# On the use of artificial neural networks as quality control tool

Zeghal, M.; Khogali, W. E. I.



**Publisher’s version / Version de l'éditeur:**

*CANCAM 2007 Proceedings, pp. 594-595, 2007-06-03*

# On the use of artificial neural networks as quality control tool

**NRCC-49224**

Zeghal, M.; Khogali, W.E.I.

A version of this document is published in / Une version de ce document se trouve dans:
CANCAM 2007, Toronto, Ontario, June 3-7, 2007, pp. 594-595

# On the Use of Artificial Neural Networks as Quality Control Tool

*M. Zeghal*[1], *W.E.I. Khogali*[2]

[1] National Research Council Canada, Morched.Zeghal@nrc.ca
[2] National Research Council Canada, Walaa.Khogali@nrc.ca

**1. Abstract**

This paper presents the use of artificial neural networks (ANN) technique as a potential quality control tool for detecting erroneous data in databases. Inconsistent behaviour observed while constructing an ANN application for a resilient modulus database pointed out to the possible existence of incorrect entries in the laboratory generated data set. Subsequent removal of suspected faulty data improved prediction capabilities of the developed network and enabled the determination of the optimum number of nodes to be used in the hidden layer.

**2. Introduction**

The use of mechanistic properties to characterize road construction materials has been the focus in recent years to effectively utilize these materials. The resilient modulus ($M_r$) of unbound materials, commonly used in base and subbase layers, represents one of the mechanistic inputs that current pavement design procedures require. The $M_r$ parameter, which is determined using laboratory repeated loading tests, is defined as the ratio of the axial deviator stress ($\sigma_d$) to the axial resilient (or recoverable) strain ($\varepsilon_r$):

$$M_r = \frac{\sigma_d}{\varepsilon_r} \qquad (1)$$

Determination of $M_r$ in the laboratory is elaborate and requires capital investment and special training. Currently, few road jurisdictions in Canada have the required testing capabilities to determine the resilient behaviour of their unbound materials. Accordingly, the Transportation Association of Canada (TAC) has recently recommended the creation of a national $M_r$ database that includes test data obtained from various regions. This is intended to reduce the amount of testing required by each jurisdiction and also provide other jurisdictions, which have no testing capabilities at the present time, with an access to the database. However, assembling data generated by various labs with different equipment and personnel training give rise to the possibility of having erroneous and sometimes conflicting results. Consequently, a mechanism (tool) to detect such flawed data should be made available to ensure the integrity of the developed database. The focus of this paper is to present an ANN scheme that can be used as a quality control tool to serve this objective.

**3. Development of ANN**

Artificial Neural networks are analytical techniques that simulate the way human brain cells function to solve problems [1]. These techniques have the capability of recognizing and capturing pattern(s) that define the relationship between input and output contained in a data set. An ANN that has learned the pattern(s) can later be used to predict new conditions for which the results (output) are not known.

An ANN is made up of three or more layers. The first layer contains the input parameters while the last layer contains the output (solution) parameters. One or more layers known as hidden layers are usually placed between the input and output layers. The development of a network scheme involves determination of the number of inputs, outputs, and nodes in one or more hidden layers. The input layer size is generally predetermined based on the number of parameters known to affect the targeted output(s). However, the number of hidden layers as well as their nodes is usually determined by trial and error. Determination of the number of hidden layers and their nodes involves training and testing of the built network against test set examples with known input and output.

In the present work, a data set made up of 55 $M_r$ entries was used to build the required ANN. This data set was generated under different conditions of density (89–98%), moisture content (3–7%), deviator stress (30–85 kPa), and percentage fines passing the 75-μm sieve (2–18%), which represents factors that are known to affect the resilient modulus of unbound materials [2], [3]. Accordingly, the inputs used for the network were those four factors while the outputs were the corresponding $M_r$ values. Prior to initiating the exercise of building the ANN network, the $M_r$ data set was split into three subsets; a main set (O), made up of 45 entries, used for building and training the network and two smaller subsets (A & B), made up of 5 entries each, used for verification of the adequacy of the built ANN.

By trial and error, it was found that using more than one hidden layer did not improve the prediction capability of the developed network, leaving the number of nodes in the single hidden layer as the only unknown parameter to be determined. Using set 'O', seven networks, each containing a single hidden layer, were created with number of nodes in the hidden layer corresponding to 8, 10, 11, 12, 13, 14 and 15. Thereafter, data from the two subsets A and B was utilized to perform predictions using the seven developed networks. The criterion used to gauge the adequacy of the built networks was the percentage "Absolute value of the Relative Error" ( $|$ARE$|$ ):

$$|ARE| = \frac{\left|\left(x_{predicted} - x_{actual}\right)\right|}{\left|x_{actual}\right|}100\% \quad \text{.................................} \quad (2)$$

Results of the above exercise, displayed in Figure 1, reveals that data from set 'A' produced predictions with high ARE value (≥ 30% for the majority of nodes) with no observable minimum. On the other hand, data from set 'B' shows that the network containing 12 nodes in the hidden layer produced an acceptable level of predictions with an ARE value of 7%.

The inconsistent behaviour of the built networks noted above suggests that some of the data contained in the primary set 'O' used to create the networks may be erroneous. This prompted visual examination of the data entries contained in set 'O' and subsequent removal of some of the entries that are suspected to be incorrect. The modified set 'O' was then used again to build new seven networks at the previously selected number of nodes (8 to 15). Results obtained from these runs are displayed in Figure 2, which now show that the new built networks produce consistent prediction trends for data from both 'A' and 'B' sets. Furthermore, the new network built with 12 nodes, referred to hereafter as $ANN_{12m}$, is identified as the one that produces an optimum level of predictions with an ARE value ranging from 2% to 5% (or an average of 3.5%).

The $ANN_{12m}$ was retained as the final network that could be used as a quality control tool for the national $M_r$ database described previously.

## 4. Summary and Conclusions

This paper presented an analytical investigation aimed at using the artificial neural networks technique to detect the existence of faulty or erroneous data contained in resilient modulus databases. The study involved the development and verification of a specific ANN scheme as the required quality control tool. In addition, findings from the study revealed that an ANN scheme with a single hidden layer containing 12 nodes is sufficient to produce a network with good prediction capabilities.

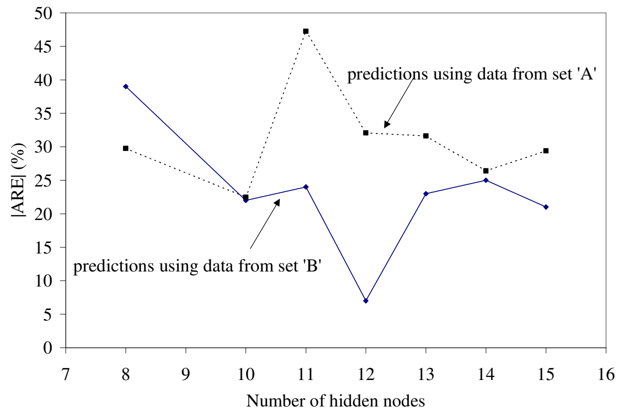


Figure 1. Prediction results for networks built using original set 'O'

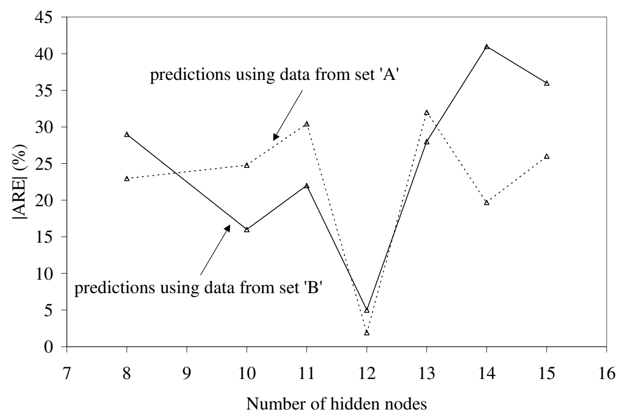


Figure 2. Prediction results for networks built using modified set 'O'

## References

[1] I.A. Basheer and M. Hajmeer, *Artificial Neural Networks: Fundamentals, Computing, Design, and Application*, Journal of Microbiological Methods, No. 43, pp. 3–31 (2000).

[2] A.R. Dawson, N.H. Thom and J.L. Paute, *Mechanical Characteristics of Unbound Granular Materials as a Function of Condition*, Flexible Pavements, Proc. European Symposium Euroflex. 199, A. G. Correia, editor, Balkema, Rotterdam, The Netherlands, 35-44, (1996).

[3] W.E.I. Khogali and ElHussein H. Mohamed, *Novel Approach for Characterization of Unbound Materials*, In Transportation Research Record No. 1874, pp. 38 – 46, Journal of the Transportation Research Board, (2004).