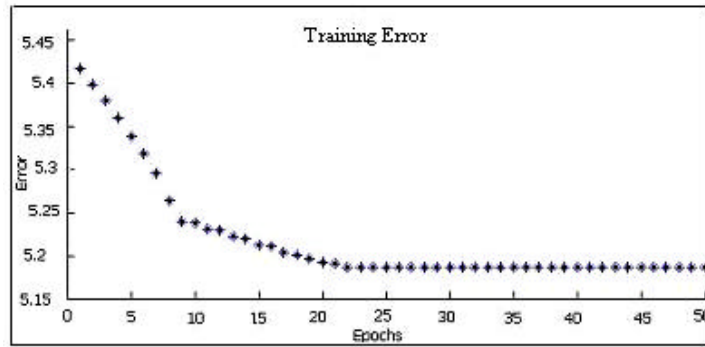
Figure 4. (a) Correlation between predicted and observed value of *p*-wave velocity using ANN, (b) Performance graph of the neural network used for prediction.

confidence represents the probability that the values will fall inside the area marked by the ellipse, thus, it is a very good measure of distribution of points around the best fit line. Hence, smaller the

minor axis and the angle between the its major axis and the 1:1 line better the prediction accuracy. A visual inspection reveals that confidence ellipse for ANN model is having its minor axis longer and its



(a)



(b)

Figure 5. (a) Correlation between predicted and observed *p*-wave velocity using Neuro-fuzzy, (b) Performance graph of the Neuro-fuzzy model.

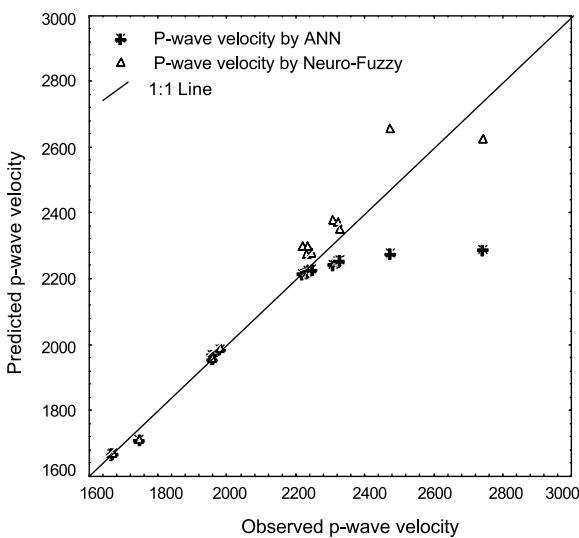


Figure 6. Comparison graph between predictions from two networks against observed values.

major axis is involving the bigger angle with 1:1 line, in comparison to the confidence ellipse of Neuro-fuzzy model. Visual inspection of comparison graph (figure 6) reveals that ANN model shows excellent prediction accuracy for lower values of *p*-wave velocity but is unable to maintain its accuracy for higher values, hence losing its generalization capabilities, whereas the Neuro-fuzzy model maintains its excellent prediction accuracy throughout the range of *p*-wave velocity, hence showing consistency and high a degree of generalization capability. From the above discussion it can be concluded that Neuro-fuzzy model shows better prediction accuracy and better generalization capability in comparison to the ANN model.

Observations made from comparing the results are backed by the fact that results from ANN are largely dependent on architecture of the network,

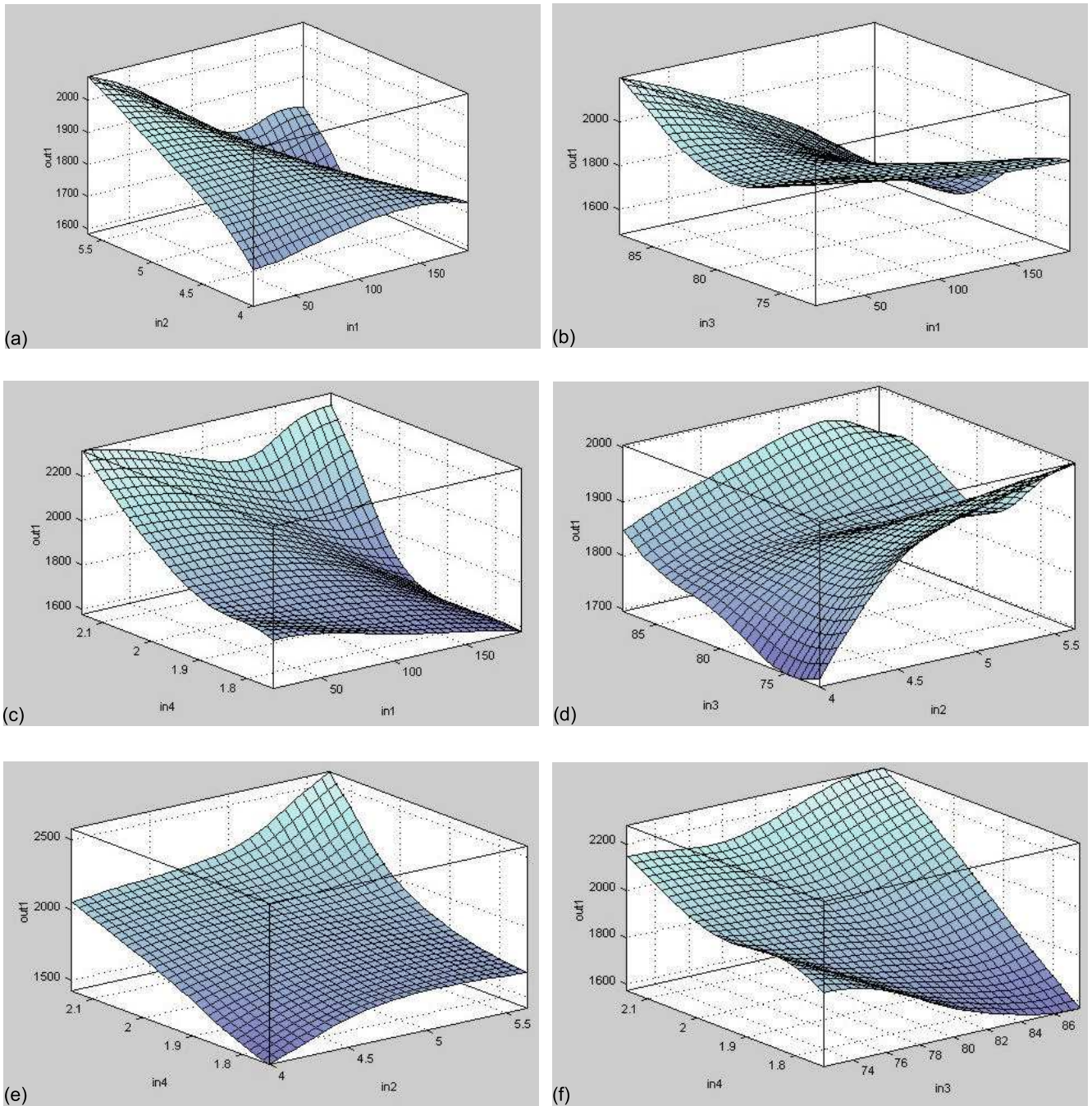


Figure 7. Surface graph showing relationship of (a) input 1 and input 2 with predicted output, (b) input 3 and input 1 with predicted output, (c) input 4 and input 1 with predicted output, (d) input 3 and input 2 with predicted output, (e) input 4 and input 2 with predicted output, and (f) input 4 and input 3 with predicted output, input 1 is compressive strength of rock, input 2 is hardness of rock, input 3 is percentage of quartz in rock, input 4 is density of rock and output is predicted p -wave velocity by Neuro-fuzzy model.

which is very hard to select as it is a complex and time-consuming task. Another limitation that ANN has is its inadequate ability to deal with fuzzy and nonlinear data, whereas Neuro-fuzzy is largely free from both of those limitations. Furthermore, computationally the Neuro-fuzzy model is more easy and efficient than the ANN model.

11. Conclusions

The Neuro-fuzzy method presented in this paper shows a good potential to model complex, nonlinear and multivariate problems. It also shows that the Neuro-fuzzy method is better than the ANN method though the results obtained from

ANN are also satisfactory. Considering the complexity of the relationship between the input and the output, results obtained are very accurate and encouraging. The lower MAPE obtained by the Neuro-fuzzy method suggests its good generalization capability. Neuro-fuzzy modeling is an emerging computational tool that combines fuzzy logic and artificial neural network methods. Perhaps the most interesting feature of this approach is that we can cope scientifically with subjectivity and uncertainty in the engineering process, rather than ignoring them.

References

- Alvarez Grima M, Bruines P A and Verhoef P N W 2000 Modeling tunnel boring machine performance by Neuro-fuzzy methods; *J. Tunnelling and Underground Space Technology* **15** 259–269.
- Aveline M 1964 Experimental results on the relation between micro-fissuration and speed of propagation of ultrasounds in the granites of Sidbore; *J. Sci. Terre.* **9** 439–448.
- Balakrishna S and Ramana Y V 1968 Integrated studies of the elastic properties of some Indian rocks; *Geophysical monograph. AGU* **12** 489–500.
- Berezkin V M and Mikhaylov I N 1964 On the correlational relationship between the density of rocks and their velocities of elastic wave propagation for the central and eastern regions of Russian platform, *Geofiz; Razvedka* **16** 83–91.
- Bezdek J C 1981 Pattern recognition with fuzzy objective function algorithms; New York: Plenum Press.
- Chiu S 1994 Fuzzy model identification based on cluster estimation; *J. Intelligent and fuzzy systems* **3** 267–278.
- Diallo M S, Prasad M and Appel E 2003 Comparison between experimental results and theoretical predictions for p -wave velocity and attenuation at ultrasonic frequency; *Wave motion* **37** 1–16.
- Gaviglio P 1989 Longitudinal wave propagation in a limestone: the relationship between velocity and density; *Rock Mech. and Rock Eng.* **22** 299–306.
- Hogstrom K 1994 A study on strength parameter for aggregate from south western Sweden rocks; Res Rep. (Chamer Univ. of Technology, Goteborg, Sweden).
- Inoue M and Ohomi M 1981 Relation between uniaxial compressive strength and elastic wave velocity of soft rock; *Int. Conf. Weak Rock Tokyo* 9–13.
- Jang J S R, Suni C T and Mizutani E 1997 Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence; New York: Prentice Hall.
- MacKay D J C 1992 Bayesian interpolation; *Neural Comput.* **4** 415–447.
- Maulenkamp F and Grima M A 1999 Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness, *J. Rock Mech. Min. Sci.* **36** 29–39.
- Nauck D, Klawonn F and Kruse R 1997 Foundations of Neuro-fuzzy Systems; New York: John Wiley and Sons.
- Singh T N, Monjezi M and Kumar D 1999 Static and dynamic properties of rocks at sub-zero temperature, *Ind. Min. Engg. J.* 27–31.
- Singh T N and Dubey R K 2000 A study of transmission velocity of primary wave (p -wave) in coal measure sandstone, *Ind. J. Sci. Ind. Res.* **59** 482–486.
- Singh T N, Kanchan R, Saigal K and Verma A K 2004 Prediction of p -wave velocity and anisotropic properties of rock using artificial neural networks technique, *J. Sci. and Ind. Res.* **63** 32–38.
- Simpson P K 1990 Artificial neural system-foundation, paradigm, application and implementation; New York: Pergamon Press.
- Takagi T and Sugeno M 1985 Fuzzy identification of systems and its applications to modeling and control; *J. Transactions on systems, man, and cybernetics*, SMC-**15** 116–132.
- Volarovich M P and Vey-tsin Fan 1962 Investigation of the elastic properties of rocks by static and dynamic methods at high hydrostatic pressure; *Akad Nauk SSSR Ins Fiz Zemli Tr.* **23** 19–24.
- Yudborovsky I and Vilenskaya S M 1962 Some results of study of the elastic properties of rocks of western central Asia; *Akad Nauk Turkmen SSR Ser Fiz-Tekh Khim Geol Nauk.* **3** 26–31.