Shin, H. S., Pande, G. N., Kim, C. Y., Bae, G. J., Hong, S. W., Neural network based approach for dissemination of field measurement information.
International Symposium on the Fusion Technology of Geosystem Engineering, Rock Engineering and Geophysical Exploration
Seoul, Korea, Nov. 18-19, 2003

Neural network based approach for dissemination of field measurement information

Hyu-Soung Shin[1], Gyan N. Pande[2], Chang-Yong Kim[1], Gyu-Jin Bae[1], Sung-Wan Hong[1]
[1]Korea Institute of Construction Technology (KICT), Ilsan, Korea; [2]University of Wales Swansea, Swansea, UK

Abstract: This paper presents a neural network based approach to disseminating information relating to experimental and field observations in engineering. Although the methodology is generic and can be applied to many areas of engineering science, attention is focussed here solely on geotechnical engineering applications. Field data relating to the settlement of foundations presented by Burland and Burbidge (1985) which led to their well known equation for calculation of settlement, now included in most text books, is re-visited. A part of the data, chosen randomly, is used to train an Artificial Neural Network (ANN), which relates foundation settlement to various causes as identified by the authors. Predictions are made for situations for which data were not used in training. These indicate sufficient accuracy when compared to the original field data. Accuracy of predictions is further improved when all the data are included in the training set. The finally trained ANN is shown to represent these data more accurately than the Burland and Burbidge equation.

Based on the above heuristic example, an ANN is presented as an alternative to developing equations and design rules in geotechnical engineering practice. Significant advantages are shown to arise by using this methodology. Ease of updating the ANN, as and when additional data becomes available, being the most important one. Loss of transparency, however, seems to be the main disadvantage.

1. Introduction

Advances in engineering practice, in general, take place along three parallel strands. For the solution of a real problem, one may try to (a) develop an analytical or computational model, (b) carry out experiments on physical models or (c) make field observations. Real engineering problems are enormously complex and researchers use a variety of combinations of the above approaches to develop a deeper understanding of the problem at hand. The research complimented by engineering experience is then embedded in the form of simple equations, which are included in a code of practice or are issued as 'guide lines' for practising engineers. National codes are developed to take into account varying construction or design practices. Such codes are of great significance in the assessment of the safety of the existing infrastructure and, in addition, to the economy of development of new or planned infrastructure.

There are two main drawbacks in following the strands (b) and (c) i.e. the route of experiments on physical models and field observations, as mentioned above. They are both expensive and time consuming. Furthermore, they suffer from the disadvantage that the results of such research, which are in general in the form of an equation, cannot be easily updated as and when further research results become available, nationally or internationally. Furthermore, to an experienced engineer, it would appear that many research projects are more or less replays of previous research carried out some 15 or 20 years ago, albeit with some changes in materials and methods. This provides at least anecdotal evidence that most research results obtained from approaches (b) and (c) cannot be enhanced through collection of further data.

This paper describes a ANN based approach of making predictions from the results of field observations (strand (c)). This is particularly relevant to geotechnical engineering where the variability of materials and loads make mathematical and computational models sometimes less viable. Indeed, 'observational methods' have been included in some codes of practice as a method of design. The ANN approach has the advantage that ANNs can be updated (or retrained) easily as and when new field data become available. This avoids the need for a specialist researcher to re-analyse 'old' and 'new' data and propose a new equation.

Section 2 explains the background and terminology of ANNs in simple terms. Section 3 presents the data of Burland and Burbidge (BB) for settlement of foundations on sand together with the equations proposed by the authors. Methodology of training and evaluation of the performance of a trained ANN are discussed in Section 4.

Sections 5 & 6 give conclusions and the authors' vision in the application of this methodology in engineering and research practice.

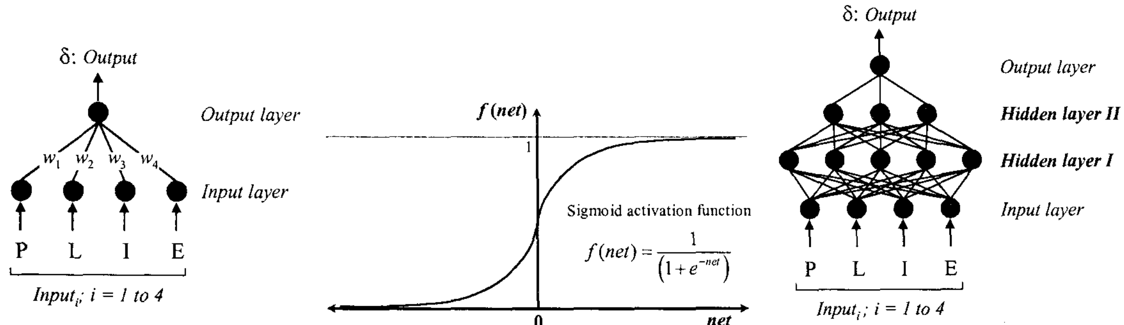
2. Artificial neural networks

ANNs attempt to simulate the working of a human brain in a very crude way. Information is passed on from one 'neuron' to all other neurons connected to it. The human brain observes the 'causes' and subsequent 'effects' of the causes and 'learns' from it. Faced with similar causes and effects for a specific problem, humans become intelligent and are then able to make predictions or anticipate the outcome (effects) due to a set of causes.

The topic of 'computational' intelligence has been researched in many disciples including electronic engineering, computer science and engineering mathematics for the past two decades or so. A number of textbooks are now available (Rumelhart and McClelland, 1986; Pao, 1989) and the computer science syllabus at most universities includes a module on this topic. New applications are being continuously discovered and reported in the literature (Toll, 1996). However, practical applications have been lacking specially in rock and geotechnical engineering perhaps due to the multidisciplinary nature of this emerging methodology. Here, a brief description of the mathematics behind ANNs is presented for the benefit of engineers to enable them to appreciate the potential of the methodology in engineering design.

Architecture of an ANN

The architecture of an ANN consists of a number of nodes, each node representing a 'cause' or 'effect'. For example let us consider the equation of deflection of a simply supported beam (δ) of uniform cross section subjected to a point load (P) at the centre of the span (L). Here, the span, load, the moment of inertia (I) and elastic modulus (E) are the 'causes' and the deflection is the 'effect'. Thus, the ANN representing the problem will have two layers; an input layer having four nodes representing each of the causes and an output layer consisting of only one node representing the effect. The architecture of the ANN will then be as shown in Figure 1(a).



(a) A simple architecture of neural network predictor

(b) Sigmoid activation function

(c) Extended architecture of neural network predictor

Figure 1. Architecture of neural network predictor for a simply supported beam.

The contribution of a particular input such as the moment of inertia, I, is represented by the value of a corresponding weight which is computed during the training process. Then the summation of the contributions of inputs is either activated or deactivated by a 'sigmoid' activation function (some other activation functions are also sometimes used). Figure 1(b) shows a plot of the sigmoid activation function, in which the output varies from 0 to 1. The data used for training are normalised so as to fall within this range through a simple transformation process. The output from the function becomes the 'effect' resulting from the network, i.e. deflection of the beam at the centre of the span, δ in this example, corresponding to the inputs (causes) given. Completion of training means that all the weight values assigned to all the connections have been fixed. The weight values are assumed optimum to minimising discrepancy between target given and output predicted by the network. For the simple case used for illustration, δ, is given by:

$$\delta = output = f(net) = \frac{1}{(1 + e^{-net})}; \quad net = \sum_i Input_i \, w_i + \theta \qquad (1)$$

where input nodes i range 1 to 4 in this example. θ is a threshold value and f is the sigmoid activation function. w_i is a weight coefficient assigned for the connection between the i^{th} input node and the output node.

To explain the philosophy of using ANNs, one has to pretend that the relationship between the deflection of the beam and the causes is unknown and perhaps too complex to be derived in a mathematical form. However, it is assumed that we have lots of observational data. The ANN can then be perceived as a 'pattern searching' algorithm attempting to find the unknown relationship. This, in other words, means finding the best values of w_i and θ to 'smooth' fit the available data. This process is termed 'training'. One can perceive training an ANN as a process of fitting a hyper surface in a multi-dimensional space corresponding to the number of input and output parameters. The architecture shown in Figure 1(a) is, however, too simplistic. The actual architecture of neural networks is much more complex, even for simple problems like that of deflection of a beam. Here, one or more hidden layers having a number of nodes will be added to represent the complex relationship between the input and the output (see Figure 1(c)). There is a standard nomenclature for referring to the architecture of an ANN. For Figure 1(c), it is denoted as NN (4-5-3-1). The first and the last numbers refer to the number of nodes in the input and output layers respectively. The intermediate numbers indicate the number of nodes in each of the hidden layers. The basic Equation (1) is extended in parallel to represent the more complex architecture shown in Figure 1(c) and the general expression for the extended equations can be found in textbooks (Rumelhart and McClelland, 1986; Pao 1989).

At this stage, perhaps the reader would like to know as to how the architecture of the ANN is fixed or decided? It can be done by trial and error, although there are 'auto-structuring' algorithms (Ash, 1989; Shin, 2001), which progressively add 'hidden layers' and nodes until a specified accuracy of fit of data is achieved. Engineers familiar with the finite element method can imagine this as an automatic mesh refinement technique. If the architecture of the ANN is not optimised or fine enough, the predictions will not be accurate. This, indeed, is true for any numerical technique.

Extent and richness of data

There are a number of issues, which an engineer should understand before attempting to use the ANN techniques. Firstly, a large set of data is required to be able to train an ANN. If the data are not sufficient, training will be poor leading to poor predictions. There are a number of ways to assess if the data set is adequate. The simplest is to use a fraction of the data (60 – 80%), chosen randomly, for training and make predictions for the rest. If the predictions are satisfactory, the data may be considered as adequate to define the pattern.

Furthermore, the data set should be rich or diverse, covering the full range of values of variables. Even thousands of data but with in a narrow range of variables will not be sufficient to train an ANN adequately. The main reason is that ANNs are, like humans, good in 'interpolation' but poor in extrapolation.

3. Field observation data on settlement of foundations

The BB data (Burland & Burbidge, 1985) are available in the form of an Excel Sheet (see acknowledgement). The 128 records are available in this study. Based on these data, the authors proposed the following equation for the settlement of structures δ (mm) on sand and gravel.

For normally consolidated soil;

$$\delta = qB^{0.7} \times \frac{1.71}{N^{1.4}} \tag{2}$$

For overconsolidated soil;

$$\begin{aligned} \delta &= (q - \tfrac{2}{3}\sigma'_{v0})B^{0.7} \times \frac{1.71}{N^{1.4}} \quad ; \; if\, q > \sigma'_{v0}, \\ \delta &= qB^{0.7}\frac{1.71}{3N^{1.4}} \quad ; \; if\, q < \sigma'_{v0} \end{aligned} \tag{3}$$

where B (m) is the breadth of the foundation, q (kPa) is the gross pressure applied at foundation level, σ'_{v0}, (kPa) is the maximum previous effective overburden pressure and N is the average SPT blow count over the depth of influence. The settlement, δ determined by either equation (2) or (3) above should be multiplied by the following shape factor f_s:

$$f_s = \left(\frac{1.25L/B}{L/B + 0.25} \right)^2 \tag{4}$$

where L (m) is the length of the foundation. And if the thickness of the sand stratum below foundation level, H, is less than the depth of influence, z_I, which is a constant depending on the foundation width B and can be determined from an empirical diagram given in (Burland and Burbidge, 1985; Craig, 1992). δ should also be multiplied by another factor, f_l:

$$f_l = \frac{H_s}{z_l}\left(2 - \frac{H_s}{z_l}\right) \tag{5}$$

where H_s (m) is the thickness of the sand or gravel layer (Figure 2). This Figure also illustrates the parameters involved in the equations above. In equation 5, D (m) is depth of a foundation, H_w (m) is depth of water table beneath ground level.

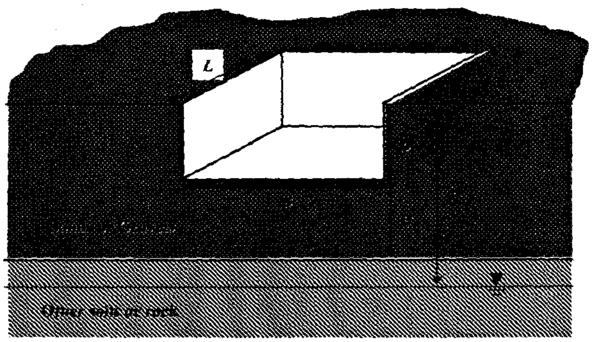


Figure 2. Definition of the parameters for a soil foundation.

4. The BB data in the form of an ANN

About 80% of the data was randomly chosen for training an ANN whose architecture is shown in Figure 3. It has 8 nodes in the input layer, one node in the output layer and two hidden layers. Each hidden layer has 8 and 3 nodes, respectively. It is noted here that unlike the BB equation, which involves 5 input parameters, the ANN consists of 8 nodes. The reason for this is that we did not filter out any variables, which may have negligible influence on the settlement, as was the case in the BB study. If there are any parameters in the current database, which have no effect or negligible effect on the output, these are automatically taken care of by the ANN. In any case, it is advisable to leave them so that if fresh data emerge indicating some dependence, the ANN can be re-trained without altering the architecture. Once the ANN was trained, the data used in training were fed in to verify that the corresponding settlements were obtained. It is, of course, not surprising that the prediction was 'good'.

Figure 4 shows the settlements of various cases chosen for training. These data have been arranged in ascending order of settlements. The field data as well as settlement predicted by the BB equation are also plotted on this graph. It is seen that the Coefficient of Correlation (CC) for the ANN is higher than that for the BB equation. It should, however, be noted that the BB equation makes use of the complete data set whilst the ANN uses only about 80% of the data. About 20% of the data, which was not used for training the ANN, is now utilised to test its prediction capability. Figure 5 shows the prediction of settlement for these cases and compares it with field data as well as the value of settlement from the BB equation. It is seen that the ANN trained on the basis of about 80% of the data makes a fairly good prediction. The CC for the ANN is higher than that for BB equation. In the next step, the ANN is re-trained with about 20% of the data not used earlier. Figure 6 shows the prediction of all cases and compares it with field data as well as that obtained from the BB equation. A considerably higher value of CC for the ANN is observed when compared with the BB equation.

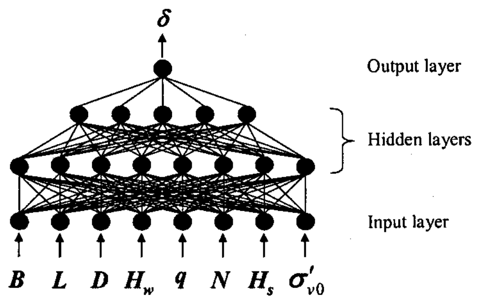


Figure 3. Architecture of the ANN for prediction of ground settlements.

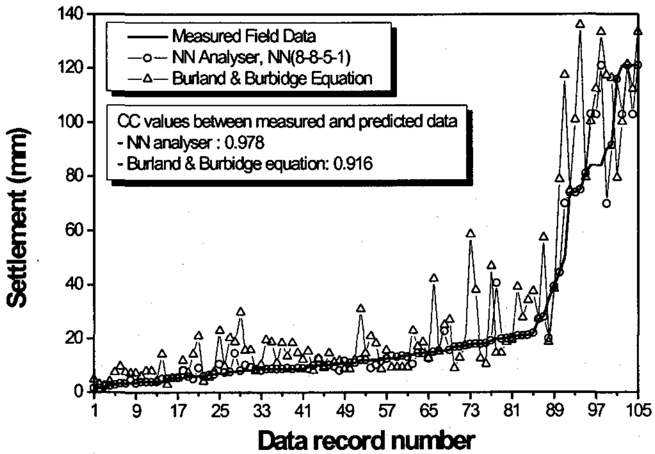


Figure 4. Comparison between results from the BB equation and the trained ANN predictor for data used in training.

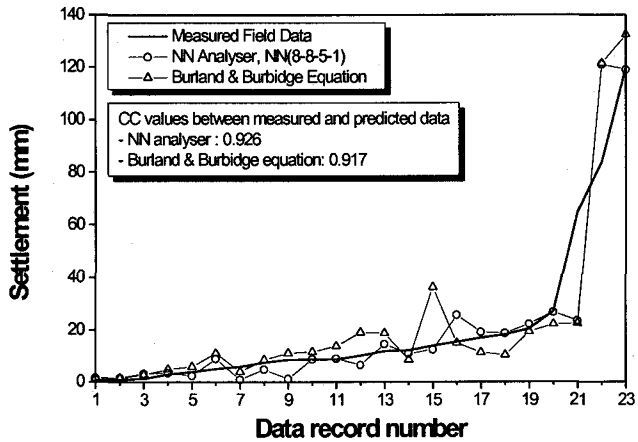


Figure 5. Comparison between results from the BB equation and the trained ANN predictor for data not used in training.

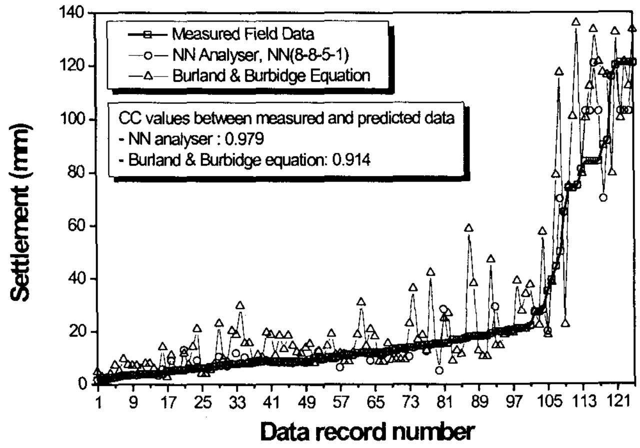


Figure 6. Comparison between results from the BB equation and the re-trained NN predictor for data used in training.

5. Comments on ANN based approach to engineering design

The above section, through an example, has shown that an ANN can be trained through field observation data and can be used in a manner similar to an equation, the BB equation in this particular case. The approach to develop simple equations is also used in engineering research where physical model experiments are conducted, either at 1g or in a centrifuge. So obviously an ANN can also be trained from the result of an experimental research programme. The question, therefore, is what are the advantages and disadvantages in adopting this methodology?

Advantages

Utilisation of future data:

The foremost advantage of this methodology is the possibility of re-training the ANN should further data become available in the future. For example, if any agency for that matter, continued to collect data on settlement at any future date, the BB equation would have to be revisited, a multivariate regression analysis would have to be carried out involving much expense and time. On the other hand, the ANN representing the BB equation, such as the one reported here, can be re-trained and used in a predictive mode with little effort. In another situation, some data may be considered inappropriate or inaccurate at a later date. They can be deleted from the database, the ANN re-trained and again used in a predictive mode. One can also easily judge the influence of including or excluding a certain set of data.

Model test results:

Enormous advantages accrue if the ANN based approach is applied to the results of physical model tests. These tests are expensive and time consuming. Due to insufficient funding, it is rarely possible to study the influence of all parameters in a single project. The conclusions of such studies, therefore, apply over a very limited range and are poor value for money spent. The ANN approach presents the possibility of enhancing the value of model test research, by adding to the database, the results of further research carried out by third parties nationally or internationally.

Complex problems:

If the number of variables in a problem were large, it becomes extremely difficult to arrive at an optimum form of an equation to fit the field or experimental data. In fact, one has to first choose a mathematical function and then determine its constants for a 'least square' or similar fit. A number of equations may have to be proposed, as is the case with the BB equation. The ANN approach does not pose such problems.

Global co-operation and databases:

A flow of information relating to field observation as well as experimental programs will become available to a large engineering and research community. The proposed methodology will encourage all concerned to pool their data and utilise it for the benefit of all. This will also encourage integration of codes of practice of member countries and felicitate good science and engineering.

Transferring experience:

This methodology in future may lead to the possibility of a 'science fiction' type of computer for engineering practice, which will eventually know more than any human being. They would allow the distilling of the experience of engineers for future generations.

The above advantages are, however, not gained without causing certain drawbacks.

Disadvantages

Loss of transparency:

The ANN based approach is not currently transparent and will perhaps never become so transparent as an equation. However, a dual approach of using a trained ANN in addition to an equation can be followed until engineers learn these techniques. After all, the parameters in the equation have to be entered in a calculator/computer to obtain the answer. The response from a trained ANN is almost instantaneous and therefore they can be routinely used in engineering practice.

Erroneous results:

The ANNs can get confused if conflicting data are input for re-training. Code updating committees will have to agree to the acceptable databases, which can then be used to re-train the ANN. All this means is that the trained ANN representing a set of accepted data will have to be 'secured'. This drawback can, however, be turned into an advantage as consulting companies could link their own databases into the ANN based equation.

Generalised training of an ANN

During training, the outputs from an ANN are adapted to approximate the target values corresponding to the inputs in the training data. This adaptation is the main concept of training of an ANN. However, the more important purpose of using an ANN is to make it more general, i.e., to have the outputs of the ANN approximate target values for inputs that are not in the training set as shown in Figure 6. Such a generalisation of an ANN may be more important than the training itself, because a trained NN should be able to produce output acceptable within its population, otherwise output from the ANN has no meaning (Shin, 2001).

Required conditions for generalising an ANN

Generalisation requires prior knowledge, as pointed out by Wolpert (1996). For any practical application, it should be decided that inputs are relevant. A restricted class of input-output cases that contains an adequate approximation to a target function should be known, because it is impossible to fit a function without any proper data points. Thus, although it is not sufficient, the following three conditions are generally required for good generalisation of an ANN (Shin, 2001):

- The first necessary condition is that input variables adopted in an ANN contain sufficient information as 'causes' to corresponding outputs, so that there exists a mathematical function relating correct outputs to inputs with the desired degree of accuracy. There is no ANN capable of learning a nonexistent function.
- The second condition is that a function to be learnt (that relates inputs to correct outputs) should be, in some sense, smooth. In other words, a small change in the inputs should, most of the time, produce a small change in the outputs. For continuous inputs and targets, smoothness of the function implies continuity and restrictions on the first derivative over most of the input space. Some ANNs can learn discontinuities as long as the function consists of a finite number of continuous pieces. Highly discrete functions such as those produced by pseudo-random number generators etc. cannot be generalised by ANNs. Therefore, in the case of a highly nonlinear function, a nonlinear transformation of the input space can increase the smoothness of the function and improve generalisation
- The third condition for good generalisation is that the training cases should be a sufficiently large and representative subset of the set of all cases to be generalised. The importance of this condition is related to the fact that there are, briefly speaking, two different types of generalisation: interpolation and extrapolation. Interpolation applies to cases that are more or less surrounded by nearby training cases; everything else is extrapolation. In particular, training cases that are outside the range of the training data require extrapolation. Training cases inside large 'holes' in the training data may also effectively require extrapolation. Interpolation can often be done reliably, but extrapolation is notoriously unreliable. Hence, it is important to have sufficient training data to avoid the need for extrapolation.

Thus, for an input-output function that is smooth, if there is a test case that is close to some training cases, the correct output for the test case will be close to the correct outputs for those training cases. If adequate cases for training set are prepared, every case in the population will be close to the sufficient number of training cases. Hence, under these conditions and with proper training, an ANN will be able to be generalised reliably within the population. Further researches for providing guidelines to achieve a good generalisation of an ANN is now in progress in authors' research group, which will be reported shortly with visible evidences.

6. Conclusions

The application of ANNs in representing equations in design codes has been presented. It has been shown that the results of field observations can be expressed in the form of a trained ANN which can represent the data better than an equation. The results of physical model tests can also be represented in the form of a trained ANN. This approach has many advantages including the possibility of re-training the ANN as and when additional data become available. Results of field observation and experimental programmes conducted by various organisations, nationally and internationally can be better co-ordinated and used for development of codes of practice. This involves much less effort as compared to constantly revising design equations. The major disadvantage is loss of transparency.

Acknowledgement

The authors are grateful to Dr. Paola Provenzano of the Department of Civil Engineering, University of Rome Tor Vergata, Rome for providing field observation data of Burland & Burbidge in an electronic form.

References

Ash, T., 1989, Dynamic node generation in back propagation networks. Connection Science, Vol 1, No. 4, pp. 365-375.

Burland, J. B. & Burbidge, M. C., 1985, Settlement of foundation on sand and gravel. Proceedings of the Institution of Civil Engineers - Part 1-Design and Construction, Vol 78 (Dec). pp. 1325-1381.

Craig, R. F., 1992, Soil Mechanics, 2nd edn. Chapman & Hall Publishing Company, UK.

Pao, Y. H., 1989. Adaptive Pattern Recognition and Neural Networks. Addison-Wesley Publishing Company, USA.

Rumelhart, D. E. & McClelland, J. L., 1986, Parallel Distributed Processing, Volume I & II. Cambridge: MIT Press.

Shin, H.S., 2001. Neural network based constitutive models for finite element analysis. Ph.D thesis, Department of Civil Engineering, University of Wales Swansea, UK, C/Ph/250/01.

Toll, D. G., 1996, Artificial Intelligence Applications in Geotechnical Engineering, The Electronic Journal of Geotechnical Engineering. Premiere Volume, http://www.ejge.com/1996/Ppr9608/Ppr9608.htm.

Wolpert, D. H., 1996, The existence of a priori distinctions between learning algorithms, Neural Computation, Vol. 8, pp. 1391-1420.