Prediction of ground settlement due to tunneling using artifical neural networks

Dhony Kurniawan Hidayat

2006

Dhony, K. H. (2006). Prediction of ground settlement due to tunneling using artifical neural networks. Master’ s thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/12217

https://doi.org/10.32657/10356/12217

PREDICTION OF GROUND SETTLEMENT DUE TO TUNNELING USING ARTIFICIAL NEURAL NETWORKS

DHONY KURNIAWAN HIDAYAT

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING
NANYANG TECHNOLOGICAL UNIVERSITY

A Thesis
Submitted to the Nanyang Technological University
in fulfillment of the requirements for the degree of
Master of Engineering

2006

Acknowledgements

In the name of Allah SWT, The Most Gracious, The Most Merciful. All praises belong to Allah SWT, Lord of the universe. This project would not have been possible without His will.

The author would like to express his gratitude to his supervisor Dr. Ashraf Mohamed Hefny for his patient guidance, advice and help throughout the course of this project.

The author would also like to thank Mr. Hon Hoy Seetoh and his staff from Land Transport Authority (LTA), Mr. Nick Osborne and his staff from contract C825 Marina Line, and Mr. Khor Eng Leong and his staff from Contract C823 Circle Line for their kind assistance during the data collection period

Table of Content

List of Figures

List of Tables

Abstract

Ground surface settlement trough associated to tunneling is characterized by two important parameters: the maximum surface settlement at the point above the tunnel centerline (S_{max}) and the width parameter (i) which is defined as the distance from the tunnel centerline to the inflection point of the trough. The estimation of these settlement parameters is a very complex problem due to uncertain nature of the soil. Over the years, many methods have been proposed to predict the tunneling-induced settlements. Most of these methods are empirical in nature. However, a method with high degree of accuracy and consistency has not yet been developed. Accurate prediction of settlement is essential since settlement is the governing factor in the design process of the tunnels. In this research, the use of artificial neural network (ANN) for the prediction of maximum surface settlement and trough width is explored. The ability of ANNs to learn from examples and generalize beyond the training data has made it a potential alternative tool for the settlements prediction. ANNs are numerical modeling techniques that are inspired by the functioning of the human brain and nerve system. ANNs have been used successfully to solve many problems in the field of geotechnical engineering and some of their applications are demonstrated in this report.

In this research, two main analyses have been performed. In the first main analysis, the feasibility of using artificial neural networks to predict the maximum settlements due to tunneling was investigated. A total of 158 case records collected from contracts NEL C705, Marina Line C825 and Circle Line C823 are used to develop and verify the ANN models. Eleven input parameters considered to have significant impact on the settlement are used in the initial analysis of the ANN models. These include cover, advance rate, earth pressure, average SPT blow count of the soil layers above tunnel crown (SPT1), average SPT blow count at tunnel springline level (SPT2), average SPT blow count at tunnel inverted level (SPT3), bulk density of soil, stiffness of soil,

ground water level, moisture content and grout pressure. Thirty two network models are developed from the combinations of these eleven parameters. The type of ANNs used is multi-layer perceptrons(MLPs) trained using the back-propagation algorithm. The results of the analysis show that the network model with eight input parameters (soil cover, advance rate, earth pressure, SPT1, SPT2, moisture content, stiffness and grout pressure) and eight hidden neuron is optimum. The effect of network parameters including momentum, learning rate, and transfer function on the performance of the model is investigated as well. In this research, the use of self-organizing map and fuzzy clustering as alternative data division methods is examined and the results are compared with those obtained using statistical consistent method. The performance of optimum ANN model is compared with the commonly used empirical method. It was found that ANN model can predict the settlement with relatively high degree of accuracy, whereas the empirical method underestimates the settlements with respect to all case records in the validation set. This shows that ANN method outperforms the empirical method and thus it can be used as an effective tool to obtain more accurate prediction of tunneling-induced maximum settlements.

In the second main analysis, neural network models were developed for the prediction of maximum settlements and trough width using the data generated from the finite element software PLAXIS. PLAXIS is commonly used for geotechnical applications in which soil models are used to simulate the soil behavior. A total of 2161 patterns were generated from the combinations of six input parameters, namely coefficient of earth pressure (K_0), bulk density (γ), cohesion (c), ratio of stiffness over cohesion (E/c), ratio of depth over diameter (H/D), and volume loss (VL). Patterns which produce failed results were discarded; hence leaving 1836 patterns for the analysis. Two cases of network training are considered in the analysis. In the first case, network models are developed using three statistically consistent data sets i.e. training, testing and validation. In the second case, the input data is divided into training and testing sets only. For each case, two types of network models were developed. The first model has two output neurons where the predictions of S_{max} and trough width (i) can be carried

out simultaneously. The second model is the networks with one output neuron which predict S_{max} and (i) separately. The analysis indicated that two separate networks performed better than the single network used to predict S_{max} and (i) simultaneously as shown by higher correlation coefficients and lower error rates when the two networks are used. The use of faster training algorithms to improve the accuracy of the network is investigated and the results are compared with the result of standard gradient descent method. The optimum network models from the two cases of training and the optimum network trained using gradient descent method are retained and validated using a set of field data to examine the generalization ability. The field data were collected from contracts C823 Circle Line and C825 Marina Line. The validation results shows that ANN models are optimal and they can be used as an effective tool to obtain more accurate prediction of maximum settlements and trough widths in the field.

Chapter 1

Introduction

1.1 Background

The use of underground space has become more and more important throughout the world. Today, besides for supply lines of gas, water, electricity, telecommunications and disposal lines, the ground also provides space for transport tunnels for rail commuter traffic, long-distance trains, motor vehicles and pedestrians. In large cities, big structures such as four or five-storey administrative building, subterranean shopping malls, storage rooms, covered watercourses, production halls, offices, and sports facilities are also set up underground. Tunnels for roads and railways are key infrastructures in transport development that provide important impulses for the economic power of a region or a nation (Haack 2000). In many countries, the construction of transport tunnels and subsurface construction in general has reached a high standard. Cities in Europe, such as London, Paris, Budapest, Hamburg, Berlin have built and developed underground metros, urban railways or rapid transport system for the past 30 to 40 years. Similar is the situation in Japan, Taiwan, South Korea as well as other countries in the world.

In a country like Singapore, going underground for infrastructural works becomes increasingly a necessity due to high population density and limited land area. Other advantages include saving up the surface for better use, direct saving of energy, protecting against natural disasters and decreasing the maintenance cost (Lathauwer 1992). Numerous tunneling activities have been carried out extensively in Singapore for the past two decades. Before 1982, the tunneling works had been carried out to some limited extent. In 1983, a 3.7 m diameter tunnel approximately 3 km long was constructed using earth pressure balance machine (EPBM) to accommodate a 3.3 m

diameter sewage pipeline. The first major use of tunnels was the construction of the MRT railway through the densely developed central areas of Singapore city commencing in 1983. The MRT network has been expanding ever since. The latest major developments are the Changi Airport Line and the Northeast Line, with 17 underground stations, which made up a total of 23.5 km of bored tunnels. Tunneling for the Deep Tunnel Sewerage System has commenced in year 2001 with a total of approximately 48 km of main tunnels and many subsidiary link tunnels. Other developments include the tunnels for the Marina Line MRT, the Cable Tunnels, Deep Tunnel Sewerage Scheme (Phase 2), Circle Line, Kallang expressway and future underground caverns.

Ground movements and consequential surface settlements associated to tunneling are major concern in the design of tunnels in urban areas. This is mainly due to the damage that ground movements may cause to overlying and nearby building and services. An engineer in charge of tunnel design must be able to estimate settlement distribution along the tunnel route so that he can identify in advance options to minimize the damage to nearby structures. Several options are re-specifying the excavation and/or lining technique to reduce those settlements, rerouting the tunnel away from 'sensitive' buildings or buried services, using injection techniques or freezing to stabilize pockets of weak ground on the original route, protecting the existing building by underpinning the foundations, relocating buried pipelines, which is an expensive option (Attewell 1977). Hence ground surface settlement prediction is very essential to assess the effects of the tunnel construction on existing structures, design of the tunnels, and method of tunneling.

In the early stage of design, estimate of ground surface settlement can be obtained using simple empirical formulas, such as those proposed by Peck (1969), Attewell(1977), Attewell and Woodman (1982), O'Reilly and New(1982), and Mair et al (1993). Simple equations based on the theory of elasticity (Uriel and Sagaseta 1989) may also be used. For the final design, a more accurate method such as finite difference

or finite element is required (Rowe and Lee 1992). In this project, a neural-network based approach is used to analyze and predict settlement due to tunneling. Over the years, Artificial Neural Network (ANN) has been widely used in many fields to predict certain output given input values. Selection of the best stocks in the market, weather prediction, identification of people with cancer risk are some tasks that have been carried out successfully using the network of prediction. One advantage of ANN is its capability to establish the non-linear relationship between a set of input variables and the corresponding output without a need for predefined mathematical equations. In geotechnical field, relationships between input and output in most problems are often difficult to model as we are dealing with unpredictable behavior of the soil. In this situation, the problems are normally solved using empirical formulae. The above advantage makes ANN a powerful alternative to solve highly non-linear and complex geotechnical problems, particularly for engineering predictions.

1.2 Objectives and Scope of study

The main objective of this research is to study the feasibility of using ANN method for predicting the maximum ground surface settlement and settlement trough width induced by tunneling. To achieve the main objective, two main analyses are carried out in the project. In the first main analysis, the feasibility of using artificial neural networks to predict the maximum settlements due to tunneling was investigated. In the second main analysis, neural network models were developed for the prediction of maximum surface settlement and trough width.

The works in the first main analysis include the following aspects:

1. Developing neural network models using the input and output data collected from site offices and Land Transport Authority database. The input data were analyzed and preprocessed so as to produce the relevant inputs required for training, testing, and validation. The neural network models were tested in order to obtain the optimum network model.

2. Conducting a sensitivity study on the effect of network parameters including number of hidden neurons, learning rate, momentum, and transfer function on the network performance so as to obtain optimum network architecture.
3. Investigating the effect of using other faster training algorithms on the accuracy of the optimum network and comparing the result with gradient-descent method.
4. Studying the effect of several data division methods and comparing the result with the commonly used statistically consistent method.
5. Conducting sensitivity analysis in order to identify which of the input variables have the most significant impact on settlement predictions.
6. Comparing the settlement predictions from the neural network and conventional methods with the actual settlements.
7. Proposing a simplified neural network model suitable for initial prediction of maximum surface settlements when only basic soil properties are known.

In the second main analysis the following aspects are covered:

1. Generating the relevant input and output data from finite element software for the training of neural network model.
2. Evaluating the effect of two stopping techniques (using three sets and two sets of data) on the performance of the network.
3. Investigating the performance of network models with two output neurons for simultaneous predictions of maximum surface settlement and trough width and comparing the results with two independent networks whereby one is to predict maximum surface settlement and the other to predict trough width.
4. Examining the use of faster training algorithm and variation of the hidden neuron number so as to obtain optimum network model.
5. Validating the reliability of optimum networks for the predictions of maximum surface settlement and trough width by testing with field data.

1.3 Layout of the report

In Chapter 2, issues related to the main structure and operation of ANNs are discussed. This includes layers of units, activation functions and learning rule. The popular back propagation algorithm used in this project with its limitations is described. Several faster training algorithms and their comparison with back propagation algorithm are briefly presented. In Chapter 3, applications of ANNs in the field of Geotechnical Engineering are reviewed to demonstrate the relative success of ANNs in this field.

In Chapter 4, the empirical method generally used to predict tunneling-induced settlements is described. This method requires the computations of few parameters namely maximum surface settlement, inflection point of settlement trough and ground loss. The mechanisms of soil settlement induced by tunneling are presented. In Chapter 5, more details of Contract C705 Northeast Line, C823 and C825 Circle Line were given including geological section and bored tunnel construction. The input data for this project were collected from site offices of the three projects.

In Chapter 6, the analysis of neural network models is presented. The database used to develop ANN models is the combination of field data from contracts C705, C823, and C825. The performance of several softwares is assessed and the most accurate one is selected for the ANN analysis. The potential factors, which may have significant effect on the settlement, are highlighted. Data division, pre-processing of data, model architecture and stopping criteria for the initial phase of the analysis are discussed. ANN models with different input combinations are investigated in order to obtain the optimum input combinations. The network model with the best input combinations (referred to as ANN model) is used throughout the remaining analysis. The effect of learning rate, number of hidden nodes, momentum term, and transfer functions on the accuracy of ANN model is examined. The performance of ANN model using various training algorithms is measured and the best algorithm is used for ANN model. Three data division methods for the development of ANN models are presented and the

results obtained using each method are compared and evaluated. The relative importance of the input factors affecting settlement is investigated. A comparison of the results obtained using ANN model and the most commonly used empirical method is presented. An alternative model of ANN which can give quick predictions of settlements is proposed.

In Chapter 7, the neural network models are developed using the input generated from finite element program PLAXIS. The outputs of the models consist of maximum surface settlement (S_{max}) and trough width (i). Two cases of training are considered, namely training using three data sets and two data sets. For each case, two types of network are tested, namely network with two outputs neuron and network with one output neuron. The former is used to predict both S_{max} and (i) simultaneously, while the later is used to predict S_{max} and (i) separately. The use of faster training algorithms to improve the accuracy of the network is investigated and the results are compared with the result of standard gradient descent method. The best network models are retained and validated using the field data to examine the generalization ability.

In Chapter 8, the research work is summarized and conclusions are presented. Recommendations for future works are given as well.

Chapter 2

Artificial Neural Networks

2.1 Introduction

Artificial neural networks are composed of simple elements operating in parallel, which are designed to simulate the behavior of biological neural networks for several purposes. These include pattern recognition, identification, classification, speech, vision and control systems. The surge of interest in neural networks is mainly based on the wish to build machines that are capable of performing complex tasks for which the programmable computers invented by von Neumann (1946) are not suitable. The research in the field of artificial neural networks has started ever since the general theory of information processing based on the so-called neurons was proposed in 1943 by McCulloch and Pitts. In the 1950's, several neural network models such as the perceptron (Rosenblatt 1962) and Adaline (Widrow and Hoff 1960) were invented. Due to the failure of perceptrons to be successfully applied to more complex sets, the research in the field of artificial neural networks almost stopped around 1970. The interest in neural networks was revived with the discovery of the so-called back-propagation algorithm by Rumelhart et al. (1986). Today, the neural networks have been rapidly developed through extensive research that they have been applied in many important areas such as business, aerospace, automotive, banking, defense, electronics, entertainment, industrial, medical, securities, telecommunications, and transportation.

The architecture of neural networks is originally inspired by that of the human brain. The human brain forms a massive communication network, consisting of billions of nerve cells, also known as neurons (Figure 2.1). The neuron receives incoming

impulses via the dendrites and the impulses are transmitted via the axon and synapses to other neurons.

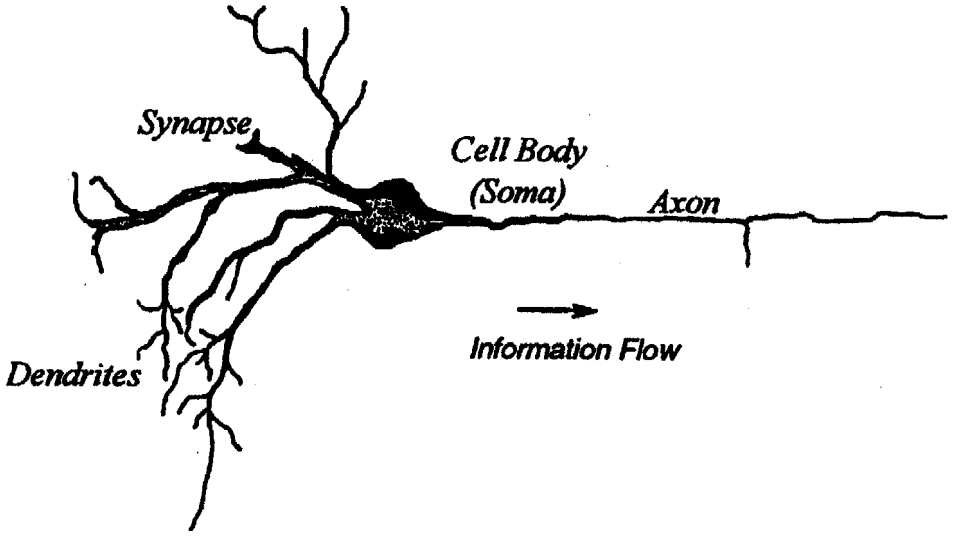


Figure 2.1: Sketch of a biological neuron (after Tsoukalas and Uhrig, 1997)

2.2 Structure and Operation of Artificial Neural Networks

The first modeling of neurons was carried out by McCulloch and Pitts in 1940s. They proposed the model of a synthetic neuron with its inputs and outputs are Boolean values. In neural network, a synthetic neuron is also called a processing element (PE), a unit, or a node. The schematic diagram of a processing element is shown in Figure 2.2. Each PE receives inputs from other PEs, performs a weighted summation, applies an activation function to the weighted sum, and outputs its results to other neurons in the network (Sundararajan and Saratchandran, 1998).

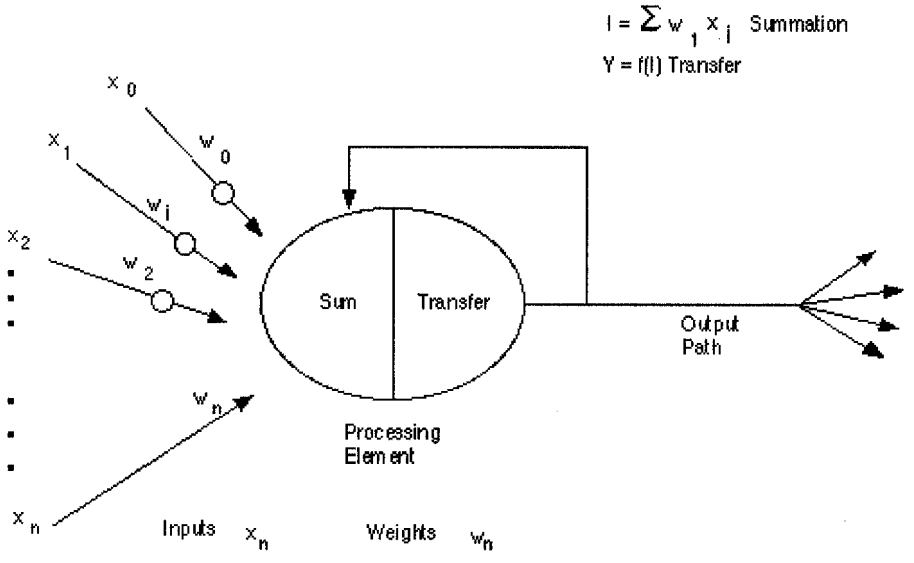


Figure 2.2: A processing element of ANN and its operation (Anderson and McNeil, 1992)

2.2.1 Layers of units

Units, PEs, or nodes are usually arranged in layers, each layer consisting of at least one neuron. A single-layer neural network in which there is no hidden layer is the characteristic of the simple perceptron model. Multi-layer feed-forward neural networks consist of multiple layers of neurons with input, hidden, and output layers (Figure 2.3). Each layer other than input layer, receives their inputs only from all neurons in the previous layer and from one bias signal source. Bias or threshold is a value that the summation of the input signals must exceed before it can be transmitted (Weijters and Hoppenbrouwers 1995). A two layer feed forward network (only one hidden layer) has been used in most applications.

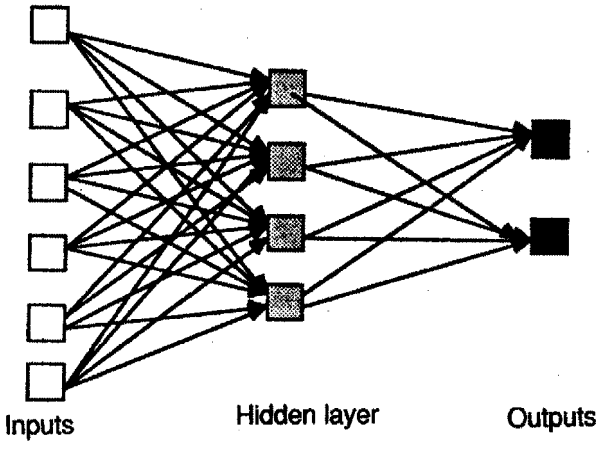


Figure 2.3: Structure of a Multi-Layer Feedforward ANN

2.2.2 Activation functions

The net input of units in neural networks is transformed by using a scalar-to-scalar function called an "activation function" to yield unit's activation value. The activation value is then fed via synaptic connections to one or more other units. The activation functions or transfer functions are very essential to introduce the nonlinearity to the network. Without nonlinearity, hidden units would not make nets more powerful than just plain perceptrons (which do not have any hidden units, just input and output units) (Sarle 2002). The nonlinearity will make the network able to represent nonlinear functions and this is why multilayer networks can be so powerful. Three commonly used activation functions are logistic, tanh, and linear activation functions.

Linear activation function:

$$y = D*x \tag{2.1}$$

where x is the input to the neuron, y is the final value of the neuron and usually D = 1.

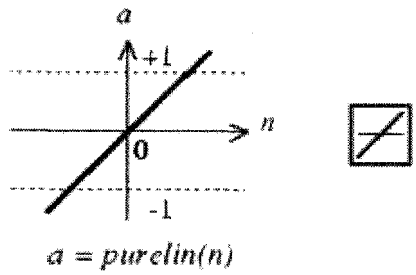


Figure 2.4: Linear transfer function

Most functions are difficult to approximate using linear function, hence nonlinear functions such as logistic and tanh are used. Nonlinear activation functions are used in most backpropagation networks because they are differentiable.

The standard sigmoid (or logistic) runs from 0 to 1 and it is:

$$y = 1 / (1 + \exp(- x)) \tag{2.2}$$

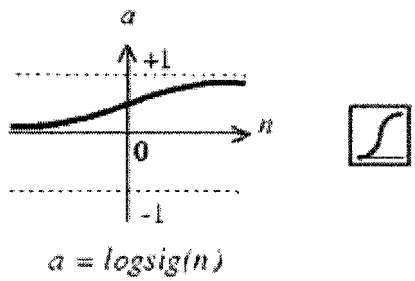


Figure 2.5: Log-Sigmoid transfer function

The function tanh has outputs in the range -1 to 1 and can be written as:

$$y = 2 / (1 + \exp(-2 * x)) - 1 \tag{2.3}$$

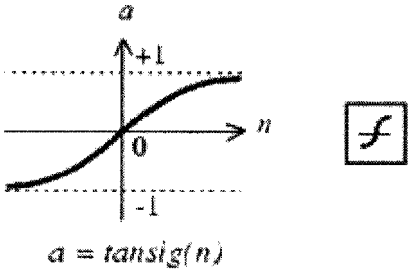


Figure 2.6: Tan-Sigmoid transfer function

Jordan (1995) stated that the logistic function is an excellent choice for binary (0/1) targets. Since its values are centered around 0, using tanh will result in faster training (Brown et. al, 1993). However, experiments by Tveter (1998) show that sometimes tanh is better but sometimes it is not. The backpropagation networks with linear output units and a single layer of non-polynomial hidden layer units can represent closely most reasonable functions (Leshno et al. 1993). Other transfer functions include Gaussian function, hard-limit transfer function, radial basis function, triangular basis function, softmax transfer function, satlin transfer function.

2.2.3 Learning Rule

Learning paradigm or learning rule is a procedure for modifying the weights and biases of a network. The neural network learns to solve a problem as its weights changes. Two learning rules commonly used are supervised learning and unsupervised learning. In supervised learning, the network is fed with the inputs and the resulting outputs are compared against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process is repeated over and over until the minimum error or desired accuracy is reached. In unsupervised training, the network is provided with inputs but not with desired outputs. The hidden neurons must find a way to organize themselves without help from the outside. Unsupervised learning is more representative of a real life

learning where the input-output sets do not exist. Kohonen (1995) developed a network known as self-organizing map using this learning method. It is sometimes called an auto-associator, which learns without the benefit of knowing the right answer. Clustering operations, where the input patterns are categorized into a finite number of classes, are performed mostly using unsupervised learning. An example of such application is vector quantization.

Several mathematical algorithms are used to update the connection weights and biases during network training. Hebb (1949) introduced the first learning rule later best known as Hebb's rule. According to this rule, if a neuron receives an input from another neuron, and if both are highly active (mathematically have the same sign), the weight between the neurons should be strengthened. Hebb's rule forms the foundation of later learning laws such as Hopfield Law and the popular Delta Rule. Developed by Widrow and Hoff (1960), the delta rule formed the basic concept of the well-known neural network type called Feed forward, Back-propagation.

2.3 Backpropagation

Since the publication of the Parallel Distributed Processing volumes by Rumelhart et al. in 1986, learning by backpropagation has become the most popular method of training neural networks. This is due to the relative power and the underlying simplicity of the algorithm. It is powerful because, unlike its precursors, the perceptron learning rule and the Widrow-Hoff learning rule, it can be employed for training nonlinear networks of arbitrary connectivity (Rumelhart et al. 1995). It is simple because the basic idea is to define an error function and use gradient descent to find a set of weights which optimize performance on a particular task (Rumelhart et al. 1995). In fact, backpropagation is little more than an extremely judicious application of the chain rule and gradient descent (le Cun 1988). It is a straightforward but elegant application of the chain rule of elementary calculus (Werbos 1994). The name backpropagation actually comes from the term used by Rosenblatt (1962) for his effort to generalize the

perceptron learning algorithm to the multilayer case. Unlike simple delta rule or the perceptron rule, which can be used for a single layer network only, backpropagation learning rule is mostly used in multilayer networks. The typical back-propagation network has an input layer, an output layer, and at least one hidden layer, as shown in Figure 2.7.

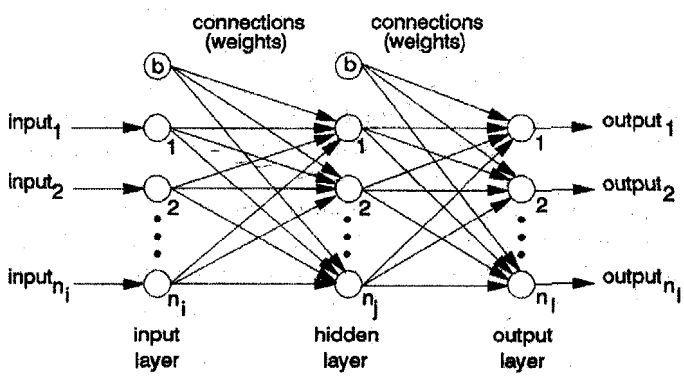


Figure 2.7: The typical back-propagation network structure

The network most commonly used with the backpropagation algorithm is the multilayer feedforward network. As mentioned previously, each input to the network is multiplied by respective weight and the weighted sum of the inputs and the bias are then fed to the transfer functions (f) to generate the output. Any differentiable function can be used as transfer function. The most commonly used functions for backpropagation are log-sigmoid and tan-sigmoid transfer functions

2.3.1 The Back Propagation Rule

In a multilayer feed-forward network with backpropagation learning, the objective is to find a set of weights that minimizes the network error function. Backpropagation rule is based on the idea of continuously modifying the strengths of the input connections to reduce the difference (the delta) between the desired output value and the actual output of a neuron. The connection weights are adjusted to minimize the error of the network. The error is back propagated into previous layers one layer at a time. The process of back-propagating the network errors continues until the first layer is reached. The main steps to train the network with backpropagation are as follows (Freeman and Skapura 1991):

1. Apply an input vector to the network and calculate the corresponding output values.
2. Compare the actual outputs with the correct outputs and determine a measure of the error.
3. Determine in which direction (+ or -) to change each weight in order to reduce the error.
4. Determine the amount by which to change each weight.
5. Apply the corrections to the weights.
6. Repeat items 1 through 5 with all the training vectors until the error for all vectors in the training set is reduced to an acceptable value.

Saratchandran and Sundararajan (1998) divided the backpropagation learning phase into a forward phase and a backward phase. Forward phase starts from initializing the weights to calculating the actual outputs. Based on the difference between actual and desired outputs (error), weights are adjusted to reduce the difference; this is the backward phase. Details of the two phases during training for a single training-vector pair are given in the following sections.

- ***Forward Phase***

The input vector $X = (x_1, x_2,..., x_{Ni})^T$ is multiplied by the hidden layer weight matrix W_h to obtain the net-input values to the hidden layer units (Eq. 2.1). The net-input values are passed through activation functions to produce outputs from the hidden layer (Eq. 2.2). Equation 2.1 is applied to these outputs again to calculate the net-input values to each unit in the output layer (Equation 2.3). The net-input values are fed to output activation functions to obtain actual outputs. (Eq. 2.4).

$$\text{net} - \text{input}_j = i_{h,j} = \sum_{i=1}^{N_i} w_{h,ji}\, x_i \tag{2.1}$$

$$\text{output}_j = y_{h,j} = f(i_{h,j} - b_j) \tag{2.2}$$

$$\text{net-input}_k = i_{o,k} = \sum_{j=1}^{N_j} w_{o,kj}\, y_{h,j} \tag{2.3}$$

$$\text{output}_k = y_{o,k} = f(i_{o,k} - b_k) \tag{2.4}$$

where: $w_{h,ji}$ is the weight connecting input unit i to unit j in the hidden neuron layer, function f is a nonlinear activation function, $w_{o,kj}$ is the weight connecting hidden unit j to unit k in the output neuron layer, N_i is the number of neurons in input layer, N_j is the number of neuron in hidden layer, b_j and b_k are the bias term of the unit in the hidden and output layer . The indices i, j, and k are used to denote the input, hidden, and output neuron layers, respectively.

- ***Backward Phase***

The actual output and the desired output are compared and the difference (error) is used to adapt the weights to reduce the overall error measure E for a training set of P patterns. The error measure E_p for a training pattern p is the sum of the squared errors of the actual output and desired output of the neurons in the output layer (Eq. 2.5).

$$E_p = \frac{1}{2}\sum_{k=1}^{N_o}(d_{p,k} - y_{p,o,k})^2 \tag{2.5}$$

where $d_{p,k}$ and $y_{p,o,k}$ are the desired output and the output of the neuron in output layer for training pattern p respectively, N_o is the number of neurons in output layer.
The overall error measure for a training set of P patterns is:

$$E = \sum_{p=1}^{P} E_p \tag{2.6}$$

In the following expressions, the pattern index *p* has been omitted on all variables to improve clarity. The error terms δ for the output units and hidden units can be shown to be (Rumelhart et al. 1986):

$$\delta_{o,k} = y_{o,k}(1 - y_{o,k})(d_k - y_{o,k}) \tag{2.7}$$

$$\delta_{h,j} = y_{h,j}(1 - y_{h,j})\sum_{k=1}^{N_o}\delta_{o,k}\, w_{o,kj} \tag{2.8}$$

If learning by epoch is applied, changes of weight on the output layer and hidden layer are:

$$\Delta w_{o,kj} = \eta\, \delta_{o,k}\, y_{h,j} \tag{2.9}$$

$$\Delta w_{h,ji} = \eta\, \delta_{h,j}\, x_i \tag{2.10}$$

where η is the learning rate coefficient. The output layer weights and hidden layer weights are updated accordingly.

$$w_{o,kj}' = w_{o,kj} + \Delta w_{o,kj} \tag{2.11}$$

$$w_{h,ji}' = w_{h,ji} + \Delta w_{h,ji} \tag{2.12}$$

The method used for finding the correct weight changes (Δw) is known as gradient descent. Henseler (1995) explained this method as follow. If E is considered as a function of weights ($w_1,\ldots,w_n$), then the gradient of E with respect to w denotes the slope of the "error-surface". By descending this surface downhill, i.e., in the direction of the negative gradient, we will finally reach at the bottom of the surface. At that point the error can no longer be decreased and the procedure finishes. Details of the backpropagation learning rule are described in Rumelhart et al. (1986).

- ***Momentum***

Typical characteristic for many gradient descent methods is its slow convergence. Hence, efforts have been made to speed up the learning process of back propagation. One method is to include the momentum term in the back propagation learning rule. Imagine a ball rolling down a hill. As it does so, it gains momentum, so that its speed increases and it becomes more difficult to stop. Similar is the function of momentum term. A little of the previous iteration's weight changes are added to the weight changes for the current iteration. How small to make the additional changes is controlled by a parameter α called the **momentum**, which is set to a value between 0 and 1. The weight changes on the output layer then become:

$$\Delta w_{o,kj}(p+1) = \eta \delta_{o,k}(p+1)\, o_{h,j}(p+1) + \alpha\, \Delta w_{o,kj}(p) \qquad (2.12)$$

Where p is the training pattern index and α is the momentum term. The weights are then updated

$$w_{o,kj}'(p+1) = w_{o,kj}(p) + \Delta w_{o,kj}(p+1) \qquad (2.13)$$

The equations for updating of the hidden weights can be derived in the same manner. Here, delta weight equation (Δw) is modified so that a portion of the previous delta weight is fed to the current delta weight. Hagan et al. (1996) stated that momentum allows a network to respond not only to the local gradient, but also to recent trends in

the error surface. Acting like a low-pass filter, momentum allows the network to ignore small features in the error surface. Without momentum a network may get stuck in a shallow local minimum. With momentum a network can slide through such a minimum. If momentum is used, the learning rate can be increased without leading to oscillation. Higher learning rate means more rapid learning as the changes in the weights become larger.

- ***Faster Training Algorithms***

The commonly used backpropagation training algorithm, gradient descent with momentum, is often too slow for practical problems. Attempts have been made to discover new training algorithms which can converge faster than the algorithms discussed above. From the analysis of standard descent algorithm, two more techniques are developed, namely variable learning rate backpropagation and resilient backpropagation. Other high performance algorithms are derived from the standard numerical optimization techniques. Three types of numerical optimization techniques for neural network training are conjugate gradient, quasi-Newton, and Levenberg-Marquardt.

1. Variable learning rate backpropagation

With gradient descent method, the learning rate is held constant throughout training. The chosen learning rate may not be the optimal one as it is unlikely to determine the optimal learning rate before training. Too high learning rate will cause the algorithm to oscillate and become unstable. Too small learning rate will cause the algorithm to converge very slowly. In fact, the optimal learning rate changes during the training process, as the algorithm moves across the performance surface. Variable learning rate backpropagation allows the learning rate to change during the training process. This will improve the performance of gradient descent algorithm as the learning rate is made responsive to the complexity of the local error surface.

2. Resilient backpropagation

The problem of using logistic transfer function is that the slope approaches zero when the input gets large. As a result, the gradient can be very small and this causes small changes in the weights and biases. Yet the weights and biases are still far from their optimal values. Resilient backpropagation technique tackles this problem by taking only the sign of the derivative to determine the direction of the weight update. The magnitude of the derivative plays no role on the weight update. Instead, a factor is used to update the value for each weight and bias. The update value is increased by the factor whenever the derivative of the performance function with respect to that weight has the same sign for two successive iterations. The update value is decreased by the factor whenever the derivative with respect to that weight changes sign from the previous iteration.

3. Conjugate gradient

In backpropagation, the performance function (error) decreases most rapidly in the steepest descent direction (negative of the gradient). However, this does not mean that the fastest convergence is already achieved. In standard gradient descent method, the length of the weight update (step size) is determined by learning rate and thereby a constant. In the conjugate gradient algorithms, a search is made along the conjugate gradient direction to determine the step size, which minimizes the performance function along that line. All of the conjugate gradient algorithms start out by searching in the steepest descent direction (negative of the gradient) on the first iteration. Then the next search direction is determined so that it is conjugate to previous search directions. The general procedure for determining the new search direction is to combine the new steepest descent direction with the previous search direction. The step size is adjusted at each iteration; this produces generally faster convergence. There are four different variations of conjugate gradient algorithms namely Fletcher-Reeves update, Polak-Ribiére update, Powell-Beale restarts, and Scaled Conjugate Gradient.

4. Quasi-Newton Algorithm

Newton's method is used to update values of the weights and biases. This algorithm often converges faster than conjugate gradient methods. The most successful quasi-Newton method is the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) update. Since the BFGS algorithm requires more storage and computation in each iteration than the conjugate gradient algorithms, there is need for a secant approximation with smaller storage and computation requirements. The one step secant (OSS) method is an attempt to bridge the gap between the conjugate gradient algorithms and the quasi-Newton (secant) algorithms. It can be considered a compromise between full quasi-Newton algorithms and conjugate gradient algorithms.

5. Levenberg-Marquardt

The Levenberg-Marquardt algorithm is a general non-linear downhill minimisation algorithm for the case when derivatives of the objective function are known (McLauchlan, 2002). It is derived from further modification of Newton's method where it dynamically combines Gauss-Newton and gradient-descent iterations. This algorithm ensures that the performance function will always be reduced at each iteration. Levenberg-Marquardt appears to be the fastest method for training moderate-sized feedforward neural networks (up to several hundred weights) (Hagan et al. 1996).

All the methods discussed above are local optimization methods; they normally find local optima. There is no guarantee that a global optimum will be obtained using the above methods. One approach to find global optimum for any of the above methods is to use numerous random starting points. Another way is to use more complicated methods designed for global optimization such as simulated annealing or genetic algorithms. Experiments conducted by Hagan et al. (1996) show that generally the Levenberg-Marquardt algorithm will have the fastest convergence for networks that contain up to a few hundred weights. However, this applies on function approximation problems only. As the number of weights in the network increases, the performance of this algorithm decreases. Levenberg-Marquardt does not perform well on pattern

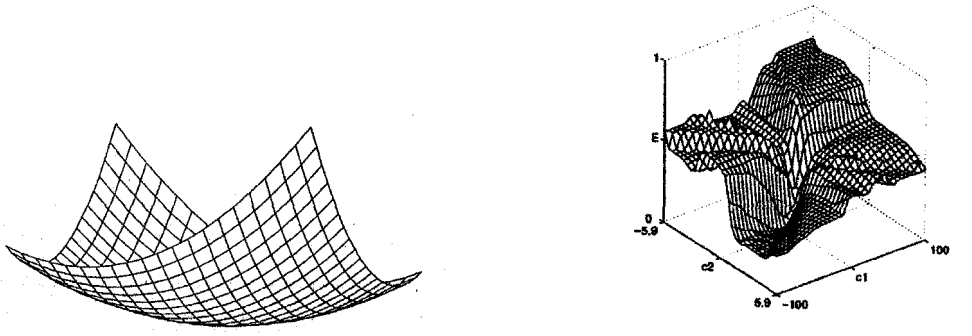
recognition problems. The fastest algorithm on these problems is Resilient backpropagation (Hagan et al. 1996). Yet, its performance is relatively poor on function approximation problems. For a wide variety of problems, the conjugate gradient algorithms seem to perform well particularly for networks with a large number of weights (Hagan et al. 1996).

2.3.2 Limitations & Cautions

There are some problems associated with back-propagation algorithm, which have not been resolved up to now.

- **Local Minima**

In a gradient descent procedure, the system will follow the contour of the error surface and always move downhill in the direction of steepest descent. For single-layer linear model (e.g. the *least-mean-square* (LMS) learning paradigm), this is not a problem as it always have bowl-shaped error surfaces (Figure 2.8). Hence, the global minima will always be found. However, in multilayer networks, the error surfaces become more complex with many minima (Figure 2.8). Many algorithms that rely on a sequential search over the error surface may become trapped in local minima. In the gradient-descent technique, the use of momentum term helps to minimize this problem. Another approach to overcome local minima and make training more efficient is proper initialization of weights. It is recommended that the neural network is reinitialized and retrained several times to ensure that the best solution is obtained. Two more powerful techniques to avoid local minima are simulated annealing and genetic algorithm. Simulated annealing is easy to understand and implement and has low memory requirements; whilst genetic algorithm is more complex and has quite large memory needs but is generally superior (Masters 1993).



(a) (b)

Figure 2.8: Error surfaces: (a) for single-layer network (after Henseler 1995), (b) for multilayer network (after Kordos and Duch 2003)

- **Slow Training**

The standard back-propagation algorithm consumes large quantities of computing time due to extensive calculations of error derivatives and updating the weights. However, many methods have been proposed to speed up the training of backprop, such as Quickprop (Fahlman 1989) and RPROP (Riedmiller and Braun 1993). Another approach is to use faster algorithm for nonlinear optimization such as conjugate gradients, Levenberg-Marquardt, etc.

- **Underfitting and Overfitting**

The critical issue in developing a neural network is generalization. It measures how accurate the neural network in predicting cases that are not in the training set. Neural network model can also suffer from either underfitting or overfitting depending on the complexity of the network. A network that is too simple will be unable to detect fully the pattern in a complicated data set, leading to underfitting. On the other hand, a very complex network will fit the pattern and the noise as well, leading to overfitting. The best way to avoid overfitting is to use as many data as possible for the training set. The most commonly used method to overcome underfitting and overfitting problem is Early Stopping. In Early Stopping, the available data is divided into training, testing, and

validation sets. During training, the error rate of testing test is monitored periodically and training is stopped when the error rate "starts to go up". The network is then run on the third set of data, validation set, to estimate its generalization ability.

- **Network Architecture**

To date, there is no hard and fast rule or even a satisfactory empirical formula which can be used to determine the dimension of a network (no. of hidden layer, no. of hidden units) for a particular problem. However, in Multi Layer Perceptrons with any of a wide variety of continuous nonlinear hidden-layer activation functions, one hidden layer with an arbitrarily large number of units suffices for the "universal approximation" property (Hornik et al. 1989). For the vast majority of practical problems, there is no reason to use more than one hidden layer (Masters 1995). Number of hidden neurons in hidden layer is also vital to the performance of the network. Too few neurons will render the network incapable of solving the problem due to lack of resources; whilst too many neurons will cause over fitting problem where the network displays low training error but high generalization error. Many books and articles offer "rules of thumb" for choosing number of hidden units. However, the reliability is doubted as they ignore the number of training cases, the amount of noise in the targets, and the complexity of the function. An intelligent choice of the number of hidden units depends on whether we are using early stopping or some other form of regularization. If early stopping is used, it is essential to use many hidden units to avoid bad local optima (Sarle 2002). The common way adopted to determine the optimum number of hidden unit is by trial and error. Many networks with different numbers of hidden units are tested and generalization error for each one is recorded. The network with minimum estimated generalization error gives the optimum number of hidden neurons.

Chapter 3

Applications in Geotechnical Engineering

3.1 Introduction

Geotechnical engineering is a field involving large number of uncertain parameters due to the fact that we are dealing with the product of nature (the ground). In many circumstances, prediction of ground behavior is a major problem because of limited fundamental understanding of soil and rock behavior (Toll 1996). In this situation, many researchers resort to empirical approaches to solve the geotechnical problems. As mentioned previously, Artificial Neural Network is very useful for problems where there is no direct relationship between the input and the output. Thus it should be ideally suited for application in the field of geotechnical engineering. Artificial Neural Network applications was started to be used in the field of geotechnical engineering in 1991. The literatures reveal the successful use of ANNs in pile capacity prediction, predicting the settlement of structures, modeling soil properties and behavior, determination of liquefaction potential, site characterization, modeling earth retaining structures, evaluating stability of slopes and the design of tunnels and underground opening. Some of them are described in this chapter.

3.2 Liquefaction

When earthquake occurs, the soil will lose its strength due to shaking (liquefaction). Most often, soil liquefaction causes extensive damage to infrastructure and serious loss of life. The prediction of soil liquefaction is difficult due to many critical factors, which influence liquefaction, such as the properties of the soil, the depth of the soil deposit, the magnitude and intensity of the earthquake, the distance from the source of the earthquake, and the seismic attenuation properties (Goh 1995). A study by Goh (1995)

illustrated the potential use of neural networks to predict the seismic liquefaction. The popular back-propagation algorithm was adopted in the study, and the training was carried out using actual field records. A total of 85 case records, representing 42 sites with liquefaction and 43 sites without liquefaction, were evaluated using the neural networks. Eight input variables were tested in the models namely, earthquake magnitude (M), total vertical stress (σ_o), effective stress (σ_o'), the standard penetration test (SPT) value, normalized peak horizontal acceleration at ground surface (a/g), the equivalent dynamic shear stress, fines content, and the mean grain size of the soil. Different number of input variables were tried in order to determine the most reliable model. Table 3.1 summarizes the performance of several neural network models.

Table 3.1: Summary of neural networks performance (Goh, 1995)

Model	Input variables	Success Rates (%) Training	Testing
M4	M, $(N_1)_{60}$,τ_{av}/σ'_o,F	95	65
M5	M, $(N_1)_{60}$,τ_{av}/σ'_o,F, a/g	93	85
M6	M, $(N_1)_{60}$,τ_{av}/σ'_o,F, a/g, σ'_o	95	85
M7	M, $(N_1)_{60}$,τ_{av}/σ'_o,F, a/g, σ'_o, D_{50}	95	92
M7A	M, $(N_1)_{60}$,τ_{av}/σ'_o, a/g, σ_o, σ'_o, D_{50}	95	81
M8	M, $(N_1)_{60}$,τ_{av}/σ'_o,F, a/g, σ_o, σ'_o, D_{50}	97	92

The result of model M8 indicates high correlation between the input and output data for both training and testing set; thus, it is considered as the best model. Totally, there were 2 errors in the training data and 2 errors in testing data. This means a 95% success rate which is higher than the result using the Seed et al. procedure (84% success rate or 14 errors). This shows that neural network perform better than the more conventional method for evaluating liquefaction potential.

Ural and Saka (1998) used artificial neural network to evaluate soil liquefaction potential and resistance from the CPT data. For this study, eleven soil and seismic

parameters are selected as input, and seven different models are constructed by changing the input parameters. The best model gave an overall success rate of 92%. They concluded that comparisons between the neural network approach and simplified liquefaction procedure indicate that network results are as reliable as conventional methods.

A fuzzy adaptive neural network called "Fuzz-ART", based on adaptive resonance theory combined with fuzzy set theory is developed by Chern et al. (2001) to evaluate liquefaction potentials induced by Chi-Chi earthquake in Yuan-Lin area. The system is a combination of back propagation algorithm for parameter learning and the Fuzzy ART algorithm for structure learning. The training and testing data used in Goh's network were also used for the proposed network. The network produced only one error in the training data and also one error in the testing data. This indicates an overall success rate of 97.6 %, much higher than those of Seed et. al (84%) and Goh (95%). Furthermore, the present Fuzz-ART model converges much faster than Goh's model. The study shows that Fuzzy-ART neural network model is more reliable than the methods of conventional three-layer neural networks.

3.3 Pile Capacity

Determination of the axial load-bearing capacity of driven piles is a complex problem involving large number of uncertain parameters. One widely used approach to this problem is to use empirical design methods, which establish correlation between the soil parameters and pile capacity. Although it is not very accurate, this approach is simple and able to give quick estimate of pile load capacity. Neural network is a suitable alternative for this problem because of its inherent ability to incorporate the uncertainties associated with the controlling parameters.

Nawari et al. (1999) proposed neural network models to predict the axial and lateral load capacity of piles, using only simple input data such as SPT-N values and the

geometrical properties. Feedforward Backpropagation(BPNN) and Generalized Regression Networks(GRNN) are utilized in this study. The data were derived from 60 load test records and 23 full-scale laterally loaded drilled shafts tests. Complete input parameters include the SPT-values with depth, Pile length, cross-sectional area, circumference and the amount of steel reinforcement. Figure 3.1 presents the measured axial capacities of H-piles versus the results predicted by the BPNN, GRNN, AASHTO and SPT91 respectively. Overall, the results indicate that BPNN and GRNN can predict satisfactorily the total pile capacity, especially that of the H-piles. Their results are even better than those of AASHTO and the SPT91, which are considered here because they often provide better results than other empirical design formulas. The correlation coefficient for the neural network models ranges between 0.88 and 0.94, while for AASHTO and the SPT91, it varies between 0.65 and 0.78.

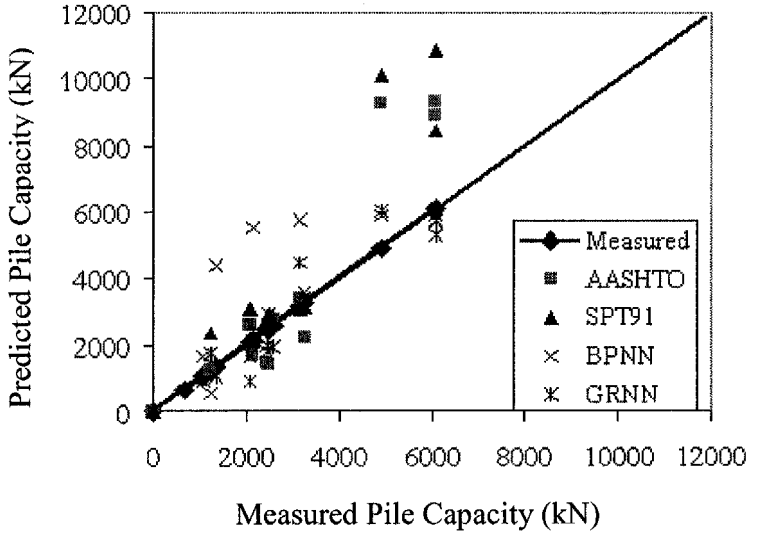


Figure 3.1: Comparisons of predicted and measured pile capacity for various methods (Nawari et al. 1999)

In another study by Teh et al. (1997), a back-propagation neural network model was used to estimate static pile capacity from dynamic stress-wave data. The database for training and testing of the network comprise 37 records of precast reinforced concrete

(RC) piles from 21 different sites. The CAPWAP procedure, proposed by Rausche et al. (1972), was used to determine the target values. Three neural network models, denoted as NN1, NN2, and NN3, were developed during the training phase. NN1 is set up to predict the total pile capacity only; the layer output has a single neuron. The resistance distribution of piles is predicted by NN2, which has, in output layer, 20 neurons representing the shaft resistance and one neuron for the toe resistance. NN3 predicted the damping and quake parameters as well as the soil resistance distribution. The output layer consists of 21 neurons to represent the resistance, and 4 neurons to keep the information on the damping and quake parameters. The results indicate that neural network models can provide good estimation of the total static capacity based on digitized force and velocity information alone. In addition, determination of resistance profile from stress-wave data becomes feasible as demonstrated by models NN2 and NN3. The study also shows that networks trained on RC pile data are capable of predicting the capacities of non-RC piles as well, but the accuracy is less.

McKinley (1996) studied the applicability of artificial neural networks to the interpretation of bearing capacity data. He concluded that artificial neural networks can successfully characterize the underlying pattern in scattered and uncertain data such as might be obtained from bearing capacity tests.

Other applications of ANNs for pile capacity include prediction of load bearing capacity of piles from Statnamic Pile Test Data (Javadi et al. 2001) and prediction of pile axial capacity and pile driving analysis (Hoi 1999).

3.4 Settlement of Foundations

For shallow foundations, the designs are usually governed by settlement factor. As a result, prediction of settlement is a major step in the design process. Shahin (2003) investigated the feasibility of using artificial neural networks for settlement prediction of shallow foundations on cohesionless soils. A total of 189 records were used to calibrate and validate the neural network models. The input parameters for the models include footing width, footing net applied pressure, average SPT blow count over the depth of influence of the foundation, footing geometry and footing embedment ratio. The models were trained with the back-propagation algorithm and the results were compared to traditional methods proposed by Meyerhof (1965), Schultze & Sherif (1973), and Schmertmann et al. (1978). Table 3.2 shows the comparisons of the results using ANN model and the three traditional methods.

Table 3.2: Results of ANN and traditional methods for settlement prediction

Performance measure	ANN	Meyerhof	Schultze and Sherif	Schmertmann et al.
r	0.905	0.44	0.729	0.798
RMSE	11.04	25.72	23.55	23.67
MAE	8.78	16.59	11.81	15.69

The above results indicate that ANN method surpasses the traditional methods, based on measures of correlation coefficient (r), root mean square error (RMSE), and mean absolute error (MAE).

Sivakugan et al. (1998) proposed a network with one hidden layer and 11 hidden nodes to predict the settlement of shallow foundations on sands. The study was carried out on data of 79 records, and five inputs were used to train the network. The optimum model has shown good performance compared to methods proposed by Terzaghi and Peck (1967) and Schmertmann (1970).

A neural network was developed by Goh (1994) to predict the settlement of a vertically loaded pile foundation in a homogeneous soil stratum. The input parameters were ratio

of the elastic modulus of the pile to the shear modulus of the soil, pile length, pile load, shear modulus of the soil, Poisson's ratio of the soil and radius of the pile; the output parameter was the pile settlement. Finite element and integral equation method by Randolph and Wroth (1978) was used to obtain the desired output. The results indicate that the neural network can be used to model the settlement of pile foundations effectively.

3.5 Tunneling

In the design of a tunnel, ground surface settlement due to tunnel excavation is a major consideration as it can severely disrupt the function of nearby structures and utilities. Many empirical and semi-empirical formulae are available to predict ground surface settlement. However, the predictions are frequently inaccurate because they do not take into consideration all the relevant factors. Kim et al. (2001) utilized the capabilities of pattern recognition and memorization of ANN to solve the problem. From literature reviews, three categories of major factors affecting ground movements in tunneling are Tunnel geometries, Ground conditions, Excavation and support conditions. Each major factor incorporate a number of parameters with some parameters can be further divided into several detailed items which serve as inputs to the neural networks. As a result, a total of 47 nodes were used in input layer and 2 nodes were used in output layer to predict i and δ_{smax}. The training data is composed of '113' field results, which have been collected from Seoul subway sites. Twenty seven candidate models were developed, and a model with three hidden layers of 47 neurons each was found to give optimal result. This model was then verified on the real database. With further modification in the training, the average errors for i-values and δs_{max} can be reduced to approximately 0.6 and 0.2 %, respectively. To confirm the generality of a trained ANN, simple examples are undertaken with two different sets of 12 data in each, extracted from '113' original data. The first 12 data have relatively large values of maximum

surface settlements, while the second 12 data have relatively small values of maximum surface settlements. For both cases, the remaining '101' data were used to train the network. It is noted that the '101' data are not the same data in each other case. After separate training in both cases, the extracted two sets composed of 12 data in each are used for prediction using the corresponding ANNs to the data sets. Figures 3.2 and 3.3 show the predicted results for the maximum settlements using the first and second data set. Overall, it is shown that ANN model is able to predict with high confidence (less than about 16% on inference error) and the generality is guaranteed for further predictions. The shortcoming of the ANN model proposed by Kim et al. is that it requires a lot of input parameters in the input layer (47 nodes). Most of these parameters are detailed items which might not be readily available from the site. Hence, in this project, an attempt is made to develop ANN models which requires less input parameters but with high accuracy for the prediction of maximum surface settlement and trough width (i).

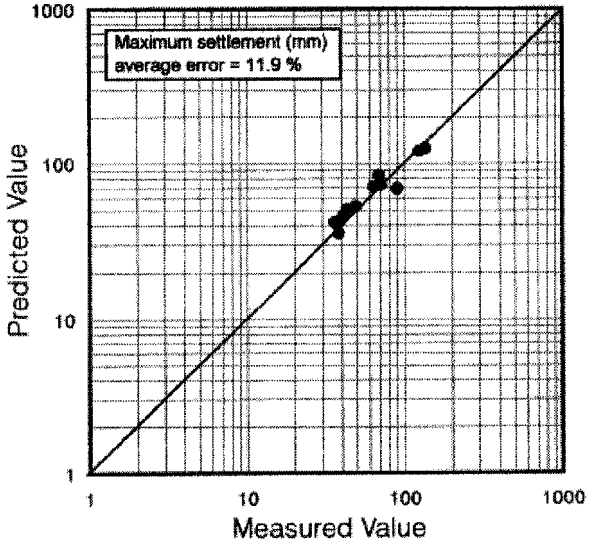


Figure 3.2: Results for untrained 12 data in the first data set (Kim et al. 2001)

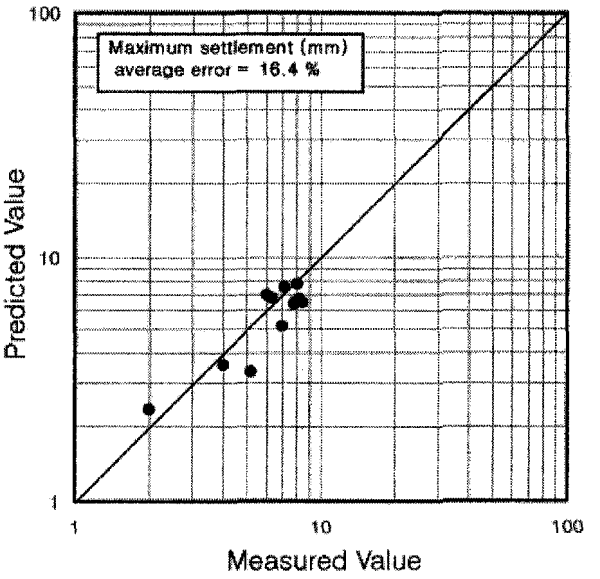


Figure 3.3: Results for untrained 12 data in the second data set (Kim et al. 2001)

Duan (2001) used a theoretical approach, empirical approach, numerical simulation approach, and artificial intelligence approach to study ground settlement caused by microtunneling. Microtunneling is a trenchless technology for construction of pipelines. Its process is a cyclic pipe jacking operation. In the numerical approach, a commercial finite element software FLAC3D was used to simulate the ground settlement caused by microtunneling. In the artificial intelligence approach, a three-layer back propagation neural network is developed to predict the ground settlement caused by microtunneling using the numerical simulation results. The results indicate that the neural network provides a means of rapid prediction of the surface ground settlement curve based on the soil parameters, project geometry and estimated ground loss. The predictions matched closely the results of FLAC3D over the full range of parameters studied and have a reasonable correspondence to the field results with which it was compared.

Another neural network model for predicting settlements during tunneling was proposed by Shi et al. (1998). The input layer consists of 11 nodes representing 11

major affecting factors while the output layer consists of 3 nodes representing three settlement parameters; maximum settlement at tunnel crown level, maximum settlement at tunnel inverted arch level, and maximum final settlement after tunnel excavation. The neural network model is trained and tested using the actual collected data from the 6.5 km Brasilia Tunnel in Brazil. The results show poor agreement between the predicted and measured values for test pattern with an average error of 70 mm (Figure 3.4).

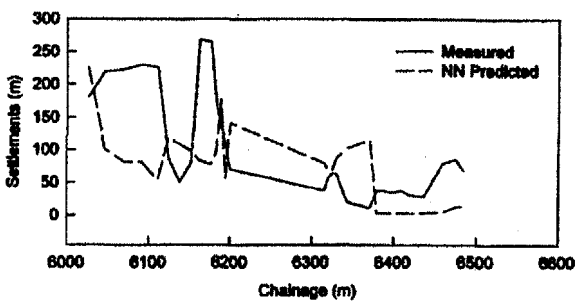


Figure 3.4: Results of Test Patterns for Maximum Settlements after Stabilization (Poor Agreement) (after Shi et al., 1998)

To improve the prediction accuracy, modular neural network models were used. A modular network consists of multiple NN modules, each of which only models one specific category of expertise. Each module is trained and tested separately using the data patterns in its category. All predictions from the trained modules are then combined into one report before it is deployed to the end user. By using modular NN models, the average prediction error is reduced to 33.4 mm; thus improving the agreement between the measured and predicted values for testing patterns (Figure 3.5). This shows that modular NN model has better performance than the general network model used in the beginning.

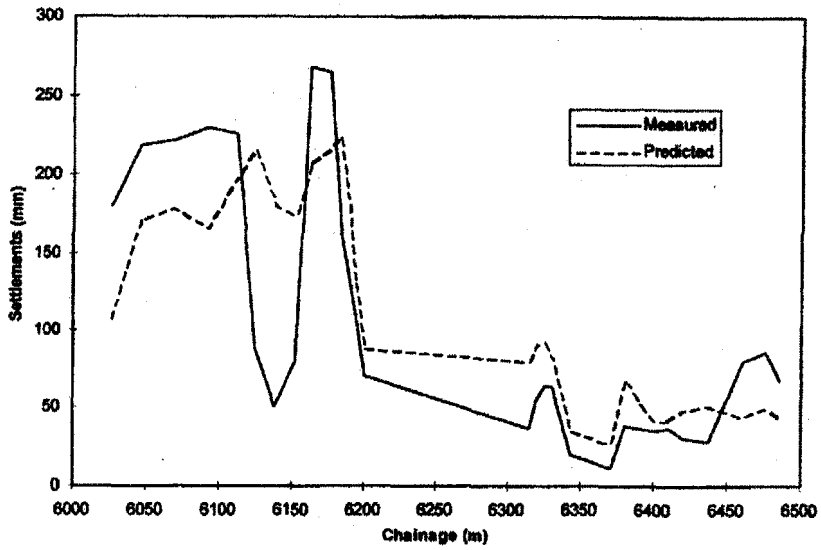


Figure 3.5: Test Results from Modular ANN Approach (after Shi et al., 1998)

3.6 Researches on Artificial Neural Network (ANN) at Nanyang Technological University

Various ANN applications, especially in geotechnical engineering, have been developed at Nanyang Technological University and some of them were mentioned in previous sections. The applications are Reliability assessment of serviceability performance of braced retaining walls using a neural network approach (Goh 2005), Neural network approach to model the limit state surface for reliability analysis (Goh 2003), Nonlinear modeling in geotechnical engineering using neural networks (Goh 1994), Seismic liquefaction potential assessed by neural network (Goh 1995), Pile driving records reanalyzed using neural network (Goh 1996), and Prediction of pile capacity using neural networks (Teh et al. 1997). In addition, a text book titled Parallel architectures for artificial neural networks: paradigms and implementations (Sundararajan and Saratchandran 1998) has been published and an ANN software NNGeo (Hefny 2000) has been developed. The research carried out for this thesis focused on the application of ANN to predict the ground surface settlement caused by tunneling works. Project of the same title has been carried out previously but with

lesser amount of data for training, testing, and validation sets. In this thesis's project the generalization ability of ANN has been improved greatly by adding more data from several tunneling projects in the training set. Furthermore, a special feature in this project include the Finite Element analysis for the development of reliable ANN model which can be used to predict the maximum surface settlement (S_{max}) and the trough width (i); whereas, the previous project concentrated only on the prediction of maximum surface settlement. The two parameters (S_{max} and i) are important in the assessment of the shape of a settlement trough. In this context, the produced neural network, which can predict both parameters, will be of great benefit to study the impact of the tunneling works on the ground surface settlement.

Chapter 4

Empirical methods for ground surface settlements

4.1 Introduction

Difficulty of predicting surface movements has always been the major problem in the study of ground subsidence induced by underground excavation. Many authors have tried to solve the problem using different approaches. Let's consider the simplest case of a circular tunnel in a homogeneous ground with horizontal surface (Figure 4.1).

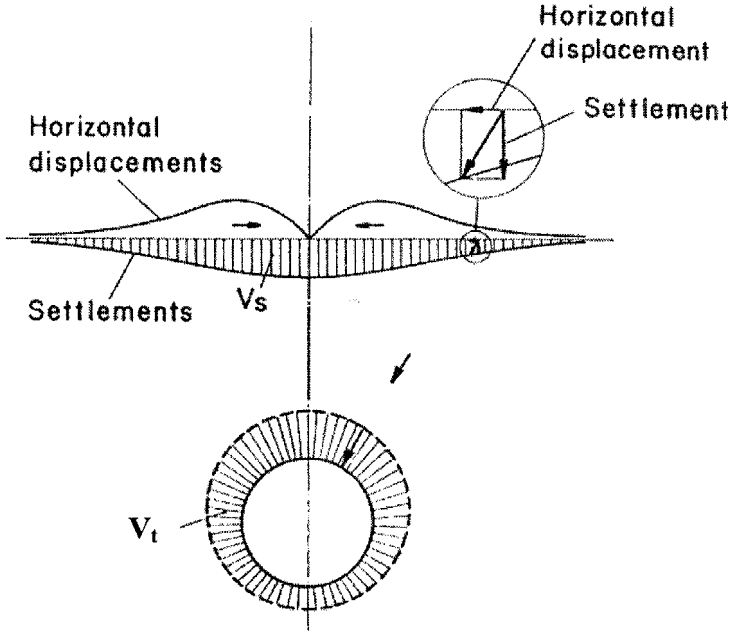


Figure 4.1: Soil movements and Ground loss (after Uriel and Sagaseta, 1989)

Opening of an underground excavation causes local relaxation of existing stress. As a result, the soil moves inward, more or less towards the center of the opening. The excavation line deforms to a new shape, which is smaller than the theoretical line. The

area enclosed between the theoretical and the deformed excavation lines is referred to as the 'ground loss', V_t (Uriel and Sagaseta 1989). The ground loss depends on many factors such as soil type, presence of water and method of construction. Ground surface settlements are attributed to this ground loss which occurs during tunnel construction. Ground loss indicates the volume of ground relaxing into the tunnel excavation, expressed in terms of unit distance advance of the excavation that causes the relaxation (i.e. cubic meters per meter advance). The final result of this loss is the formation of subsidence troughs.

4.2 Ground surface settlement profile

To date, the most popular method adopted in engineering practice to predict the surface settlement profile is the Gaussian distribution curve proposed by Peck (1969). Two main parameters used in the curve to determine the settlement trough are: the maximum surface settlement at the point above the tunnel centerline (δ_{smax} or S_{max}) and inflection point *i* (Figure 4.2).

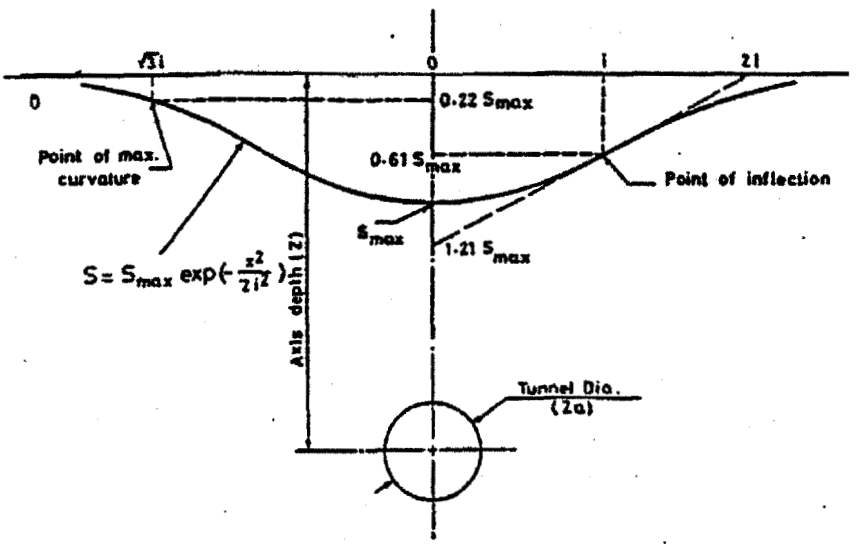


Figure 4.2: Surface settlement profiles (after Peck, 1969)

When the two parameters are available, the vertical settlement (displacement), δ_s, along the transverse line to tunnel axis can be calculated by (Attewell et al, 1986):

$$\delta_s = \delta_{s\max} \exp\left(-\frac{y^2}{2i^2}\right) \tag{4.1}$$

where y is the coordinate distance measure from the tunnel centerline. Equation (4.1) is a simplified form of Equation (4.2) in the case where x is much smaller than x_i.

$$\delta_s = \delta_{s\max} \exp\left[-\frac{y^2}{2i^2}\right]\left\{G\left(\frac{x-x_i}{i_x}\right) - G\left(\frac{x-x_f}{i_x}\right)\right\} \tag{4.2}$$

$$G(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{\alpha} \exp\left(\frac{u^2}{2}\right) du \tag{4.3}$$

where x is the distance between a point of concern and tunnel face, x_i is the initial position of tunnel face, x_f is the final position of tunnel face and i_x can be replaced by the inflection point, i. Maximum surface settlement is mainly governed by type of ground, size of tunnel, depth of tunnel, method of tunneling, and type of lining (Attewell 1977).

4.3 Estimation of total volume loss, V_t, and the surface settlement volume, V_s

In *cohesive soils*, the ground loss is expected to be between about 0.5 % and 2.5% of the tunnel face excavated area, depending upon the stiffness of the soil and the speed at which the initial support is installed (Attewell 1977). A more satisfactorily approach is to estimate total volume loss,V_t and surface settlement volume,V_s base on an 'Simple Overload Factor' defined as the ratio of overburden pressure σ_z minus any supporting internal pressure σ_i (e.g. compressed air pressure) to the undrained shear strength c_u under conditions in which the external, pre-existing stress field is uniform ($\sigma_z = \sigma_y = \sigma_x$) (Deere et al, 1969).

$$OFS = (\sigma_z - \sigma_i) / c_u \quad (4.4)$$

The overburden pressure is normally expressed as the product of bulk unit weight γ and depth of tunnel axis z_0. Figure 4.3 shows V_t and V_s as a function of the overload factor.

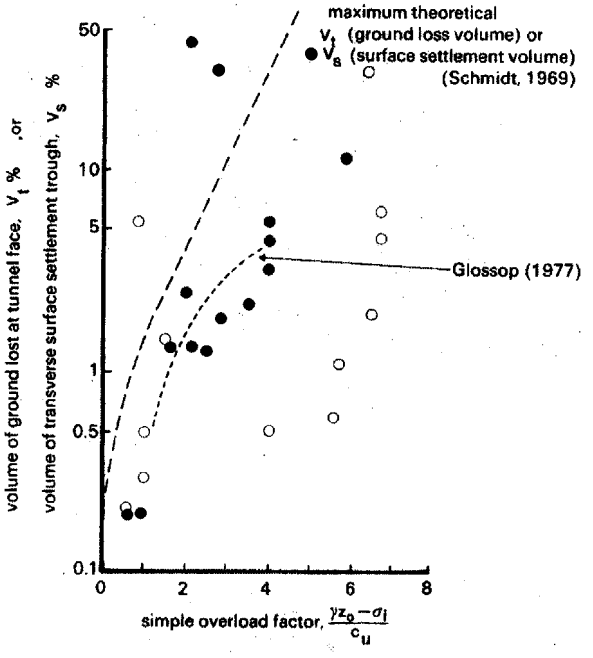


Figure 4.3: Estimation of ground losses and surface settlement volumes from the overload factor for tunnels in cohesive soils. Open circles are Schmidt's(1969) data from shield-driven tunnels. Solid circles are as presented in Attewell and Yeates (1984) (after Attewell et al, 1986)

For tunneling in most clay soils, Equation (4.5) will give reasonable estimate of V_s for $1.5 \leq OFS \leq 4$ (Glossop 1978):

$$V_s\% = 1.33 \times (OFS) - 1.4 \quad (4.5)$$

For tunneling in *granular soils* above a water table, approximate V_t will be of 2-5 percent. If tunneling is carried out below the water table, compressed air will be necessary to control the stability of the granular soil. In that case, the range of V_t will

be between 2 and 10 percent. For preliminary calculations, a 5 percent approximate could be adopted in both cases (Attewell et al, 1986).

If the settlements occur without any change in volume of the soil (undrained behaviour), then volume of the settlement trough along a unit length of tunnel having the same shape as the Gaussian distribution curve is given as (Attewell and Farmer, 1974):

$$V_s = 2.5\ i\ \delta_{smax} \tag{4.6}$$

Equation (4.6) allows maximum settlement δ_{max} to be estimated from surface settlement volume V_s, and inflection point i. The magnitude of V_s depends on the type and strength of the soil, the tunnel depth and the method and quality of construction. For most single tunnels in firm-to-stiff clay soils, the volume V_s of the settlement trough is approximately equal to the volume V_t of ground lost at the tunnel, whilst in granular non-cohesive soil, dilation may occur through arching above the tunnel crown if depth of tunnel axis is equal or greater than three times the excavated radius of a tunnel (Hansmire, 1975).

Attewell et al. (1986) concluded that Gauss-function settlement troughs was confirmed by the settlement data compiled from monitoring programs of soft ground tunnels around the world. However, when the magnitude of the maximum surface settlement exceeds about 0.5% of the tunnel axis depth, the edges of the trough do not expand in line with increasing settlement, and the profile deviates from the normal probability curve (Attewell 1977). The normal probability profile assumptions require that the width of the fully-developed transverse settlement trough is infinite. However, in practice, it can be approximated as $2\sqrt{2\pi}i$ (i.e. 5i)(Attewell et al, 1986).

To date, the most powerful tool to simulate tunneling is the finite element method. This is mainly due to the availability of powerful codes and rapid improvement in computer efficiency. Burd et al (2000) used finite element for a three-dimensional analysis of tunnel, soil and a building interaction. The study provides a method to estimate the extent of crack damage caused to masonry structures by nearby shallow tunneling.

be between 2 and 10 percent. For preliminary calculations, a 5 percent approximate could be adopted in both cases (Attewell et al, 1986).

If the settlements occur without any change in volume of the soil (undrained behaviour), then volume of the settlement trough along a unit length of tunnel having the same shape as the Gaussian distribution curve is given as (Attewell and Farmer, 1974):

$$V_s = 2.5\ i\ \delta_{smax} \tag{4.6}$$

Equation (4.6) allows maximum settlement δ_{max} to be estimated from surface settlement volume V_s, and inflection point i. The magnitude of V_s depends on the type and strength of the soil, the tunnel depth and the method and quality of construction. For most single tunnels in firm-to-stiff clay soils, the volume V_s of the settlement trough is approximately equal to the volume V_t of ground lost at the tunnel, whilst in granular non-cohesive soil, dilation may occur through arching above the tunnel crown if depth of tunnel axis is equal or greater than three times the excavated radius of a tunnel (Hansmire, 1975).

Attewell et al. (1986) concluded that Gauss-function settlement troughs was confirmed by the settlement data compiled from monitoring programs of soft ground tunnels around the world. However, when the magnitude of the maximum surface settlement exceeds about 0.5% of the tunnel axis depth, the edges of the trough do not expand in line with increasing settlement, and the profile deviates from the normal probability curve (Attewell 1977). The normal probability profile assumptions require that the width of the fully-developed transverse settlement trough is infinite. However, in practice, it can be approximated as $2\sqrt{2\pi}i$ (i.e. 5i)(Attewell et al, 1986).

To date, the most powerful tool to simulate tunneling is the finite element method. This is mainly due to the availability of powerful codes and rapid improvement in computer efficiency. Burd et al (2000) used finite element for a three-dimensional analysis of tunnel, soil and a building interaction. The study provides a method to estimate the extent of crack damage caused to masonry structures by nearby shallow tunneling.

However, the finite element method is complex and it requires dedicated computing facilities.

4.4 Settlement trough dimension parameter, *i*

Normally, the maximum ground surface settlement (δ_{smax}) can be measured directly, whilst it is very difficult to determine an inflection point *i* on a trough curve. Tunnel geometry and soil conditions are several factors which govern the location of inflection point (Kim 2001). A common way to determine inflection point is to take a distance x corresponding to settlement value of $0.61\delta_{smax}$ (Attewell et al, 1986). Cording and Hansmire (1975) proposed the 'best-fit method', where the inflection point is determined from the plot of recorded settlements in logarithm scale against the square of the transverse distance (log δ_s-x^2 diagram). By utilizing a linear-fitting curve, a point corresponding to settlement value of $0.61\delta_{smax}$ can be determined. Equation (4.6) can be used as well to obtain the inflection point. This is known as the 'volume method'. All three methods use Gaussian normal probability function as their basis. In cases where the settlement trough perfectly matches the Gaussian curve, the three methods will produce identical *i* value. In addition to the above methods, several empirical relations are available as well to approximate the parameter i (Peck 1969, O'Reilly and New 1982, Leach 1985, Atkinson and Potts 1976).

Based on the stochastic theory and empirical evidence, a simple expression for soil, clay and rock incorporating normalized factors (i.e. Z_o/D and i/R) is derived as follows (Attewell 1977):

$$i/R = K(Z_o/D)^n \qquad (4.7)$$

where K and n are constants depending on the ground conditions. R is the radius of the tunnel and Zo is the depth of the tunnel axis from the ground surface. For clayey soils, Attewell and Farmer (1977) suggested values of K=1.0 and n=1.0, while Clough and Schmidt (1981) proposed values of K=1.0 and n=0.8. Peck (1969) synthesized the

range of K and n-values with respect to certain ground conditions, as shown in Figure 4.4.

The following values of K and n suggested by Attewell (1982) will be used in Chapter 5 to calculate the surface settlement of the validation set by using empirical method:

1) K = 1.0 and n = 1.0 for clays
2) K = 0.63 and n = 0.97 for granular soils irrespective of water table

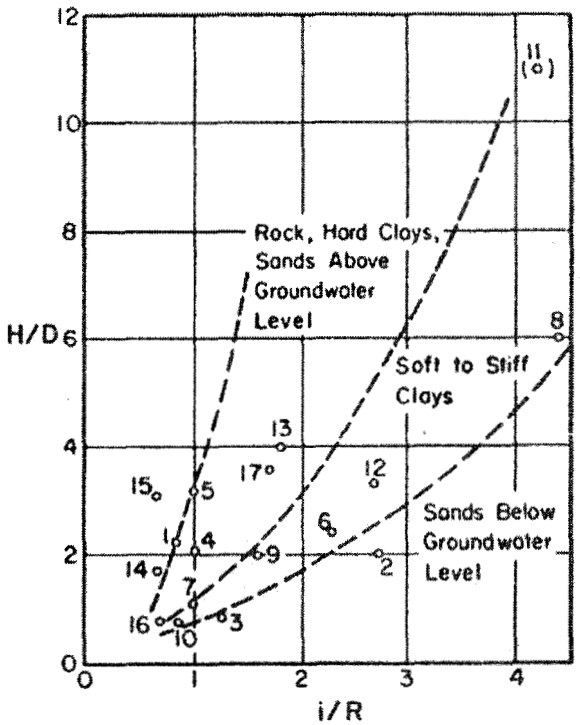


Figure 4.4: Peck's (1969) zoning for *i* as a function of tunnel depth and soil type.

From Equations (4.1), (4.6), and (4.7), the steps to estimate settlement for a real tunnel can be formulated as follows (Gunn (1992)):

(a) the engineer estimates a figure for ground loss on the basis of experience with similar tunneling techniques in similar soils.

(b) a value of *i*/D is assumed, based on Peck's chart (Figure 4.5) or Equation (4.7) or a similar relationship.

(c) Equation (4.6) is now used to find δ_{smax} and hence Equation (4.1) can now be used to predict the surface settlement at any point.

4.5 Settlement profile along tunnel drive

A significant ground surface settlement is usually induced during the tail void closure, which occurs immediately after the passage of the shield machine (Ou and Cherng ,1995). The settlement due to the tail void closure can be obtained from the settlement history curve, as shown in Figure 4.5. Line a-b denotes the tail void closure, while line b-c, which has a gentler slope, is considered as the consolidation stage.

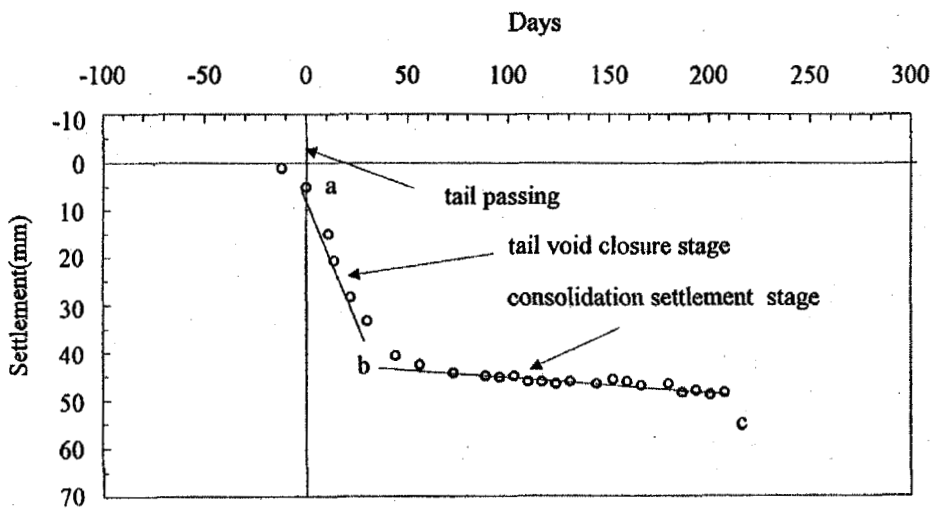


Figure 4.5: Typical settlement history due to shield tunneling (Ou et al. 1998)

According to Attewell et al. (1986), ground loss is a total summation of face loss, shield loss, pregrout loss, and postgrout loss. Face loss is associated with the axial loss at the tunnel face during the intrusion of the shield. Shield loss is radial loss over the shield through closure of a bead or other overcutting device. Pregrout loss is radial groud loss occurs behind the tailskin due to closure of an ungrouted space between lining and ground. Postgrout loss is radial loss occurs behind the tail of the shield due to closure of the grouted space (Figure 4.6). Attewell and Woodman (1982) developed equations to predict the three-dimensional ground movements and strains caused by tunneling based on the analysis of some case history data. They concluded that the soil movement ahead and at the sides of a tunnel face can be quantified by assuming that there is no change in the volume of soil at ground surface and ground movement is the sum total of elementary movements of assumed form resulting from increments of tunnel advance.

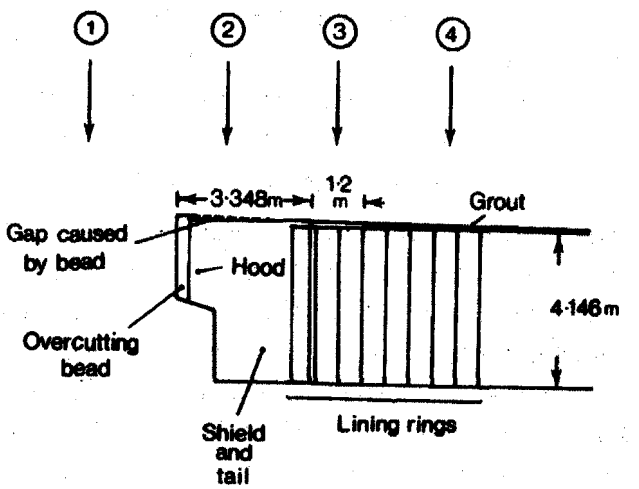


Figure 4.6: Major areas of ground loss around a shield-driven tunnel in soil: (1) intrusion at the face; (2) radial loss over the shield through closure of a bead or other overcutting device; (3) and (4) post-shield radial losses. (3) represents closure of an ungrouted space between lining and ground, (4) represents closure of the grouted space (Attewell and Farmer 1974).

Ground displacement caused by tunneling is a three-dimensional problem. Figure 4.7 shows, in a three dimensional view, how settlements develop as the tunnel heading advances. The soil in front of the face tends to move inwards producing settlements well ahead the heading. At a certain distance from it, the ground deformation does not occur yet; whilst at a certain distance behind it the ground displacements have reached their maximum values.

Along the direction of shield driving, the surface settlement profile may be divided into five different regions, as shown in Figure 4.8. The mechanisms of soil settlement induced by shield driving for each region are given in Table 4.1 (Sun 2000). According to the study by Woodman and Attewell (1982), a cumulative probability function can approximate fairly well the settlement profile along the tunnel axis.

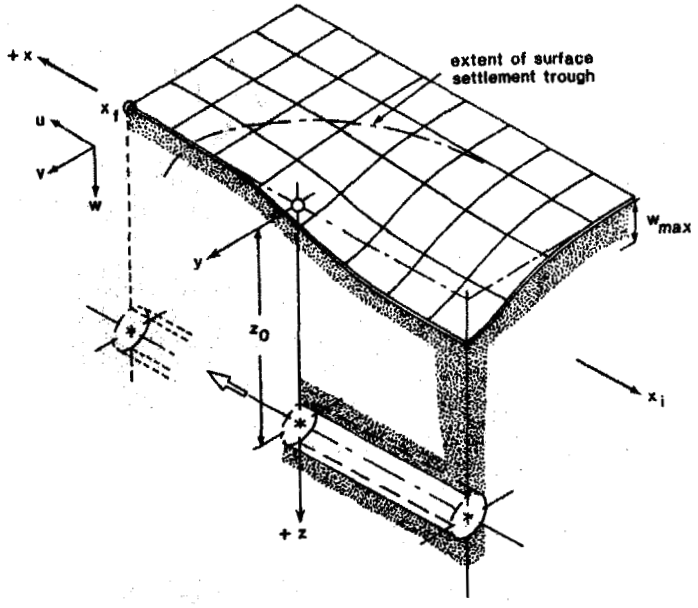


Figure 4.7: 3-D view of ground subsidence (from Attewell et al, 1986)

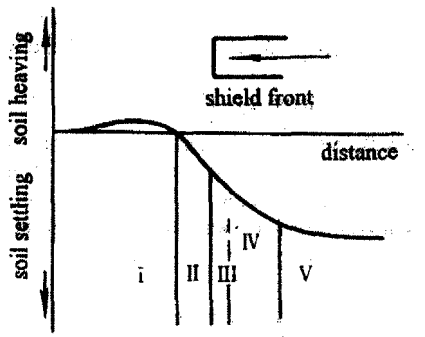


Figure 4.8: Different regions of ground surface displacement along direction of shield driving (Sun 2000)

Table 4.1: Mechanism of soil settlement induced by shield driving (modified from Sun 2000)

Region	Types of soil settlement	Cause/reason	Stress Disturbance
I	Soil heaving at a certain distance in front of shield working face	when, pi>pz+pw, soil heaving when, pi<pz+pw, soil settling pi-excerting slurry pressure in closed cabin pz- static external earth pressure pw- ground water pressure	increasing of pore-water pressure; increasing of total soil stress
II	Initial settlement	soil consolidated due to squeezing & compaction under shield driving	dissipating of pore water pressure; increasing of effective soil stress
III	Settlement just at the while shield passing through	soil disturbance during construction; shearing-dilatation between shell of shield & its surrounding soils; excess amount of soil excavation	stress release of soils
IV	Settlement caused by gap-space existing at shield tail	soil losses shield supporting, loss of earth strata caused by "constructional gap-space" at shield-tail; grouting behind tunnel lining segment not in time	stress release of soils
V	Settlement due to secondary consolidation of soils	time dependent viscous soil (delayed) deformation (follow-up soil creep)	stress relaxation of soils

Chapter 5

Details of Contract C705 Northeast Line, C825 Marina Line and C823 Circle Line

In the first main analysis presented in Chapter 6, the feasibility of using artificial neural networks to predict the maximum settlements due to tunneling is investigated. For this analysis, the case records used to develop and verify the ANN models were collected from contracts North East Line C705, Marina Line C825 and Circle Line C823. The details of the contracts and the relevant information of Singapore geological profile are presented in the following sections.

5.1 Geological Profile

The geological formations of Singapore encountered during tunneling and excavation works can be broadly grouped into 4 main types (Moss 2000):

1) Kallang formations. These are recent deposits of soft clays and fluvial sands, including the Singapore Marine Clay
2) Jurong Sedimentary. These are sedimentary series of sandstones, siltstones and mudstones with degrees of weathering varying from fresh rock through to residual soil
3) Old Alluvium. It mainly consists of medium dense to very dense clayey coarse sand and fine gravel.
4) Bukit Timah Granite. It is fresh rock at depth and completely weathered approaching the ground surface, producing silty clay.

5.2 Contract C705, Singapore North East Line

C705 is one of the eleven designs and construct contracts of the Northeast Line project, which dealt with the construction of bored running tunnels from Boon Keng Station to

Potong Pasir Station. The Northeast Line provides a predominantly underground mass transit railway from the World Trade Centre in the south to Punggol in the northeast of Singapore Island. The contract was awarded to a joint venture comprising Kumagai Gumi, Sembawang and Mitsui (KSM) in 1997. Babtie Tunnelling Division in the United Kingdom carried out the design for the tunneling works. Geology for C705 comprised mainly Old Alluvium overlain by soft deposits of the Kallang Formation.

The Old Alluvium is composed of very dense silty sands with layers of hard clays and silts, while the Kallang formation includes sands and soft to firm clays of marine/alluvial origin. Erosion in the past is responsible for the undulating surface of the Old alluvium. Due to removal of overburden, the Old Alluvium is considered to be overconsolidated, however, the marine clay of the Kallang formation is typically normally or very slightly overconsolidated. The twin bored tunnels were driven mainly within the Old Alluvium, while encounter with marine clay is experienced locally. A geological section along the northbound tunnel is shown in Figure 5.1.

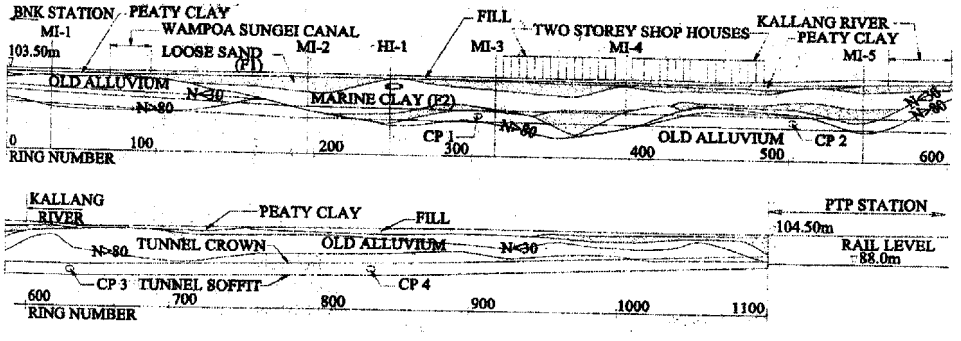


Figure 5.1: Geological section along North Bound Tunnel (South bound tunnel is similar) (Izumi et al. 2000)

5.2.1 Bored Tunnel Construction

The twin tunnels run approximately parallel to each other with the horizontal spacing between the tunnel axes varying from about 15 m to 24 m. The depth to the tunnel axis varies between 12 and 25 m below ground level. The internal diameter of both tunnels is 5800 mm and the thickness of the pre-cast segment lining is 250 mm. Two Earth Pressure Balance Shield machines (EPBM) with an external diameter of 6440 mm were used to construct the tunnels. To support the face, the EPBMs used a cutter pressure chamber filled with plasticised soil. During the excavation time, polymer was injected to lubricate the material and avoid the blocked nozzles. By balancing polymer against material removed, a continuous pressure was maintained. The EPBMs were equipped with three pressure gauges mounted on the front face of the machine and the gauge at center elevation is used to control the slurry pressure. Initially, one tail grouting device was used to inject grout (at ground receiving pressure plus around 0.5 to 2 bar) into the tail void as the tunnel shield advanced and the lining was installed. Later on, due to faulty ports, the grout was injected through the upper lining segment through a grout nozzle screwed into the lifting sockets. Grouting started immediately after the ring was installed and continued until the next ring was placed upon which the grout nozzle was moved forward. Figure 5.2 shows the EPBM used in C705, while its specification is given in Table 5.1. Figure 5.3 shows the typical settlement profile due to shield tunneling for contract C705.The graph displays the similar trend as the graph of settlement profile along tunnel drive in Figure 4.5. The ground surface settled immediately after the first passage of the shield machine followed by the consolidation settlement stage. There was a slight heave before the tail passing.

Table 5.1: Specifications of C705 Earth Pressure Balance Machine (Izumi et al., 2000)

Specification of Shield			
Out side diameter	6440mm		
Overall length	7500mm		
Propulsion force	3500tf		
Shield Jack speed	10.0 cm/min		
Shield Jack	24pcs 1471KN*1800stroke		
Specification of Cutter			
Cutter torque	4714 kNm		
Rotational speed	0.9 rpm		
Copy cutter jack	2pcs 161.8KN*150 stroke		

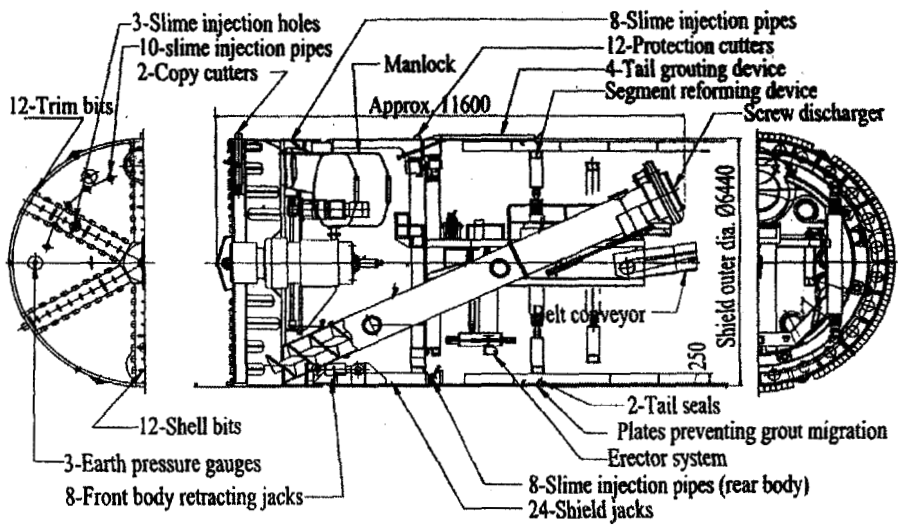


Figure 5.2: C705 Earth Pressure Balance Machine (Izumi et al. 2000)

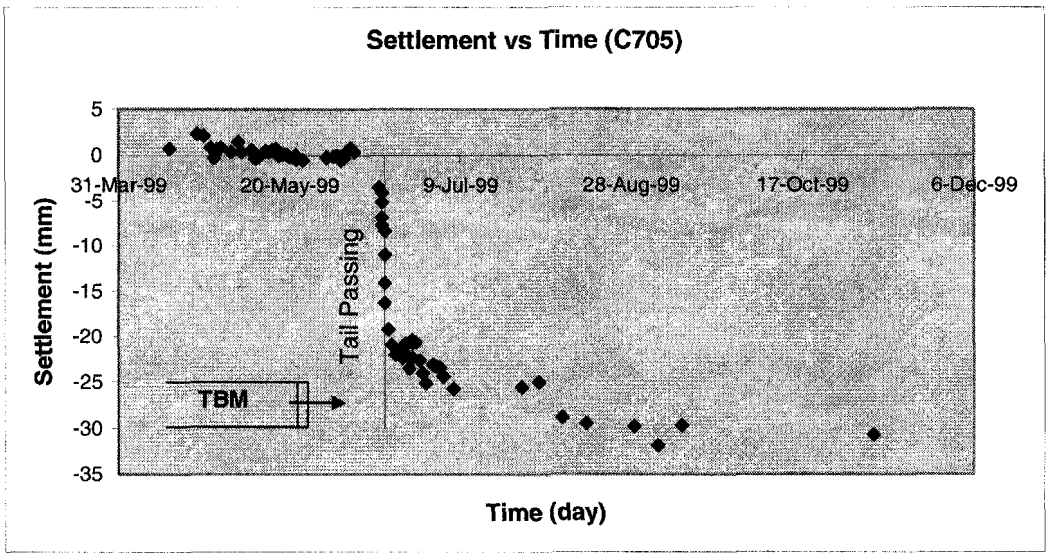


Figure 5.3: Typical settlement profile due to shield tunneling for contract C705

5.3 Contract C825, Marina Line

The Marina Line (MRL) is part of the Land Transport Authority's ongoing effort at building a comprehensive rapid transit system network. The MRL serves as the first stage of the Circle Line (CCL), an orbital line that will link all MRT lines which run into the city (East West Line, Nort South Line, North East Line). The Marina Line comprises six stations and runs underground from Dhoby Ghaut Interchange Station on the North East Line to end at Boulevard Station at Stadium Boulevard. The project of the MRL is divided into two contracts C824 and C825 with C824 awarded to the JV Nishimatsu-Lam Chum and C825 to the JV WoHup-Shanghai Tunnel Engineering. Both projects are currently under construction.

Contract C825 consists of 4 stations and tunnels in between (Figure 5.4). They are Dhoby Ghaut Station (DBG), Museum Station (MSM), CVC Station, and Millenia Station (MLN). All the four stations are built using the top-down construction method. This method was opted in order to reduce soil movements and ground water lowering. The length of tunnels constructed using TBM is approximately 1400m. The soil

conditions vary considerably along the project from soft Marine Clay at Millenia Station, Old Alluvium at CVC Station, Fort Canning Boulder Bed at Museum station to Jurong Formation at Dhoby Ghaut. From previous tunneling projects within Singapore, Old Alluvium has proved to be an ideal tunneling medium in terms of settlement as volume losses of less than 1% can be attained easily. Old alluvium is a Pleistocene deposit of medium dense to very dense clayey coarse sand and fine gravel, containing lenses of silt and clay. It originated from weathered granite and thus contains a high proportion of slightly rounded quartz.

The TBM Herrenknecht machines for the bored tunnels are operating in closed mode according to the EPB method. The TBM starts from a temporary shaft at the east end of the project (Figure 5.4). Initially, the tunnels were arranged vertically one above the other to minimize impact on piled foundations. The lower tunnel was driven ahead of the upper tunnel. This arrangement was changed from vertical to become horizontal (side by side) prior to reaching CVC Station. The TBMs are driven from the first shaft through CVC station before station excavation to be picked up at a temporary shaft at the eastern end of Museum Station. The TBMs are then transported to Dhoby Ghaut station and driven in opposite direction to the western end of Musem Station (Osborne et al., 2004). The specifications for the C825 EPBM are shown in Table 5.2.

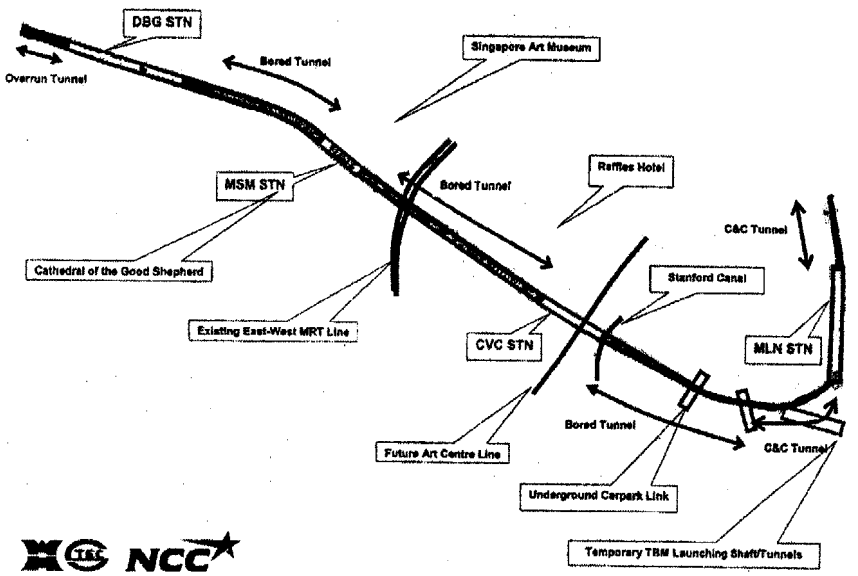


Figure 5.4: C825 comprising 4 Stations and Tunnels (Osborne et al. 2004)

Table 5.2: Specifications of C825 Earth Pressure Balance Machine (Osborne et al., 2004)

EPBM Specifications	
TBM Manufacturer	Herrenknecht
Outer Diameter (mm)	6580
Overall TBM Length (mm)	7400
Overall Length (inc Back-Up) (m)	55
Cutterhead Drive	Hydraulic
Power (kW)	945
Cutterhead Motors	8
Torque (rpm)	2.7
Face Injection Ports (No.)	8
Bulkhead Ports (No.)	4
Total Cutter Picks (No.)	72
Total Discs (No.)	40
Copy Cutters (No.)	2
Overcut amount (mm)	30
Thrust Rams (No.)	16
Stroke (mm)	2200
Equipped Propulsion (kN)	42575
Articulation Rams (No.)	14
Total capacity (kN)	73892

5.4 Contract C823, Circle Line

Contract C823 is part of Circle Line project Stage 2. This contract was awarded to the JV Nishimatsu-Lam Chum. The contract comprises 3 stations and bored tunnels in between. The stations are Old Airport Road Station, Tanjong Katong Station, and Paya Lebar Station. In this project, the settlement readings from this contract were obtained from level instruments along the tunnel line between Old Airport Road Station and Tanjong Katong Station. The soil investigation on site was carried out by Soil & Foundation (Pte). Based on the borehole data and in-situ tests results obtained from site, the underlying subsoils can be broadly classified into the following layers: Fill, Kallang Formation, and Old Alluvium. The fill layer consists typically of clayey sand, sandy silt and clayey silt mixed with foreign debris such as woods and concrete. The Kallang Formation was found beneath the fill layer in most boreholes. It comprises peaty Clay, fluvial sand, fluvial clay, and marine clay. The Old Alluvium occurred as basement subsoils in this area. Tunneling was carried out mostly within clay sediments such as marine clay, estuarine, and fluvial clay. The estuarine and fluvial clay were found to be mainly underlying the marine clay or as the intermediate layer within the two marine member. Marine clay is soft compressible normally consolidated clay that can contain pockets and lenses of silt and sand, but behaves in an undrained manner during tunneling. A typical near surface strength is 22 kPa with minor strength gain as the strata goes deeper. Therefore, positive support pressure from the EPBM needs to be maintained at all times. The old alluvium with SPT less than 30 was encountered for a short distance near the Old Airport Road station. Hitachi Zosen EPBMs were used to construct the bored tunnels on C823. The specifications for the EPBM are shown in Table 5.3. Figure 5.5 shows the map of Circle Line project including C825 and C823.

Table 5.3: Specifications of C823 Earth Pressure Balance Machine (Osborne et al., 2004)

EPBM Specifications	
TBM Manufacturer	Hitachi Zosen
Outer Diameter (mm)	6630
Overall TBM Length (mm)	8615
Overall Length (inc Back-Up) (m)	51
Cutterhead Drive	Electric
Power (kW)	660
Cutterhead Motors	12
Torque (rpm)	1.06
Face Injection Ports (No.)	6
Bulkhead Ports (No.)	4
Total Cutter Picks (No.)	92 bits
Total Discs (No.)	0
Copy Cutters (No.)	2
Overcut amount (mm)	150
Thrust Rams (No.)	26
Stroke (mm)	2050
Equipped Propulsion (kN)	39000
Articulation Rams (No.)	12
Total capacity (kN)	24000

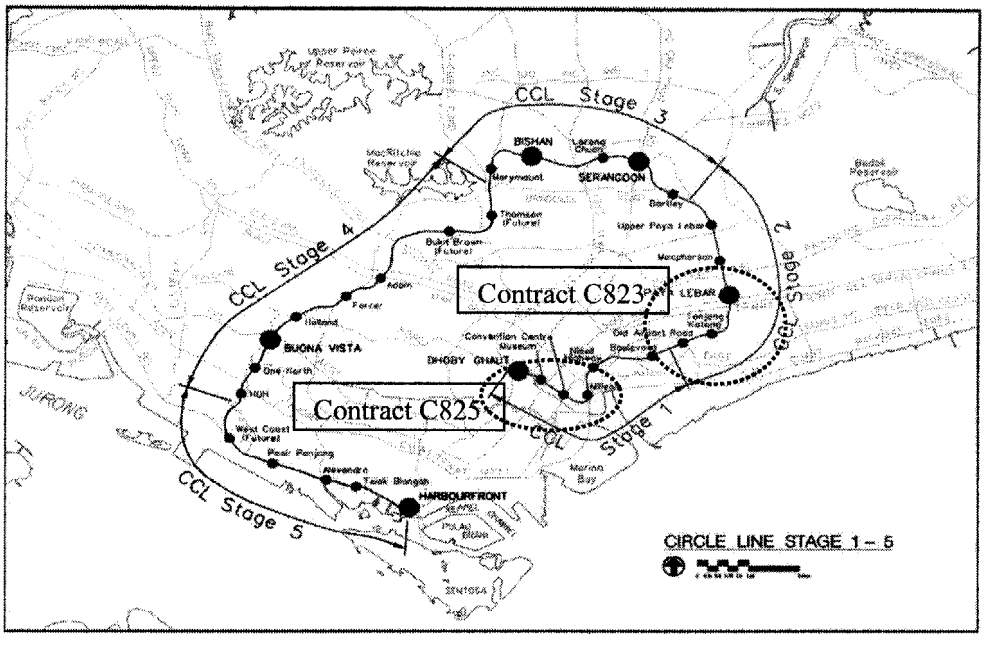


Figure 5.5: Map of Circle Line including project C825 and C823 (from LTA)

Chapter 6

Prediction of Maximum Surface Settlement by Multi-layer Perceptrons (MLP)

The settlement trough, as described by Gaussian distribution, is determined by two main parameters: the maximum surface settlement at the point above the tunnel centerline (S_{max}) and the width parameter (i) defined as the distance from the tunnel centerline to the inflection point of the trough. Due to insufficient information of width parameters from the field, this chapter focused only on the development of ANN models to predict maximum surface settlements induced by single tunneling. The most commonly used neural network type is Feed-forward multi-layer perceptrons (MLPs) trained with the back-propagation algorithm, as they have a high capability of data mapping. MLPs trained with back-propagation algorithm have been used successfully in many geotechnical engineering problems; hence they are used in this research.

6.1 Comparison of softwares

Several software systems for neural network operation are tested and the results are compared. The best software is then used for network analysis in this project. This section presents a comparison of performance of available back propagation neural network softwares , namely NNGeo, Pittnet , and MatLab. Results and details of the tests using NNGeo and Pittnet were obtained from the report by Robin (2001). NNGeo was developed by Hefny (2000) and applicable in Windows platform. The program code was written in Visual Basic. It can read the input data from both, normal text files and Microsoft Excel files and output of the training can be written to these two file formats as well. Pittnet was developed by Smith (1997) at the University of Pittsburgh. This computer program works in a DOS environment allowing the user to construct, train and test different types of artificial neural networks. Matlab is a technical computing software developed by The MathWorks, Inc. It has been used by technical people all over the world for their engineering and

scientific work. The neural network toolbox contained in Matlab is used for the network analysis. The three softwares are used to predict the friction capacity of driven piles. The data was obtained from load test records compiled by Goh (1995). A total of 45 patterns were used to train the neural network and 20 patterns to test the predictive ability of the neural network. Summary of the training and testing data are shown in Appendix A.

For NNGeo and MatLab, the learning rate and momentum used were 0.2 and 0.9, respectively. For Pittnet, the learning rate was 0.9 while the momentum was set at a default value of zero since it could not be changed by the user. For all three programs, the number of training iterations used was 30000 and the number of hidden neurons was three. The accuracy of the three networks was compared with that obtained by Goh (1995). Table 6.1 shows the comparison of prediction accuracy of the different networks. It can be seen that MatLab showed the highest correlation coefficient and lowest error rate for testing set compared to other networks, whereas Pittnet displayed lowest correlation coefficient and highest error rate. Hence, MatLab is used for the subsequent analysis of neural network in order to obtain more accurate settlement prediction.

Table 6.1: Comparison of prediction accuracy of different back propagation networks using pile driving data from Goh (1995)

Back propagation Network	Neural network inputs	No. of hidden neurons	Correlation Coefficient Testing	Error Rate Testing
NNGeo	S_u , σ'_v	3	0.962	4.56
Goh	S_u , σ'_v	3	0.963	4.72
Pittnet	S_u , σ'_v	3	0.959	5.53
MatLab	S_u , σ'_v	3	**0.972**	**3.97**

6.2 Preliminary analysis to obtain optimum ANN Model

• Model Inputs and Outputs

Eleven parameters are accepted to have significant impact on the settlement of ground surface due to tunneling, and are thus tested as the ANN model inputs. The model output is the initial maximum surface settlement immediately after the tail passing. The input and output parameters used for the network analysis are listed in Table 6.2.

Table 6.2: Input and Output parameters for ANN analysis

	Input Parameters	Symbol	Units
1	Cover	H	m
2	Earth Pressure	EP	kPa
3	Advance Rate	AR	mm/min
4	Mean SPT above Crown level	S1	N300
5	Mean SPT at Springline level	S2	N300
6	SPT at Invert level	S3	N300
7	Mean bulk density above crown level	BD	kN/m^3
8	Mean stiffness around tunnel circumference	E	Mpa
9	Moisture Content	MC	%
10	Ground Water Level	GWL	m
11	Grout Pressure	GP	kPa
	Output Parameter	**Symbol**	**Units**
1	Initial Maximum Surface Settlement	S	mm

1. Soil Cover

Soil cover is measured from tunnel crown to ground surface level. Since Equation 4.7 uses depth of tunnel axis from ground surface (z_o) as a parameter, soil cover is obviously an important factor to estimate surface settlement.

2. Advance Rate

Attewell et al (1986) used the average rate of tunnel advance to calculate three components of ground loss namely face loss, radial loss, and postshield lost. This signifies high correlation between advance rate and final ground surface settlement.

3. Earth Pressure

As discussed before, EPBM provides the ability to maintain full support to the excavation face by applying earth or face pressure. Excessive pressure on the tunnel-excavated face reduces the life span of the TBM's cutter head while insufficient pressure will cause ground surface settlement.

4. SPT 1, SPT 2, SPT 3

SPT "N" values describe geotechnical properties of the soils where tunnels are excavated. They are included in the log of boring produced from site-specific soil investigation and testing. Three types of SPT value are used in this project, namely SPT 1, SPT 2, and SPT 3. SPT 1 is the mean SPT N value of soil layers above the tunnel crown up to ground surface. SPT 2 is obtained by taking the average of the SPT values at the crown, middle and invert levels, as shown in Figure 6.1. SPT 3 is the SPT N value at tunnel inverted level.

5. Bulk Density, Stiffness

Soil stiffness, E, soil density, and coefficient of earth pressure at rest, K_o, are the key parameters for the design of tunnel lining. The LTA Design Criteria used K_o =

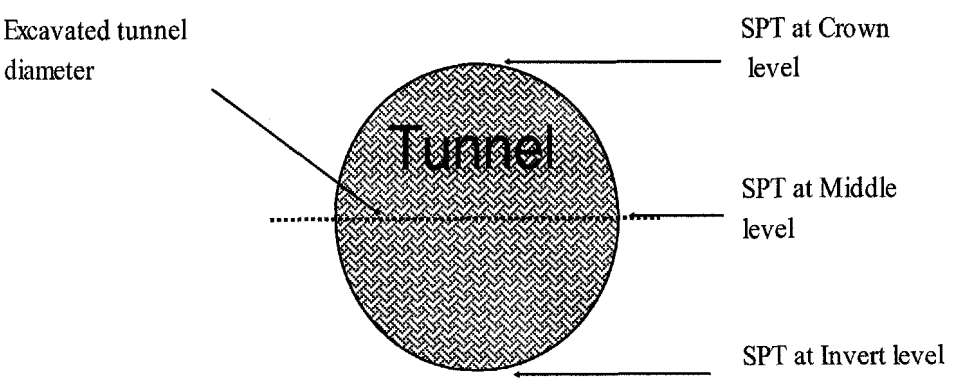


Figure 6.1: SPT 2 – The average of the SPT values at the crown, middle and invert levels

1.0, considered to be conservative for the Old Alluvium, Marine Clay and Kallang soils. The values of bulk density and stiffness adopted in this project are given in Table 6.3. The parameters were used by Babtie Tunnelling Division for the tunnel design of contract C705 North East Line.

Table 6.3: Geotechnical Parameters adopted for C705 Tunnel design

Soil		SPT "N" Value	Bulk Density (kN/m^3)	Stiffness MPa
Marine Clay	M		15.5	$0.3C_u$
Fluvial Sand	F1		18.0	20
Old Alluvium	O	N<25	19.0	1.5N
		N 25-80	20.0	1.5N
		N>80	21.0	120

C_u = Undrained Cohesion kPa

6. Moisture Content, Ground Water Level

Presence of water in the soil is indeed a factor that has significant effect on the magnitude of ground loss (Attewell et al (1986), Uriel and Sagaseta (1989)). Moisture content and ground water level gauge the amount of water present in the soil excavated. For neural network analysis, the moisture content of soil layer, driven through by the tunnels, is used as an input. The ground water table is measured from the ground surface level to the water table.

7. Grout Pressure

As the tunnel shield advanced, the space around the exterior of the tunnel was immediately filled by the grout, which was injected at ground receiving pressure into the tail void. Since postgrout loss occurs due to closure of the grouted space, the grout pressure used for the injection will have an impact on the potential ground loss. Hence, it is included as one of the input factors.

• Data Division

Neural network models were developed using different input combinations and the results were compared to obtain the network model with optimum input combinations. A total of 158 patterns are used to develop the models. The data are divided into three sets: training, testing and validation, as is the standard practice in the development of ANN models in geotechnical engineering. 90 patterns are used for training set, 39 patterns for testing set and 29 patterns for validation set. The three sets are divided in such a way that they are statistically consistent and thus represent the same statistical population. In order to achieve this, several random combinations of the training, testing and validation sets are tested until three statistically consistent data sets are obtained. The data and statistics of the training, testing and validation sets are shown in Appendix B. The statistical parameters considered are mean, standard deviation, minimum, maximum and range.

• Pre-processing of Data

To increase the efficiency of neural network training, certain preprocessing steps are performed on the network inputs and targets. Before training, the input and output variables are scaled to eliminate their dimension and to ensure that all variables receive equal attention during training. Scaling has to be commensurate with the limits of the transfer functions used in the hidden and output layers. In this analysis, input and output variables are scaled so that they fall in the range -1.0 to 1.0. This is because tan-sigmoid transfer function is used for both hidden and output layers for all models. For each variable x with minimum and maximum values of x_{min} and x_{max}, respectively, the scaled value x_n is calculated as follows:

$$x_n = 2\times(x-x_{min})/(x_{max}-x_{min}) - 1 \qquad (6.1)$$

6.3 Model Architecture and Stopping Criteria

A total of thirty two network models with different input combinations were tested to obtain the optimum one. All the models used one hidden layer with eight neurons. Prior to this, a preliminary test was conducted whereby the network of eleven input parameters is trained using different number of hidden neurons. The network obtained the lowest testing error when the number of hidden neuron is eight, as shown in Table 6.4. Hence, eight hidden neuron is considered optimal and is used for the remaining thirty one networks. This initial step is necessary to avoid under fitting which may occur when small number of hidden neurons is used for the network with many input parameters. On the other hand, the optimum number of hidden neuron from the network of maximum number of input parameter can be used for the network with smaller number of input parameters. In this case, over fitting is rarely a problem since early-stopping is used as the stopping criteria. The learning rate and momentum used were 0.2 and 0.9 respectively. The performance of each model is quantified from two measures; the coefficient of correlation (r) and error rate. The expression for error rate is given as follows:

$$\text{Error rate} = \Sigma\, \text{error}_i/N_p \quad \text{for } i = 1,2,\ldots N_p \tag{6.2}$$

$$\text{error}_i = |(T_i - O_i)| \tag{6.3}$$

where error_i is the error of the ith pattern, T_i is the target output value, O_i is the predicted output value and N_p is the number of patterns.

Table 6.4: Performance of ANN model NN1 with different hidden layer nodes

Model	No. of Hidden Neurodes	No. of input units	Correlation Coefficient			Error Rate (unscaled)		
			Training	Testing	Validation	Training	Testing	Validation
NN1	2	11	0.849	0.912	0.933	7.114	6.238	5.150
	4		0.906	0.892	0.913	5.608	5.935	5.360
	6		0.912	0.917	0.935	5.385	5.704	5.510
	8		0.945	0.936	0.916	5.158	5.226	6.457
	10		0.916	0.914	0.881	5.274	5.602	8.344
	12		0.908	0.871	0.928	5.758	7.077	5.932
	20		0.884	0.887	0.891	6.598	6.440	7.138
	30		0.950	0.836	0.929	4.933	7.522	5.685
	40		0.842	0.894	0.827	7.144	6.517	8.771

The performance results of the 32 models are shown in Table 6.5. Four network models NN6, NN7, NN16, and NN26 were selected for further analysis as they have high coefficient of correlation, relatively low error rates, and consistent performance on the training, testing, and validation sets. The four models were retrained with different number of hidden layer nodes. The maximum number of hidden layer nodes tested was 12 as the aim is to obtain a model, which provides satisfactory performance coupled with a small number of hidden nodes. The performance results of the four models with different hidden layer nodes are shown in Table 6.6 to 6.9. It can be seen that ANN model NN6 with 8 hidden neurons performs well, as it showed high coefficient of correlation, low error rate on the testing set and validation set, and consistent performance on the training, testing, and validation sets. Thus, from here onwards, the input variables considered for the ANN model analysis are cover (H), earth pressure (EP), advance rate (AR), average SPT above tunnel Crown level (SPT1), average SPT at tunnel Springline level (SPT2), stiffness (E), moisture content (MC) and grout pressure (GP).

Table 6.5: Comparison of Neural Network Models of different input combinations

Model	No. of input units	No. of hidden neurons	Neural Network Input Parameters	Coefficient of correlation			Error rate		
				Training	Testing	Validation	Training	Testing	Validation
NN1	11	8	H, EP, AR, S1, S2, S3, BD, E, MC, GWL,GP	0.89	0.92	0.88	6.46	5.50	7.33
NN2	10	8	H, EP, AR, S1, S2, S3, BD, E, MC, GWL	0.83	0.89	0.91	7.72	6.22	5.97
NN3	9	8	H, EP, AR, S1, S2, S3, BD, E, MC,	0.94	0.90	0.85	5.32	6.55	6.47
NN4	9	8	H, EP, AR, S1, S2, BD, E, MC,GP	0.93	0.89	0.86	4.53	6.53	6.76
NN5	8	8	EP, AR, S1, S2, S3, BD, E, MC	0.99	0.94	0.86	2.53	4.63	8.84
NN6	8	8	H, EP, AR, S1, S2, E, MC,GP	0.91	0.95	0.92	5.37	4.38	4.77
NN7	8	8	H, EP, AR, S1, S2, GWL, E, MC	0.92	0.93	0.91	6.16	5.19	6.07
NN8	8	8	H, EP, AR, S1, BD, E, MC,GP	0.83	0.88	0.91	7.57	6.62	8.11
NN9	7	8	EP, AR, S1,S3, BD, E, MC	0.93	0.91	0.74	5.00	5.77	8.52
NN10	7	8	MC,GWL,E,EP,H,S1,S3	0.95	0.87	0.78	3.66	6.51	9.21
NN11	7	8	AR,EP,S1,S3,MC,E,GP	0.94	0.91	0.73	5.39	6.39	9.08
NN12	7	8	AR,EP,S1,S3,MC,GWL,GP	0.96	0.89	0.89	4.48	6.59	7.46
NN13	6	8	EP, S1, S3, BD, E, MC	0.93	0.75	0.77	5.11	7.57	11.00
NN14	6	8	H, S1, S2, BD, E, MC	0.88	0.91	0.86	6.62	5.95	7.89
NN15	6	8	AR, S1, S2, GP, E, MC	0.88	0.83	0.92	6.38	7.77	5.56
NN16	6	8	AR, S1, E, GP, MC,BD	0.93	0.95	0.88	4.83	4.66	7.10
NN17	6	8	H, S2, BD, MC, GWL, GP	0.84	0.88	0.92	7.07	7.32	6.84
NN18	6	8	H,EP,S1,S3,MC,E	0.80	0.87	0.79	7.65	6.76	10.37
NN19	5	8	EP, S1, BD, E, MC	0.95	0.77	0.68	3.57	7.63	10.92
NN20	5	8	H,EP,S1,S3,MC	0.83	0.83	0.88	7.07	7.59	6.21
NN21	5	8	S1, S2, GP, E, MC	0.75	0.85	0.81	8.15	6.95	7.47
NN22	5	8	AR, S1, S2, E, MC	0.87	0.91	0.86	5.79	5.65	7.62
NN23	5	8	AR, S1, S2, BD, E	0.80	0.85	0.83	8.07	6.99	8.78
NN24	5	8	EP, AR, S1, S2, MC	0.85	0.83	0.88	7.14	7.83	7.92
NN25	4	8	EP, S1,BD,MC	0.75	0.48	0.67	8.21	10.57	9.74
NN26	4	8	H,S1,MC,GWL	0.94	0.93	0.87	4.69	5.23	5.59
NN27	4	8	S1,S3,MC,E	0.92	0.91	0.76	5.50	5.51	9.14
NN28	4	8	EP, AR, S1, S2	0.80	0.87	0.88	8.08	7.32	7.57
NN29	4	8	H,S1,S2,MC	0.93	0.91	0.88	5.69	5.93	7.49
NN30	3	8	BD, EP, S1	0.80	0.62	0.63	8.88	9.91	10.92
NN31	3	8	BD, MC, EP	0.80	0.79	0.80	7.72	7.40	8.94
NN32	3	8	H,S1,S2	0.86	0.86	0.91	6.83	6.88	4.88

Table 6.6: Performance of ANN model NN6 with different hidden layer nodes

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN6A	2	0.82	0.85	0.75	7.20	7.46	9.91
NN6B	4	0.90	0.92	0.91	5.84	5.53	6.40
NN6C	6	0.89	0.89	0.84	5.55	5.99	7.67
NN6D	8	0.91	0.95	0.92	5.37	4.38	4.77
NN6E	10	0.92	0.90	0.78	5.78	6.26	9.56
NN6F	12	0.84	0.90	0.89	7.20	6.02	9.37

Table 6.7: Performance of ANN model NN7 with different hidden layer nodes

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN7A	2	0.83	0.85	0.86	7.65	7.46	8.21
NN7B	4	0.89	0.94	0.90	6.73	4.40	5.74
NN7C	6	0.91	0.89	0.80	5.68	6.39	9.78
NN7D	8	0.92	0.93	0.91	6.16	5.19	6.07
NN7E	10	0.86	0.95	0.89	6.97	4.78	5.48
NN7F	12	0.90	0.93	0.91	6.46	5.60	6.08

Table 6.8: Performance of ANN model NN16 with different hidden layer nodes

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN16A	2	0.86	0.93	0.89	6.67	4.74	6.13
NN16B	4	0.87	0.93	0.87	6.36	5.11	5.59
NN16C	6	0.97	0.94	0.80	3.43	4.59	9.04
NN16D	8	0.93	0.95	0.88	4.83	4.66	7.10
NN16E	10	0.90	0.86	0.90	5.57	6.20	6.16
NN16F	12	0.90	0.88	0.91	5.84	5.48	5.30

Table 6.9: Performance of ANN model NN26 with different hidden layer nodes

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN26A	2	0.87	0.91	0.81	6.70	5.77	9.02
NN26B	4	0.94	0.91	0.83	5.59	6.67	10.10
NN26C	6	0.84	0.88	0.87	6.97	6.83	7.35
NN26D	8	0.94	0.93	0.87	4.69	5.23	5.59
NN26E	10	0.95	0.90	0.73	4.14	5.78	7.44
NN26F	12	0.91	0.92	0.88	5.99	5.95	6.73

6.4 Optimum ANN model

As discussed previously, a common way to obtain the optimum network architectures is by using a trial-and-error approach. A network with one hidden layer can approximate any continuous function, provided that sufficient connection weights are used (Cybenko 1989). Thus, one hidden layer is used in the analysis. The steps adopted for finding the best network architectures can be summarized as follows:

1. A number of trials are carried out with one hidden layer and varying number of hidden layer nodes.
2. The network that performs best with respect to the testing set is then retrained with different combinations of momentum terms, learning rates and transfer functions in order to improve model performance.
3. The model that has the optimum momentum term, learning rate and transfer functions is retrained several times with different initial weights until no further improvement occurs.

The third step is necessary to avoid getting trapped in a local minimum. Since back-propagation algorithm with gradient descent method is used to adjust the connection weights, problems with local minimum may be encountered if the initial starting point in weight space is unfavorable. Step 1 has been carried out in previous section which yield ANN model NN6 with 8 hidden layer nodes as the optimal model. The effect of momentum, learning rate and transfer function on model performance is summarized in Table 6.10. Figure 6.2 shows graphically the effect of momentum term on model performance. It can be seen that momentum does not affect much the performance of NN6 model in the range 0.01 to 0.7. The best prediction was obtained with a momentum value of 0.9. The effect of different learning rates on model performance is shown graphically in Figure 6.3. It can be seen that the error rate is the smallest at learning rate of 0.2. Table 6.10 shows that the model performs well when both the hidden layer and output layer use tan-sigmoid (tansig) transfer

function. The model produces poor results for logsig-logsig, logsig-tansig, and tansig-logsig. Other two combinations (tansig-linear and logsig-linear) produced satisfactory coefficients of correlation and error rates. However, the results are not as optimal as those when tansig-tansig combination was used for hidden and output layer.

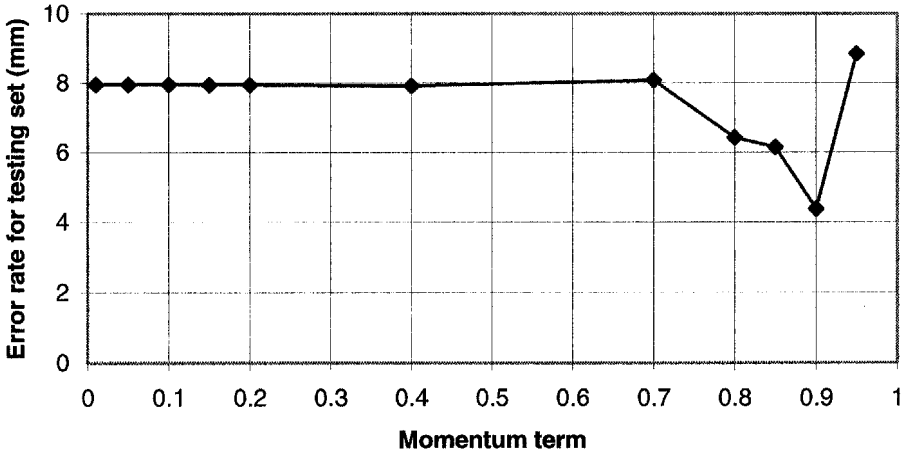


Figure 6.2: Effect of various momentum terms on ANN performance (L.R = 0.2)

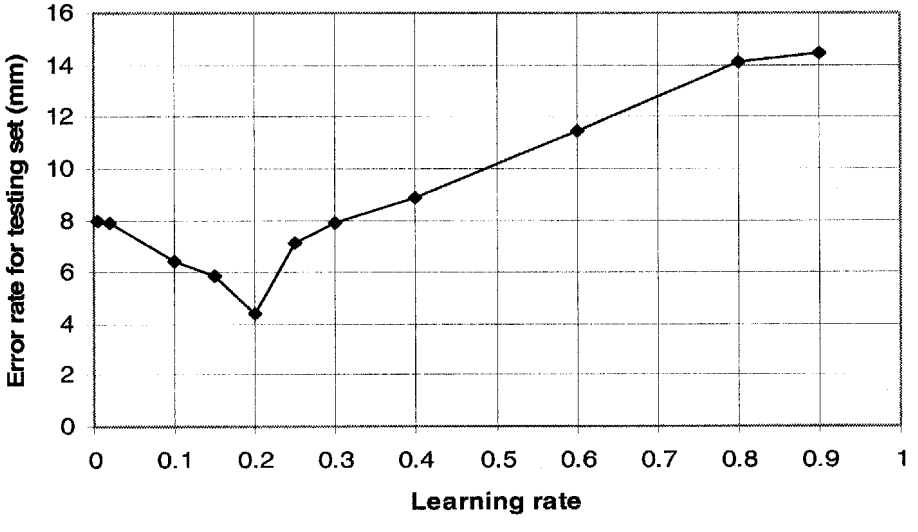


Figure 6.3: Effect of various learning rates on ANN performance (Momentum = 0.9)

Table 6.10 : Structure and Performance of ANN model NN6

Parameters effect	No. of input units	No. of hidden neurons	Neural Network Input Parameters	Learning rate	Momentum	Transfer function hidden layer	Tranfer function output layer	Coefficient of correlation			Error rate		
								Training	Testing	Validation	Training	Testing	Validation
Momentum	8	8	H, EP, AR, S1, S2, E, MC,GP	0.2	0.01	Tansig	Tansig	0.85	0.82	0.81	6.78	7.96	9.46
				0.2	0.05	Tansig	Tansig	0.85	0.82	0.81	6.77	7.95	9.45
				0.2	0.1	Tansig	Tansig	0.85	0.82	0.81	6.77	7.95	9.43
				0.2	0.15	Tansig	Tansig	0.85	0.82	0.81	6.77	7.95	9.41
				0.2	0.2	Tansig	Tansig	0.85	0.82	0.81	6.77	7.94	9.38
				0.2	0.4	Tansig	Tansig	0.85	0.82	0.82	6.75	7.91	9.21
				0.2	0.7	Tansig	Tansig	0.85	0.79	0.75	6.68	8.08	11.39
				0.2	0.8	Tansig	Tansig	0.89	0.87	0.74	5.65	6.43	10.73
				0.2	0.85	Tansig	Tansig	0.89	0.90	0.75	6.18	6.15	12.33
				0.2	0.9	Tansig	Tansig	0.91	0.95	0.92	5.37	4.38	4.77
				0.2	0.95	Tansig	Tansig	0.93	0.69	0.53	5.52	8.84	12.01
Learning rates				0.005	0.9	Tansig	Tansig	0.85	0.82	0.81	6.77	7.95	9.51
				0.02	0.9	Tansig	Tansig	0.85	0.82	0.82	6.76	7.92	9.31
				0.1	0.9	Tansig	Tansig	0.89	0.87	0.73	5.70	6.41	11.35
				0.15	0.9	Tansig	Tansig	0.87	0.92	0.89	6.05	5.88	8.08
				0.2	0.9	Tansig	Tansig	0.91	0.95	0.92	5.37	4.38	4.77
				0.25	0.9	Tansig	Tansig	0.97	0.86	0.89	4.22	7.12	7.36
				0.3	0.9	Tansig	Tansig	0.90	0.81	0.61	6.16	7.92	11.44
				0.4	0.9	Tansig	Tansig	0.93	0.69	0.51	5.52	8.84	12.35
				0.6	0.9	Tansig	Tansig	0.83	0.29	0.50	5.17	11.44	11.18
				0.8	0.9	Tansig	Tansig	0.06	0.14	0.04	15.73	14.08	14.72
				0.9	0.9	Tansig	Tansig	0.07	0.14	0.04	16.02	14.47	14.96
Transfer functions				0.2	0.9	Sigmoid	Sigmoid	0.17	0.18	0.37	43.53	42.94	42.58
				0.2	0.9	Sigmoid	Tansig	0.12	0.08	0.06	13.72	12.87	12.60
				0.2	0.9	Tansig	Sigmoid	0.10	0.05	-0.01	43.53	42.94	42.58
				0.2	0.9	Tansig	Linear	0.88	0.90	0.76	5.16	5.99	11.34
				0.2	0.9	Sigmoid	Linear	0.91	0.86	0.86	5.02	7.23	8.53

For subsequent analysis, the ANN model NN 6 used learning rate of 0.2, momentum of 0.9 and tan-sigmoid transfer function for both hidden layer and output layer.

6.5 Comparison of training algorithms

In the previous analysis, the network models used back-propagation training algorithm: gradient descent with momentum to adjust the weights. In Chapter 2, several faster algorithms are discussed. This section investigates the effect of their use on the network performance. These algorithms with their acronym are listed in Table 6.11. The results of all the algorithms training including gradient descent method are tabulated in Table 6.12. The model used was NN 6 with 8 hidden neurons.

Table 6.11: Back-propagation Training Algorithms

Acronym	Algorithm
LM	trainlm - Levenberg-Marquardt
GDM	traingdm - Gradient Descent Method
BFG	trainbfg - BFGS Quasi-Newton
RP	trainrp - Resilient Backpropagation
SCG	trainscg - Scaled Conjugate Gradient
CGB	traincgb - Conjugate Gradient with Powell/Beale Restarts
CGF	traincgf - Fletcher-Powell Conjugate Gradient
CGP	traincgp - Polak-Ribiére Conjugate Gradient
OSS	trainoss - One-Step Secant
GDX	traingdx - Variable Learning Rate Backpropagation

Table 6.12: Performance of ANN model NN 6 using various training algorithms

No. of input units	No. of hidden neurons	Neural Network Input Parameters	Algorithm	No of. Cycles	Coefficient of correlation			Error rate		
					Training	Testing	Validation	Training	Testing	Validation
8	8	H, EP, AR, S1, S2, E, MC,GP	LM	123	0.87	0.84	0.74	6.10	7.15	10.74
			GDM	3709	0.91	0.95	0.92	5.37	4.38	4.77
			BFG	3376	0.91	0.88	0.77	5.66	5.97	8.94
			RP	20	0.83	0.87	0.81	7.50	7.01	9.72
			SCG	21	0.82	0.88	0.83	7.72	6.47	8.85
			CGB	3280	0.80	0.86	0.85	7.40	6.75	7.97
			CGF	1459	0.85	0.88	0.84	7.29	6.99	8.10
			CGP	27	0.86	0.88	0.83	6.31	6.49	8.06
			OSS	35	0.81	0.87	0.88	7.03	6.60	7.78
			GDX	535	0.87	0.88	0.84	6.22	6.04	7.31

Figure 6.4 shows graphically the comparison of all training algorithms with respect to the testing set. It can be seen that Gradient Descent Method (GDM) gives the lowest error rate compared to other algorithms. Thus, its use for network analysis in this project is appropriate. Other algorithms, which give low error rates as well, include One-Step Secant method (OSS) and Variable Learning Rate Back propagation (GDX). As expected, the GDM method requires the most number of cycles compare to other faster algorithms to reach the minimum error.

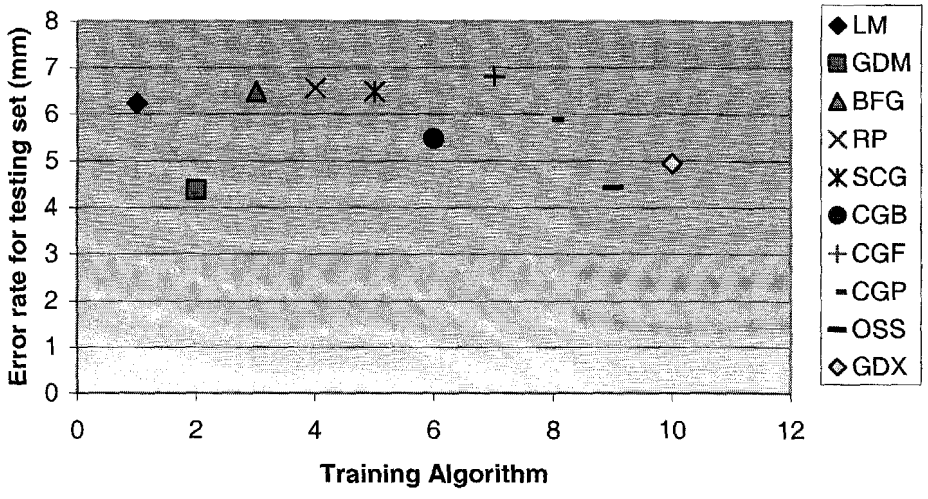


Figure 6.4: Comparison of ANN model performance using various training algorithms

6.6 Data Division for ANN Models

Recent studies indicate that the way the data are divided can have a significant impact on the performance of network model (Tokar and Johnson 1999). ANN models are developed for its interpolation ability, and not extrapolation ability. Consequently, the training set should contain all maximum and minimum data points in order to perform well. If early-stopping is used as the stopping criterion, the testing set should be representative of the training set, as the testing set is used to decide when to stop training and obtain optimal model architecture. Validation set is used to test the generalization ability of the model, thus it should be representative of the training set as well. For these reasons, the statistical properties of the various data subsets (training, testing and validation) need to be similar to

ensure that each subset represents the same statistical population (Masters 1993). In this section, the effect of data division on ANN model performance is investigated. Three data division methods are tested: (i) data division to ensure statistical consistency; (ii) data division using self-organizing maps (SOMs); and (iii) data division using fuzzy clustering. To examine how representative the training, testing and validation sets are with respect to each other, t- and F-tests are carried out. The t-test checks the null hypothesis of no difference in the means of two data sets and the F-test investigates the null hypothesis of no difference in the standard deviations of the two sets. For a given level of significance, test statistics can be calculated to test the null hypotheses for the t- and F-tests, respectively. Traditionally, a level of significance equal to 0.05 is selected (Levine et al. 1999). This means that there is a confidence level of 95% that the training, testing and validation sets are statistically consistent.

- **Approach 1: Statistically Consistent**

This method of data division has been adopted for analysis of ANN in the previous sections. In this approach, the input data are divided into their subsets in such a way that the statistical properties of the training, testing and validation are as close to each other as possible. Hence, the three data sets represent the same statistical population. The available 158 patterns are divided into three statistically consistent subsets as shown in Table 6.13. It can be seen that, for each input variable, the statistical properties of training, testing and validation sets are very close to each other. This is also confirmed by the results of null hypothesis tests in Table 6.14, which show that the hypotheses of testing and validation sets for all input parameters passed the t-test and the F-test.

Table 6.13: Input and output statistics obtained using data division to ensure statistical consistency

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Cover, H(m)					
Training set	16.9	3.9	8.5	30.0	21.5
Testing set	18.7	4.9	11.0	29.0	18.0
Validation set	17.1	4.5	11.0	29.0	18.0
Advance Rate, AR(mm/min)					
Training set	30.6	10.4	9.5	52.1	42.7
Testing set	29.4	12.4	10.5	50.0	39.5
Validation set	30.7	10.5	10.4	47.5	37.1
Earth Pressure, EP(kPa)					
Training set	189.0	83.6	11.0	370.0	359.0
Testing set	187.4	73.6	62.0	347.0	285.0
Validation set	220.8	70.9	90.0	340.0	250.0
SPT1, N					
Training set	26.6	28.8	0.7	80.3	79.7
Testing set	29.6	27.1	2.6	75.6	73.0
Validation set	21.0	26.5	0.8	78.0	77.2
SPT 2, N					
Training set	54.3	42.2	0.0	100.0	100.0
Testing set	58.5	42.9	0.0	100.0	100.0
Validation set	46.4	41.7	0.0	100.0	100.0
Moisture content, MC(%)					
Training set	27.5	18.2	6.0	66.5	60.5
Testing set	28.4	20.7	9.5	65.8	56.3
Validation set	30.7	18.9	10.6	63.3	52.7
Stiffness, E(MPa)					
Training set	70.6	51.3	5.0	120.0	115.0
Testing set	73.5	51.9	5.3	120.0	114.7
Validation set	60.1	50.0	5.0	120.0	115.0
Grout Pressure, GP(kPa)					
Training set	247.2	166.1	1.8	700.0	698.2
Testing set	243.2	144.5	35.0	500.0	465.0
Validation set	245.4	158.0	30.0	500.0	470.0
Measured settlement, S(mm)					
Training set	16.8	22.9	0.2	112.9	112.7
Testing set	15.6	20.2	0.3	94.4	94.1
Validation set	17.0	19.5	0.9	79.9	79.0

Table 6.14: Null hypothesis tests for data division to ensure statistical consistency

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
Cover								
Testing	-2.0	-2.0	2.0	Accept	0.7	0.6	1.8	Accept
Validation	-0.3	-2.0	2.0	Accept	0.8	0.5	1.9	Accept
AR								
Testing	0.5	-2.0	2.0	Accept	0.7	0.6	1.8	Accept
Validation	0.5	-2.0	2.0	Accept	1.0	0.5	1.9	Accept
EP								
Testing	0.1	-2.0	2.0	Accept	1.3	0.6	1.8	Accept
Validation	-2.0	-2.0	2.0	Accept	1.4	0.5	1.9	Accept
SPT1								
Testing	-0.6	-2.0	2.0	Accept	1.1	0.6	1.8	Accept
Validation	1.0	-2.0	2.0	Accept	1.2	0.5	1.9	Accept
SPT2								
Testing	-0.5	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	0.9	-2.0	2.0	Accept	1.0	0.5	1.9	Accept
MC								
Testing	-0.2	-2.0	2.0	Accept	0.8	0.6	1.8	Accept
Validation	-0.8	-2.0	2.0	Accept	0.9	0.5	1.9	Accept
E								
Testing	-0.3	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	1.0	-2.0	2.0	Accept	1.1	0.5	1.9	Accept
GP								
Testing	0.1	-2.0	2.0	Accept	1.3	0.6	1.8	Accept
Validation	0.1	-2.0	2.0	Accept	1.1	0.5	1.9	Accept
Settlement								
Testing	0.3	-2.0	2.0	Accept	1.3	0.6	1.8	Accept
Validation	0.0	-2.0	2.0	Accept	1.4	0.5	1.9	Accept

- **Approach 2: Self-Organizing Map (SOM)**

Teuvo Kohonen introduced the Self-Organizing Map (SOM) in 1982. The SOM algorithm is based on unsupervised, competitive learning which means that no human intervention is needed during the learning. Self-organizing feature maps (SOFM) learn to classify input vectors according to how they are grouped in the input space. In this section, SOM is used to organize the data into clusters. Once clustering has been successfully accomplished, samples are chosen from each cluster to form the training, testing and validation sets. The method proposed by Bowden et al. (2002) is followed whereby three samples are selected randomly from each cluster; one for each of the training, testing, and validation sets. If a cluster contains two data, one is chosen for training and the other for testing. If a

cluster contains only one data, it is included in the training set. SOM is a convenient method for data division as it eliminates the need to decide which proportion of the data to use for training, testing and validation sets. Furthermore, the statistical properties of the training, testing and validation sets are nearly similar. However, this method requires the controlling parameters in learning process (i.e. learning rate, neighbourhood size, size and shape of the map) to be selected in advance. To obtain optimum combinations of these parameters, a trial-and-error approach should be used. As part of this approach, SOM function in MatLab's Neural Network toolbox is used to cluster the data. The available input variables (i.e. H, EP, AR, SPT1, SPT2, MC, E, GP) and output variable (S) are presented to the SOM as inputs (Figure 6.5). As there is no precise rule to determine the optimum size of the map, three map sizes, 5×5, 8×8, and 10×10 are investigated. Training is carried out for 10000 iterations using the default parameters recommended in the Network/Data Manager (neighbourhood distance :1, Ordering phase learning rate:0.9, Tuning phase learning rate:0.02).

Map size 5×5

From the clusters of data, 119 patterns are selected for training set, 22 patterns for testing set and 20 patterns for validation set. The statistics of the training, testing and validation sets for map size 5×5 are shown in Table 6.15. It can be seen that the statistical properties of the three subsets are quite similar as expected. This is further verified by the results of t- and F-tests (Table 6.16), whereby hypothesis of all input parameters for both testing and validation sets are acceptable.

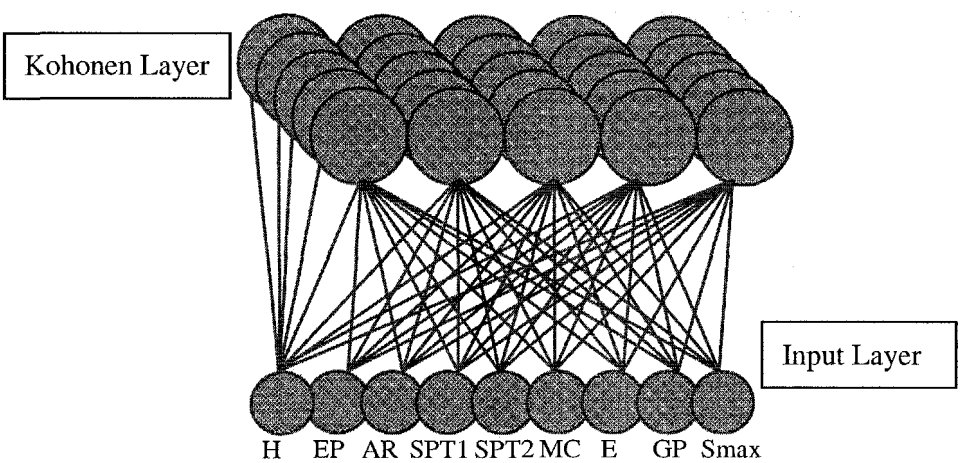


Figure 6.5: SOM for settlement data clustering

Table 6.15: Input and output statistics obtained using SOM for 5×5 map size

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Cover, H(m)					
Training set	17.1	4.3	8.5	30.0	21.5
Testing set	18.0	4.3	11.0	29.0	18.0
Validation set	18.7	3.6	13.0	25.0	12.0
Advance Rate, AR(mm/min)					
Training set	31.2	10.8	9.5	52.1	42.7
Testing set	27.5	11.3	10.5	51.0	40.5
Validation set	28.2	10.5	10.4	50.0	39.6
Earth Pressure, EP(kPa)					
Training set	199.4	82.0	11.0	370.0	359.0
Testing set	194.6	80.5	62.0	347.0	285.0
Validation set	193.6	77.1	90.0	342.0	252.0
SPT1, N					
Training set	25.2	28.2	0.7	80.3	79.7
Testing set	28.1	26.1	0.9	70.4	69.5
Validation set	30.8	28.4	2.4	80.3	77.9
SPT 2, N					
Training set	52.2	42.4	0.0	100.0	100.0
Testing set	62.8	40.7	0.0	100.0	100.0
Validation set	62.6	39.7	4.7	100.0	95.3
Moisture content, MC(%)					
Training set	29.0	19.3	6.0	66.5	60.5
Testing set	24.8	17.7	10.2	63.3	53.1
Validation set	24.2	15.4	6.1	58.7	52.6
Stiffness, E(MPa)					
Training set	67.4	51.4	5.0	120.0	115.0
Testing set	80.7	50.7	5.3	120.0	114.7
Validation set	79.6	47.0	5.3	120.0	114.7
Grout Pressure, GP(kPa)					
Training set	240.0	163.3	1.8	700.0	698.2
Testing set	277.3	144.0	13.5	500.0	486.5
Validation set	283.6	133.8	32.0	500.0	468.0
Measured settlement, S(mm)					
Training set	16.6	21.4	0.2	112.9	112.7
Testing set	16.3	23.7	0.5	94.4	93.9
Validation set	14.1	20.2	0.3	79.9	79.6

Table 6.16: Null hypothesis tests for data subsets obtained using SOM 5×5 map size

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
Cover								
Testing	-0.9	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
Validation	-1.5	-2.0	2.0	Accept	1.5	0.5	2.2	Accept
AR								
Testing	1.5	-2.0	2.0	Accept	0.9	0.5	2.1	Accept
Validation	1.1	-2.0	2.0	Accept	1.1	0.5	2.2	Accept
EP								
Testing	0.3	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
Validation	0.3	-2.0	2.0	Accept	1.1	0.5	2.2	Accept
SPT1								
Testing	-0.5	-2.0	2.0	Accept	1.2	0.5	2.1	Accept
Validation	-0.8	-2.0	2.0	Accept	1.0	0.5	2.2	Accept
SPT2								
Testing	-1.1	-2.0	2.0	Accept	1.1	0.5	2.1	Accept
Validation	-1.0	-2.0	2.0	Accept	1.1	0.5	2.2	Accept
MC								
Testing	0.9	-2.0	2.0	Accept	1.2	0.5	2.1	Accept
Validation	1.0	-2.0	2.0	Accept	1.6	0.5	2.2	Accept
E								
Testing	-1.1	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
Validation	-1.0	-2.0	2.0	Accept	1.2	0.5	2.2	Accept
GP								
Testing	-1.0	-2.0	2.0	Accept	1.3	0.5	2.1	Accept
Validation	-1.1	-2.0	2.0	Accept	1.5	0.5	2.2	Accept
Settlement								
Testing	0.1	-2.0	2.0	Accept	0.8	0.5	2.1	Accept
Validation	0.5	-2.0	2.0	Accept	1.1	0.5	2.2	Accept

Map size 8×8

The numbers of patterns for training, testing and validation sets are 98, 36 and 27 respectively. The statistical properties of the training, testing and validation sets for map size 8×8 are shown in Table 6.17. The statistics of the three subsets are consistent and this is confirmed by the outcomes of the t- and F-tests (Table 6.18).

Table 6.17: Input and output statistics obtained using SOM for 8×8 map size

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Cover, H(m)					
Training set	17.6	4.6	30.0	8.5	21.5
Testing set	17.4	3.7	25.0	11.0	14.0
Validation set	17.1	4.0	25.0	11.2	13.8
Advance Rate, AR(mm/min)					
Training set	30.4	10.8	52.1	9.5	42.7
Testing set	29.7	11.2	50.0	10.5	39.5
Validation set	30.9	11.0	50.0	10.4	39.6
Earth Pressure, EP(kPa)					
Training set	197.7	81.8	370.0	11.0	359.0
Testing set	194.6	83.9	347.0	52.0	295.0
Validation set	203.7	75.0	350.0	90.0	260.0
SPT1, N					
Training set	26.0	28.6	80.3	0.7	79.7
Testing set	25.9	25.7	70.3	0.9	69.5
Validation set	27.8	28.8	80.3	0.8	79.4
SPT 2, N					
Training set	52.2	42.5	100.0	0.0	100.0
Testing set	60.1	40.8	100.0	0.0	100.0
Validation set	57.9	41.8	100.0	0.0	100.0
Moisture content, MC(%)					
Training set	29.3	19.3	66.5	6.0	60.5
Testing set	25.1	17.3	63.3	9.5	53.8
Validation set	26.1	18.2	63.3	6.1	57.2
Stiffness, E(MPa)					
Training set	67.0	51.3	120.0	5.0	115.0
Testing set	78.0	50.4	120.0	5.1	114.9
Validation set	74.4	49.8	120.0	5.3	114.7
Grout Pressure, GP(kPa)					
Training set	242.8	161.7	700.0	1.8	698.2
Testing set	259.3	150.7	500.0	13.5	486.5
Validation set	266.5	155.3	500.0	30.0	470.0
Measured settlement, S(mm)					
Training set	16.7	22.2	112.9	0.2	112.7
Testing set	15.4	21.5	94.4	0.5	93.9
Validation set	16.1	19.3	79.9	0.8	79.1

Table 6.18: Null hypothesis tests for data subsets obtained using SOM 8×8 map size

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
Cover								
Testing	0.3	-2.0	2.0	Accept	1.5	0.6	1.8	Accept
Validation	0.5	-2.0	2.0	Accept	1.3	0.5	2.0	Accept
AR								
Testing	0.3	-2.0	2.0	Accept	0.9	0.6	1.8	Accept
Validation	-0.2	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
EP								
Testing	0.2	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	-0.3	-2.0	2.0	Accept	1.2	0.5	2.0	Accept
SPT1								
Testing	0.0	-2.0	2.0	Accept	1.2	0.6	1.8	Accept
Validation	-0.3	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
SPT2								
Testing	-1.0	-2.0	2.0	Accept	1.1	0.6	1.8	Accept
Validation	-0.6	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
MC								
Testing	1.1	-2.0	2.0	Accept	1.2	0.6	1.8	Accept
Validation	0.8	-2.0	2.0	Accept	1.1	0.5	2.0	Accept
E								
Testing	-1.1	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	-0.7	-2.0	2.0	Accept	1.1	0.5	2.0	Accept
GP								
Testing	-0.5	-2.0	2.0	Accept	1.2	0.6	1.8	Accept
Validation	-0.7	-2.0	2.0	Accept	1.1	0.5	2.0	Accept
Settlement								
Testing	0.3	-2.0	2.0	Accept	1.1	0.6	1.8	Accept
Validation	0.1	-2.0	2.0	Accept	1.3	0.5	2.0	Accept

Map size 10×10

Since its size is the largest, this map has the most clusters of data. Therefore, more data can be included in the testing set. The numbers of patterns for training, testing and validation sets are 91, 46 and 24. Table 6.19 shows the statistical properties of the three subsets while its consistency is verified from the t- and F-tests results (Table 6.20).

Table 6.19: Input and output statistics obtained using SOM for 10×10 map size

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Cover, H(m)					
Training set	17.1	4.3	8.5	30.0	21.5
Testing set	17.7	4.2	11.0	29.0	18.0
Validation set	18.3	4.3	11.2	29.0	17.8
Advance Rate, AR(mm/min)					
Training set	31.0	10.6	9.5	52.1	42.7
Testing set	29.1	11.8	10.5	51.0	40.5
Validation set	30.2	10.2	10.4	47.5	37.1
Earth Pressure, EP(kPa)					
Training set	199.2	83.2	11.0	370.0	359.0
Testing set	191.8	77.2	52.0	347.0	295.0
Validation set	206.0	80.5	70.0	350.0	280.0
SPT1, N					
Training set	24.5	27.2	0.7	80.3	79.7
Testing set	27.8	27.8	0.8	80.3	79.5
Validation set	30.2	31.0	2.7	80.3	77.6
SPT 2, N					
Training set	52.4	41.7	0.0	100.0	100.0
Testing set	59.0	42.8	0.0	100.0	100.0
Validation set	56.8	41.8	0.0	100.0	100.0
Moisture content, MC(%)					
Training set	29.0	19.0	6.0	66.5	60.5
Testing set	26.6	18.3	7.8	63.3	55.5
Validation set	25.5	18.6	6.1	63.3	57.2
Stiffness, E(MPa)					
Training set	68.5	50.4	5.0	120.0	115.0
Testing set	75.0	52.0	5.0	120.0	115.0
Validation set	71.1	51.3	5.2	120.0	114.8
Grout Pressure, GP(kPa)					
Training set	247.9	162.2	1.8	700.0	698.2
Testing set	248.7	151.6	13.5	500.0	486.5
Validation set	264.2	156.7	31.2	500.0	468.8
Measured settlement, S(mm)					
Training set	16.7	22.3	0.2	112.9	112.7
Testing set	16.4	21.4	0.5	94.4	93.9
Validation set	14.3	18.6	0.8	79.9	79.1

Table 6.20: Null hypothesis tests for data subsets obtained using SOM 10×10 map size

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
Cover								
Testing	-0.7	-2.0	2.0	Accept	1.0	0.6	1.7	Accept
Validation	-1.2	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
AR								
Testing	0.9	-2.0	2.0	Accept	0.8	0.6	1.7	Accept
Validation	0.3	-2.0	2.0	Accept	1.1	0.5	2.1	Accept
EP								
Testing	0.5	-2.0	2.0	Accept	1.2	0.6	1.7	Accept
Validation	-0.4	-2.0	2.0	Accept	1.1	0.5	2.1	Accept
SPT1								
Testing	-0.7	-2.0	2.0	Accept	1.0	0.6	1.7	Accept
Validation	-0.9	-2.0	2.0	Accept	0.8	0.5	2.1	Accept
SPT2								
Testing	-0.9	-2.0	2.0	Accept	0.9	0.6	1.7	Accept
Validation	-0.4	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
MC								
Testing	0.7	-2.0	2.0	Accept	1.1	0.6	1.7	Accept
Validation	0.8	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
E								
Testing	-0.7	-2.0	2.0	Accept	0.9	0.6	1.7	Accept
Validation	-0.2	-2.0	2.0	Accept	1.0	0.5	2.1	Accept
GP								
Testing	0.0	-2.0	2.0	Accept	1.1	0.6	1.7	Accept
Validation	-0.4	-2.0	2.0	Accept	1.1	0.5	2.1	Accept
Settlement								
Testing	0.1	-2.0	2.0	Accept	1.1	0.6	1.7	Accept
Validation	0.5	-2.0	2.0	Accept	1.4	0.5	2.1	Accept

- **Approach 3: Fuzzy Clustering**

In normal partitioning methods, each object of the data set is assigned to one and only one cluster. As a result, each object has a membership of 1 in some cluster and a zero membership in all other clusters. Frequently, we have intermediate data objects which lie approximately the same distance from some clusters. In this case, it would be very difficult to decide in which cluster to put the objects. Normally, the partitioning method would assign the objects arbitrarily to one of the clusters from which the objects have the same distance. A fuzzy clustering technique is much better equipped to handle such situations. This method assigns *membership coefficients* (range from 0 to 1) to all data objects which indicate the degree of belonging of each object to all the clusters. For instance, if object 1 belongs for 87% to cluster 1, for 6% to cluster 2, and for 7% to cluster 3, then object 1 belong mainly

to the first cluster. In the case of intermediate objects, the membership coefficients will be approximately equal with respect to several clusters. The object is then grouped to the cluster to which it has the largest membership coefficient.

The fuzzy clustering algorithm attempts to minimize the following objective function (Kaufman and Rousseeuw 1990):

$$C=\sum_{v=1}^{k}\frac{\sum_{i,j=1}^{p}u_{iv}^{2}u_{jv}^{2}d_{ij}}{2\sum_{j=1}^{p}u_{jv}^{2}} \tag{6.4}$$

where :

k = number of clusters;

d_{ij} = given distance between data point i and j

u_{iv} = unknown membership of object i to cluster v

The membership functions are subject to the constraints:

$u_{iv} \geq 0$ for i = 1,…,n ; v = 1,…,k (6.5)

$\sum_{v} u_{iv} = 1$ for i = 1,…,n (6.6)

The constraints imply that memberships cannot be negative and that each object has a constant total membership of 1, distributed over the different clusters.

The procedure for fuzzy clustering adopted in this research is as follows:

1. The optimum number of clusters is determined by using subtractive clustering method provided in MatLab's Fuzzy Logic ToolBox.
2. For the optimum number of clusters, the data records included in each cluster are ranked according to their membership values in incremental intervals of 0.1 between 0.0 and 1.0 (i.e. 0.0-0.1, 0.1-0.2,…, 0.9-1.0).
3. For each cluster and membership interval (e.g. cluster 1 and membership interval 0.0-0.1), two samples are chosen, one for the testing set and one for the validation set, and the remaining data samples are chosen for the training set. In the case where two records are obtained, one record is chosen for

training and the other is chosen for testing. If only one record is obtained, this record is included in the training set.

From step no. 1, the optimum number of clusters was found to be 8. The membership values obtained for all data records are shown in Appendix C. From step no. 3, 98 patterns are used for training set, 38 patterns for testing set and 25 patterns for validation set. Table 6.22 shows the statistics of the data in the training, testing and validation sets obtained using fuzzy clustering. The results of t- and F-tests verify the statistical consistency of the three subsets as shown in Table 6.23.

6.7 Results and Discussion

Performance of ANN model NN6 using data subsets obtained from different approaches of data division is shown in Table 6.21. It can be seen that the results obtained for the statistically consistent data division method are better than the results obtained for SOM and fuzzy clustering data division methods. The testing set has a high correlation coefficient coupled with low error rate. In addition, the generalization ability of the model is satisfactory, as shown by high correlation coefficient and low error rate of the validation set. Consequently, ANN model NN6 developed using statistically consistent data division will be used for the subsequent analysis and will be referred to as the ANN model.

Table 6.21: Performance of ANN models using data subsets obtained for different approaches of data division

Performance measures and data sets	Statistical Division	SOM			Fuzzy Clustering
		5×5	8×8	10×10	
Training					
Correlation coefficient, *r*	0.91	0.96	0.97	0.85	0.84
Error rate (mm)	5.37	4.02	3.46	6.99	6.61
Testing					
Correlation coefficient, *r*	0.95	0.89	0.81	0.86	0.90
Error rate (mm)	4.38	6.57	7.48	7.83	5.48
Validation					
Correlation coefficient, *r*	0.92	0.70	0.89	0.92	0.94
Error rate (mm)	4.77	10.94	8.73	6.22	4.75

Table 6.22: Input and output statistics for data sets obtained using fuzzy clustering

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Cover, H(m)					
Training set	17.4	4.3	8.5	30.0	21.5
Testing set	17.8	4.4	11.2	29.0	17.8
Validation set	17.0	4.0	11.0	25.0	14.0
Advance Rate, AR(mm/min)					
Training set	30.2	11.0	9.5	52.1	42.7
Testing set	29.6	10.6	9.8	50.0	40.2
Validation set	32.0	11.0	11.0	51.0	40.0
Earth Pressure, EP(kPa)					
Training set	202.7	86.8	11.0	370.0	359.0
Testing set	190.8	67.4	52.0	350.0	298.0
Validation set	190.8	76.0	70.0	350.0	280.0
SPT1, N					
Training set	25.8	27.9	0.7	80.3	79.7
Testing set	30.7	29.7	0.8	76.1	75.3
Validation set	21.5	24.5	1.4	80.3	79.0
SPT 2, N					
Training set	53.7	42.0	0.0	100.0	100.0
Testing set	57.2	43.5	0.0	100.0	100.0
Validation set	56.3	40.5	0.0	100.0	100.0
Moisture content, MC(%)					
Training set	28.0	18.7	6.0	66.5	60.5
Testing set	27.9	19.3	6.8	66.5	59.7
Validation set	27.0	18.3	7.8	63.3	55.5
Stiffness, E(MPa)					
Training set	69.4	50.1	5.0	120.0	115.0
Testing set	72.1	53.5	5.1	120.0	114.9
Validation set	73.5	51.1	5.0	120.0	115.0
Grout Pressure, GP(kPa)					
Training set	258.2	158.8	1.8	700.0	698.2
Testing set	226.9	159.0	30.0	500.0	470.0
Validation set	255.9	153.5	32.5	400.0	367.5
Measured settlement, S(mm)					
Training set	16.3	22.6	0.2	112.9	112.7
Testing set	16.2	19.1	0.3	79.9	79.6
Validation set	16.2	20.9	0.6	94.4	93.8

Table 6.23: Null hypothesis tests for data subsets obtained using fuzzy clustering

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
Cover								
Testing	-0.4	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	0.5	-2.0	2.0	Accept	1.2	0.5	2.0	Accept
AR								
Testing	0.3	-2.0	2.0	Accept	1.1	0.6	1.8	Accept
Validation	-0.7	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
EP								
Testing	0.8	-2.0	2.0	Accept	1.7	0.6	1.8	Accept
Validation	0.6	-2.0	2.0	Accept	1.3	0.5	2.0	Accept
SPT1								
Testing	-0.9	-2.0	2.0	Accept	0.9	0.6	1.8	Accept
Validation	0.7	-2.0	2.0	Accept	1.3	0.5	2.0	Accept
SPT2								
Testing	-0.4	-2.0	2.0	Accept	0.9	0.6	1.8	Accept
Validation	-0.3	-2.0	2.0	Accept	1.1	0.5	2.0	Accept
MC								
Testing	0.0	-2.0	2.0	Accept	0.9	0.6	1.8	Accept
Validation	0.2	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
E								
Testing	-0.3	-2.0	2.0	Accept	0.9	0.6	1.8	Accept
Validation	-0.4	-2.0	2.0	Accept	1.0	0.5	2.0	Accept
GP								
Testing	1.0	-2.0	2.0	Accept	1.0	0.6	1.8	Accept
Validation	0.1	-2.0	2.0	Accept	1.1	0.5	2.0	Accept
Settlement								
Testing	0.0	-2.0	2.0	Accept	1.4	0.6	1.8	Accept
Validation	0.0	-2.0	2.0	Accept	1.2	0.5	2.0	Accept

6.8 Sensitivity Analysis of the ANN Model Inputs

The purpose of sensitivity analysis is to identify which of the input variables have the most significant impact on settlement predictions. The relative importance of the input variables is obtained from the connection weights of the trained network using a simple and innovative technique proposed by Garson (1991). The optimum ANN model presented in the first year report of this project was used to illustrate the technique. The optimum ANN model had four input nodes, six hidden nodes and one output node with connection weights as shown in Table 6.24.

Table 6.24: Connection weights of a network with four inputs and six hidden nodes

Hidden nodes	Weights				
	Cover (H)	SPT1	MC	GWL	SettImnt
Hidden 1	-0.622740	0.382930	0.416470	-0.345010	-1.094900
Hidden 2	-0.333050	-0.157240	-3.320500	1.535800	2.282600
Hidden 3	1.102200	-0.731470	-0.292180	-0.105380	1.577700
Hidden 4	0.632980	0.275680	-1.634400	0.269070	-3.362100
Hidden 5	-1.294700	0.681890	0.130640	-1.014300	-1.399800
Hidden 6	0.918280	0.674620	0.582060	0.227410	-0.494610

The computational process proposed by Garson (1991) is as follows:

1. For each hidden node i, multiply the absolute value of the hidden-output layer connection weight by the absolute value of the hidden-input layer connection weight of each input variable j to obtain the products P_{ij} (j represents the column number of the weights mentioned above). As an example: $P_{11} = 0.622740 \times 1.094900 = 0.681838$. The results are tabulated in Table 6.25.

Table 6.25: Products P_{ij}

Hidden nodes	Cover (H)	SPT1	MC	GWL
Hidden 1	0.681838	0.419270	0.455993	0.377751
Hidden 2	0.760220	0.358916	7.579373	3.505617
Hidden 3	1.738941	1.154040	0.460972	0.166258
Hidden 4	2.128142	0.926864	5.495016	0.904640
Hidden 5	1.812321	0.954510	0.182870	1.419817
Hidden 6	0.454190	0.333674	0.287893	0.112479

2. For each hidden node, divide P_{ij} by the sum of all input variables to obtain Q_{ij} (Table 6.26). As an example :

   $Q_{11} = 0.681838 /(0.681838+0.419270+0.0.455993+0.377751) = 0.352398$

Table 6.26: Products Q_{ij}

Hidden nodes	Cover (H)	SPT1	MC	GWL
Hidden 1	0.352398	0.216694	0.235673	0.195235
Hidden 2	0.062292	0.029409	0.621050	0.287249
Hidden 3	0.493988	0.327833	0.130950	0.047230
Hidden 4	0.225089	0.098032	0.581196	0.095682
Hidden 5	0.414765	0.218447	0.041851	0.324937
Hidden 6	0.382239	0.280814	0.242286	0.094661

3. For each input node, sum Qij to obtain Sj (Table 6.27). As an example: S_1=0.352398+0.062292+0.493988+0.225089+0.414765+0.382239=1.93077

Table 6.27: Products Sj

	Cover(H)	SPT1	MC	GWL	Total
Sum	1.930771	1.17123	1.853007	1.044993	6

4. Divide Sj by the sum for all input variables to get the relative importance of all output weights attributed to the given input variable (Table 6.28). As an example, the relative importance for input node 1 is equal to : (1.930771×100)/6 = 32.2 %

Table 6.28: Relative importance (%)

Cover(H)	SPT1	MC	GWL
32.2	19.5	30.9	17.4

By using Garson method above, the relative importance of input parameters for the thirty two networks analyzed in section 6.3 are given in Table 6.29. The most important input for each network is indicated by the shaded box. It is observed that every network assigned relatively different relative importance with respect to an input parameter. What could be an important input for a particular network might be less important for other networks. Table 6.30 shows the relative importance of input variables for NN1, NN6, NN7, NN16, and NN26. The shaded box indicates the most important input for the network while the yellow box denotes the least important input. From the analysis in section 6.3, the last four networks are considered optimum out of the thirty two networks while NN1 is the network using the eleven parameters as input. NN1 put the highest importance on bulk density, whereas the most important input for NN6, NN7, NN16 and NN26 are earth pressure, advance rate, stiffness, and moisture content. Advance rate, which is an input of highest importance for NN7, is the least important for NN6. Similarly, stiffness is the least important for NN7, whereas it is the most important input for NN16. Overall, moisture content is considered very essential as it displays relatively high importance whenever it is used as input of the network.

Table 6.29: Relative importance (%) of inputs for 32 networks in Section 6.3

Model	No. of input units	Relative Importance										
		Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress
NN1	11	7.26	8.91	11.00	10.87	8.81	9.24	11.19	9.85	6.97	9.46	6.45
NN2	10	12.06	7.47	10.44	8.71	10.45	11.97	10.62	7.82	10.63	9.84	
NN3	9	11.03	11.43	9.03	11.51	14.16	6.65	9.65	16.76		9.79	
NN4	9	6.28	10.04	11.33	6.59	11.21		13.52	13.97		16.24	10.83
NN5	8		14.31	9.90	9.56	11.03	6.40	19.79	15.20		13.81	
NN6	8	12.29	9.94	15.64	10.23	10.13			15.54		13.03	13.20
NN7	8	13.45	18.91	12.38	11.20	12.11			11.64	10.23	10.07	
NN8	8	14.95	10.90	12.44	14.08			17.39	6.76		9.92	13.55
NN9	7		12.75	17.48	12.26		17.87	7.39	18.58		13.68	
NN10	7	11.11		11.07	13.57		14.23		18.76	13.88	17.38	
NN11	7		14.41	12.22	14.48		21.84		12.06		13.18	11.81
NN12	7		20.93	11.90	17.45		11.93		15.46	12.95		9.38
NN13	6			18.25	17.38		16.95	12.45	23.03		11.94	
NN14	6	18.53			18.60	10.61		21.35	16.73		14.18	
NN15	6		13.69		22.59	20.41			11.24		20.99	11.08
NN16	6		15.22		19.24			20.11	14.78		21.74	8.91
NN17	6	16.10				26.56		19.71	16.00	10.62		11.00
NN18	6	14.10		18.37	16.68		12.75		21.62		16.49	
NN19	5			15.50	18.06			22.27	19.58		24.60	
NN20	5	16.37		16.77	19.44		24.40		23.02			
NN21	5				16.62	27.79			16.91		21.13	17.55
NN22	5		16.82		15.09	25.66			21.65		20.78	
NN23	5		24.67		18.04	19.00		22.69			15.61	
NN24	5		16.53	23.87	20.79	20.75			18.06			
NN25	4			29.20	12.54			29.04	29.22			
NN26	4	26.37			20.91				30.79	21.93		
NN27	4				27.56		18.42		31.27		22.75	
NN28	4		31.64	30.90	14.73	22.73						
NN29	4	29.33			18.63	16.25			35.79			
NN30	3			42.12	25.90			31.98				
NN31	3			27.34				33.13	39.53			
NN32	3	38.54			25.69	35.78						

Table 6.30: Relative importance (%) of inputs for NN1, NN6, NN7, NN16, and NN26

Model	Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress
NN1	7.26	8.91	11.00	10.87	8.81	9.24	11.19	9.85	6.97	9.46	6.45
NN6	12.29	9.94	15.64	10.23	10.13			15.54		13.03	13.20
NN7	13.45	18.91	12.38	11.20	12.11			11.64	10.23	10.07	
NN16		15.22		19.24			20.11	14.78		21.74	8.91
NN26	26.37			20.91				30.79	21.93		

Table 6.29: Relative importance (%) of inputs for 32 networks in Section 6.3

Model	No. of input units	Relative Importance										
		Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress
NN1	11	7.26	8.91	11.00	10.87	8.81	9.24	11.19	9.85	6.97	9.46	6.45
NN2	10	12.06	7.47	10.44	8.71	10.45	11.97	10.62	7.82	10.63	9.84	
NN3	9	11.03	11.43	9.03	11.51	14.16	6.65	9.65	16.76		9.79	
NN4	9	6.28	10.04	11.33	6.59	11.21		13.52	13.97		16.24	10.83
NN5	8		14.31	9.90	9.56	11.03	6.40	19.79	15.20		13.81	
NN6	8	12.29	9.94	15.64	10.23	10.13			15.54		13.03	13.20
NN7	8	13.45	18.91	12.38	11.20	12.11			11.64	10.23	10.07	
NN8	8	14.95	10.90	12.44	14.08			17.39	6.76		9.92	13.55
NN9	7		12.75	17.48	12.26		17.87	7.39	18.58		13.68	
NN10	7	11.11		11.07	13.57		14.23		18.76	13.88	17.38	
NN11	7		14.41	12.22	14.48		21.84		12.06		13.18	11.81
NN12	7		20.93	11.90	17.45		11.93		15.46	12.95		9.38
NN13	6			18.25	17.38		16.95	12.45	23.03		11.94	
NN14	6	18.53			18.60	10.61		21.35	16.73		14.18	
NN15	6		13.69		22.59	20.41			11.24		20.99	11.08
NN16	6		15.22		19.24			20.11	14.78		21.74	8.91
NN17	6	16.10				26.56		19.71	16.00	10.62		11.00
NN18	6	14.10		18.37	16.68		12.75		21.62		16.49	
NN19	5			15.50	18.06			22.27	19.58		24.60	
NN20	5	16.37		16.77	19.44		24.40		23.02			
NN21	5				16.62	27.79			16.91		21.13	17.55
NN22	5		16.82		15.09	25.66			21.65		20.78	
NN23	5		24.67		18.04	19.00		22.69			15.61	
NN24	5		16.53	23.87	20.79	20.75			18.06			
NN25	4			29.20	12.54			29.04	29.22			
NN26	4	26.37			20.91				30.79	21.93		
NN27	4				27.56		18.42		31.27		22.75	
NN28	4		31.64	30.90	14.73	22.73						
NN29	4	29.33			18.63	16.25			35.79			
NN30	3			42.12	25.90			31.98				
NN31	3			27.34				33.13	39.53			
NN32	3	38.54			25.69	35.78						

Table 6.30: Relative importance (%) of inputs for NN1, NN6, NN7, NN16, and NN26

Model	Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress
NN1	7.26	8.91	11.00	10.87	8.81	9.24	11.19	9.85	6.97	9.46	6.45
NN6	12.29	9.94	15.64	10.23	10.13			15.54		13.03	13.20
NN7	13.45	18.91	12.38	11.20	12.11			11.64	10.23	10.07	
NN16		15.22		19.24			20.11	14.78		21.74	8.91
NN26	26.37			20.91				30.79	21.93		

6.9 Empirical method for settlement calculation

In this section, the empirical method described in Chapter 4 is used to obtain the predicted settlements for the same validation set used in ANN method. Since all the level instruments used in this research were positioned right above the tunnels, the maximum settlement over crown (W_{max}) calculated from empirical method, will serve as estimated settlement value.

The process to obtain the maximum ground surface settlements (W_{max}) above each tunnel is as follows (Gunn 1992):

1. The figure for volume loss is estimated and in this case it is assumed to be 3 %. This value is selected since 93 % of the sections along the North East Line tunnels show a volume loss of 2 % or less (Shirlaw et al. 2001). In contract C825, settlements predictions have been carried out based on equivalent face loss between 0.5 to 2 % (Osborne et al. 2004); while field observations indicated volume loss between 2-2.5 % for contract C823. Hence, the 3% volume loss will be on the conservative side with respect to the three contracts.
2. Point of inflection *i* is estimated from Equation 4.7 in Chapter 4 using the K and n values recommended by Attewell (1981).
3. Equation 4.6 in Chapter 4 is used to obtain maximum surface settlement (W_{max}).

   Details of the above calculations are shown in Appendix D.

Table 6.31 presents the maximum surface settlements of the validation set with their corresponding K, n and *i* values.

Table 6.31: Maximum surface settlement (W_{max}), K, n and *i* values obtained using empirical method

Instrument	K	n	Cover	Zo	*i*	W_{max}(mm)
3023	0.63	0.97	17.29	20.29	6.16	1.95
2315	0.65	0.97	19.35	22.35	6.99	1.72
2111	0.80	0.98	17.70	20.70	8.09	1.48
3033	0.63	0.97	17.22	20.22	6.14	1.95
2101	0.90	0.99	14.93	17.93	8.01	1.50
2291	0.63	0.97	17.85	20.85	6.33	1.90
2071	0.77	0.98	14.37	17.37	6.54	1.83
2281	0.63	0.97	17.74	20.74	6.29	1.91
2287	0.63	0.97	17.81	20.81	6.32	1.90
3025	0.63	0.97	15.17	18.17	5.54	2.17
2052	0.63	0.97	13.03	16.03	4.90	2.45
8077R	0.85	0.99	12.00	15.00	6.27	1.91
8083R	0.82	0.99	12.00	15.00	6.03	1.99
8086	0.87	0.99	11.00	14.00	6.00	2.00
8025	0.87	0.99	20.00	23.00	9.87	1.22
8031	0.82	0.99	21.00	24.00	9.69	1.24
8015	0.88	0.99	21.00	24.00	10.38	1.16
9804	0.76	0.98	25.00	28.00	10.37	1.16
5069	0.91	0.46	29.00	32.00	5.86	2.05
5078	0.91	0.46	28.00	31.00	5.81	2.06
1115	0.83	0.99	11.20	14.20	5.81	2.07
1103/50	0.91	0.99	16.40	19.40	8.75	1.37
1103/31	0.89	0.99	17.20	20.20	8.87	1.35
1101/06	0.93	0.99	16.00	19.00	8.74	1.37
1094/02	0.90	0.99	12.80	15.80	7.02	1.71
1099	0.87	0.99	15.20	18.20	7.86	1.53
1100	0.91	0.99	15.60	18.60	8.44	1.42
1101/08	0.93	0.99	16.00	19.00	8.74	1.37
1108/01	0.89	0.99	14.80	17.80	7.84	1.53

6.10 Comparison of ANN Model with Empirical method

Results of the validation set obtained using the optimum ANN model NN6 and the empirical method are compared in Figure 6.6. It is observed that ANN model predicts closely the measured settlements, whereas the empirical method under predicts the settlements for all the case record in validation set. Figure 6.7 shows the predicted results from ANN model against the measured settlements for all data obtained from contract C705. The figure shows that there is good agreement between the measured and the predicted settlemnts from ANN approach. Figures 6.8 to 6.10 show the separate comparisons of measured and predicted settlements obtained using optimum ANN model (a) and empirical method (b) for the training,

testing and validation sets. As shown in the figures, the superiority of ANN method over empirical one is clearly demonstrated as well for the data in the training and testing sets.

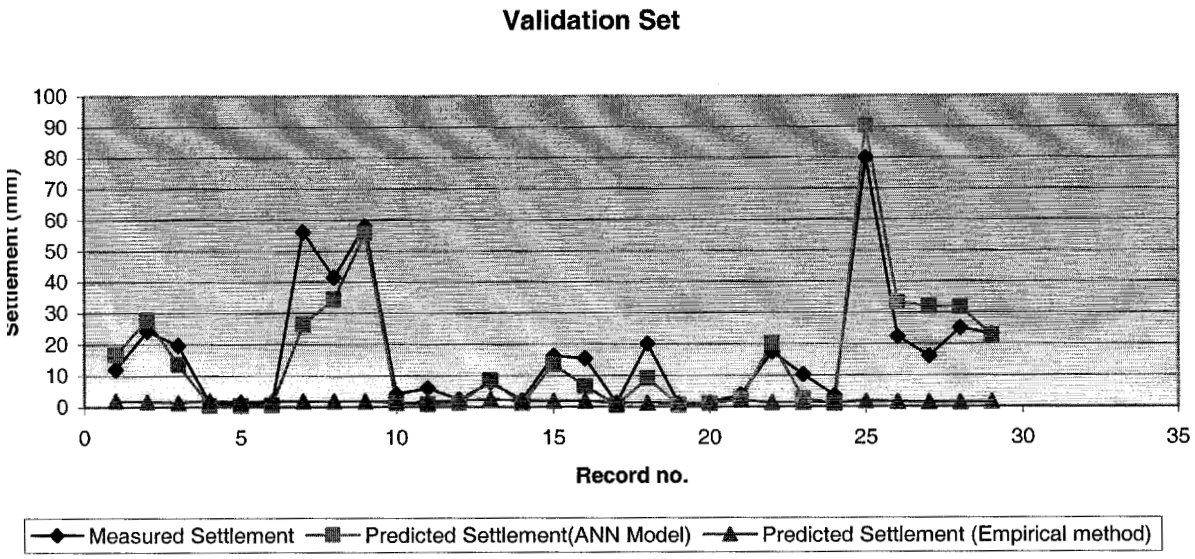


Figure 6.6: Comparison of measured settlements and the predicted settlements obtained using ANN and Empirical method

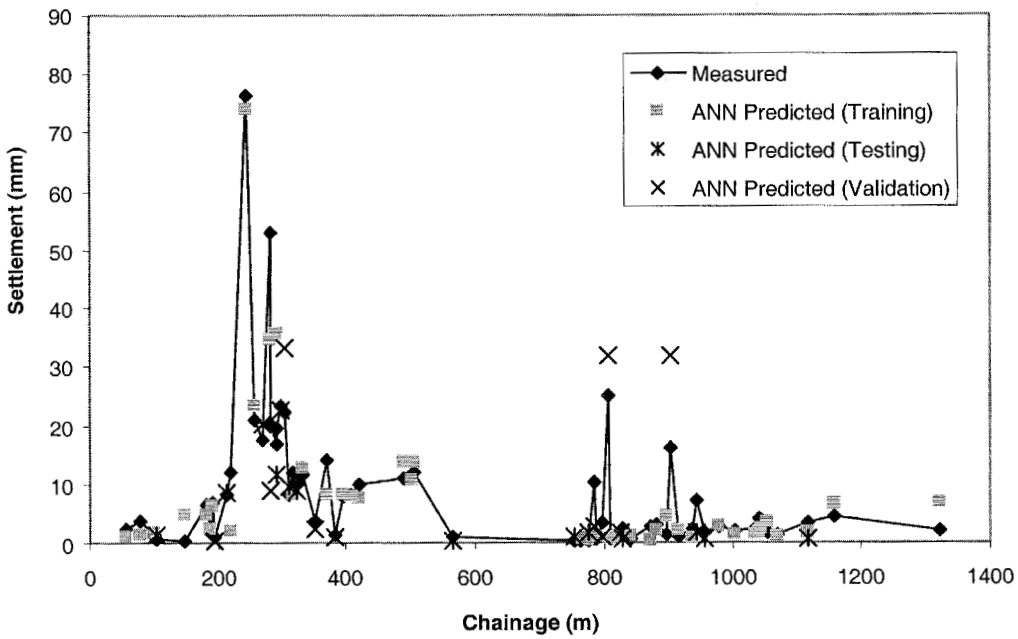
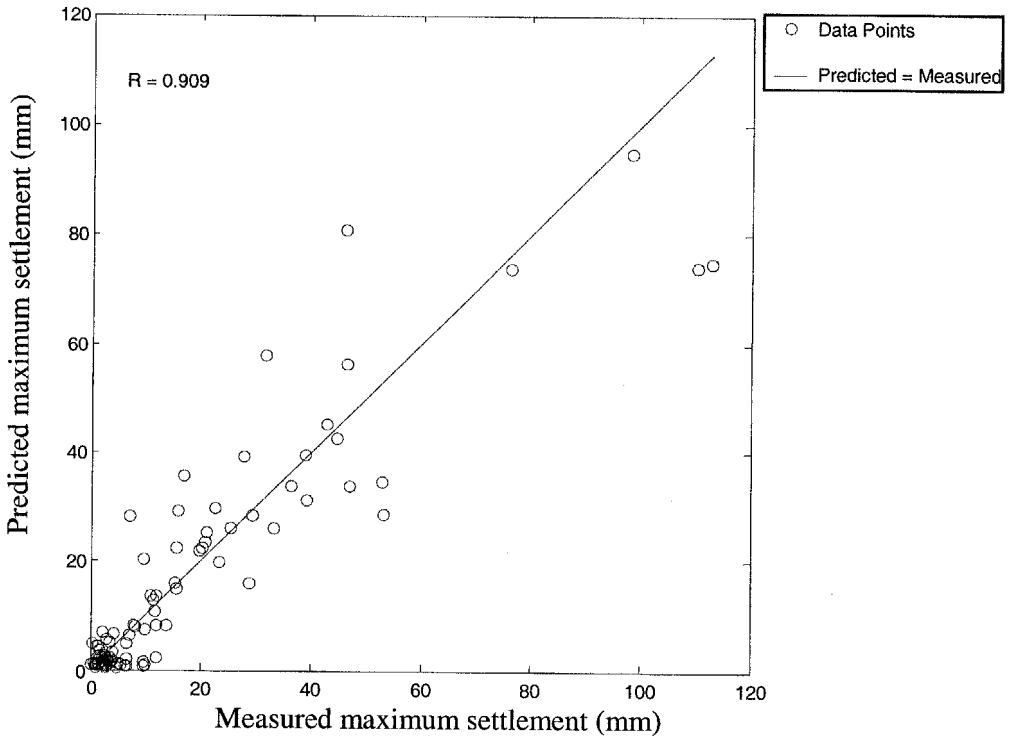
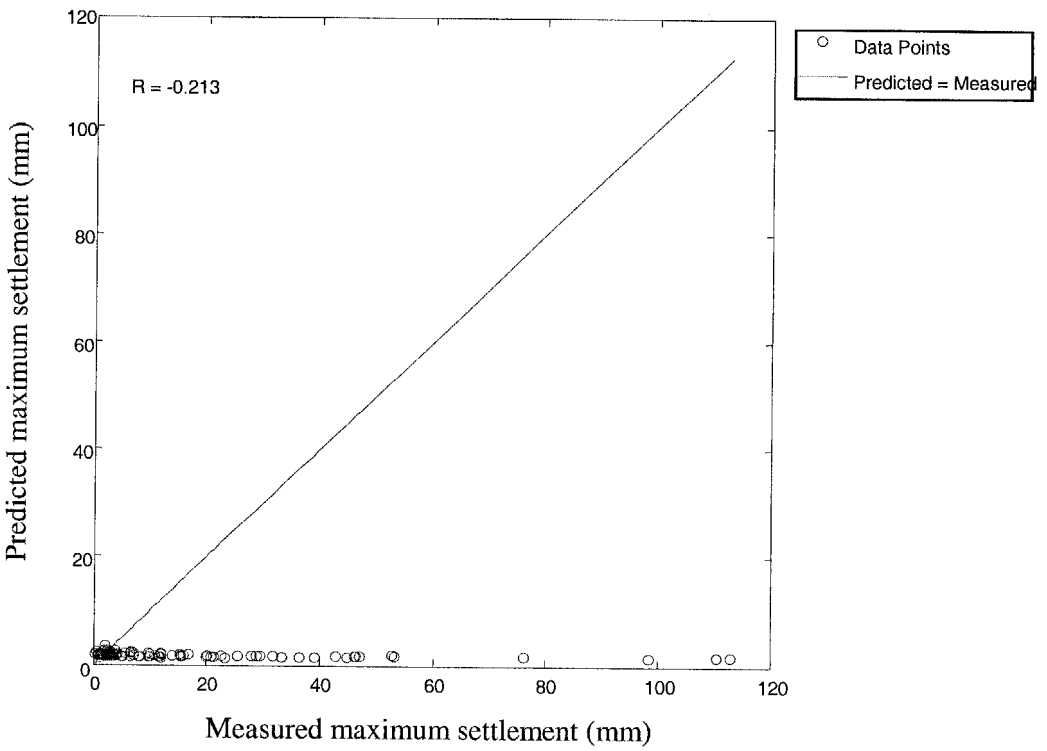


Figure 6.7: Comparison of measured and predicted settlements obtained using ANN for all data from contract C705

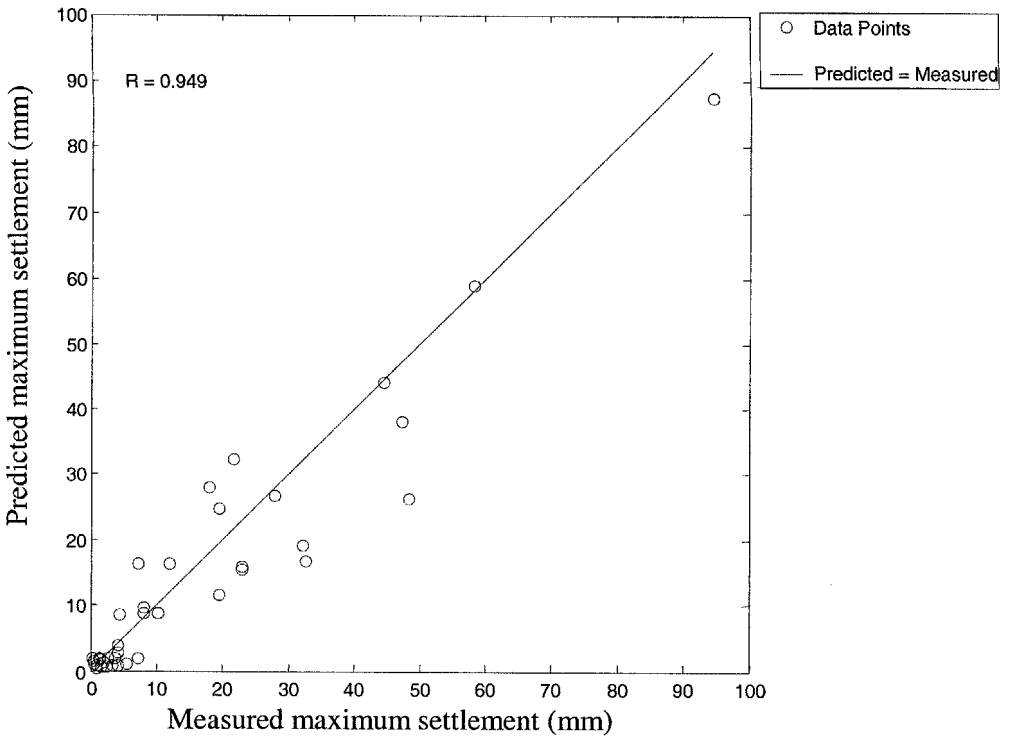


(a) ANN

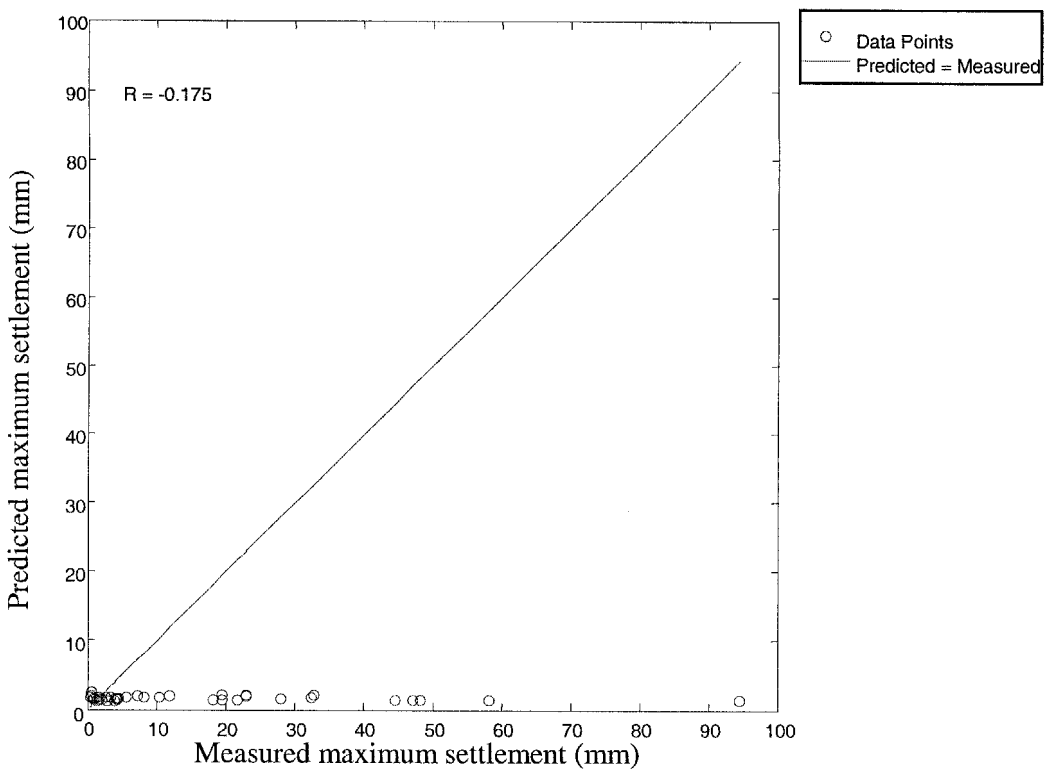


(b) Empirical Method

Figure 6.8: Comparison of measured and predicted maximum settlements for training set

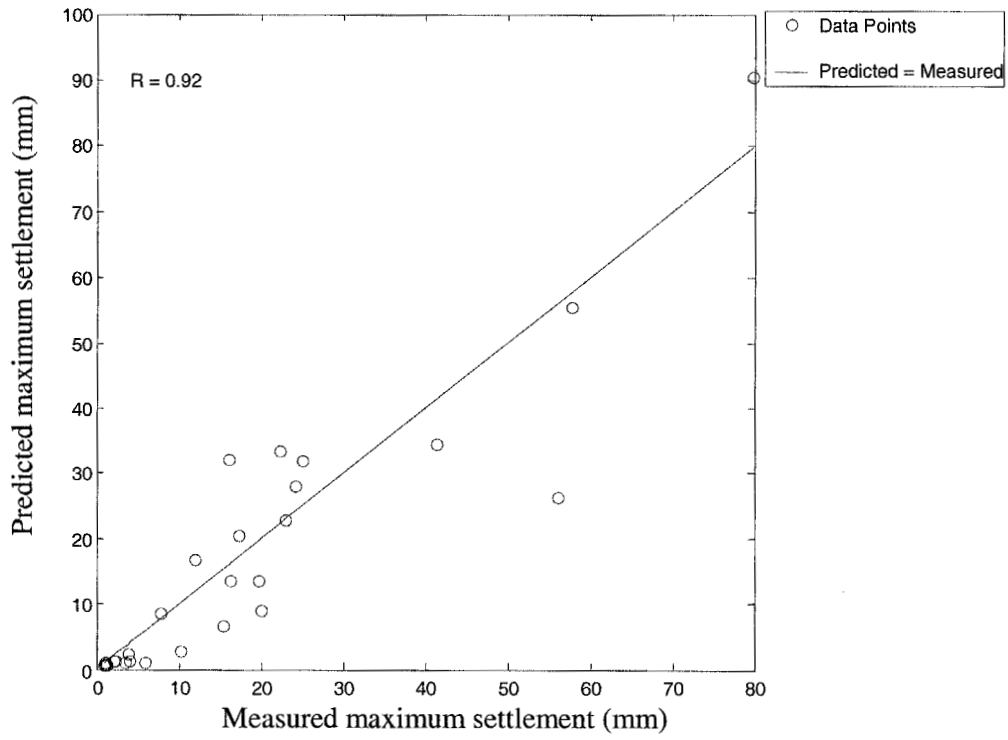


(a) ANN

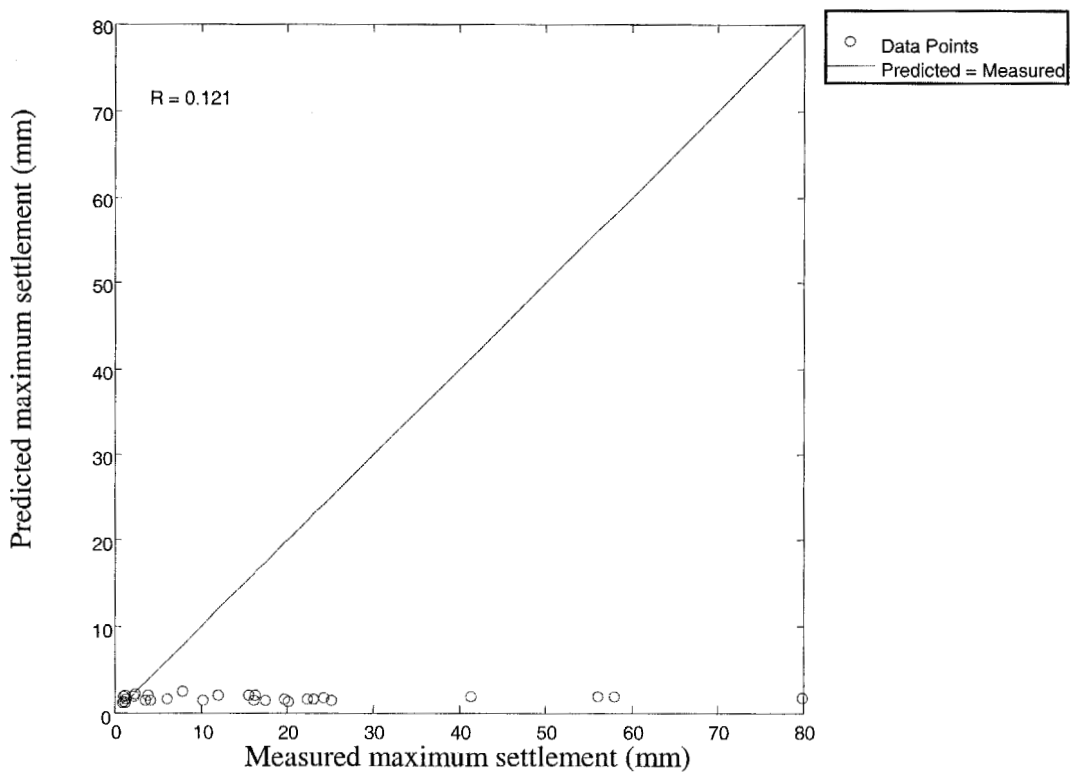


(b) Empirical Method

Figure 6.9: Comparison of measured and predicted maximum settlements for testing set



(a) ANN



(b) Empirical Method

Figure 6.10: Comparison of measured and predicted maximum settlements for validation set

6.11 Alternative ANN model for initial prediction of maximum settlement

From section 6.3 onwards, neural network model NN6 has been employed to predict the surface settlements induced by tunneling and the results mapped the field settlements quiet accurately. However, this network model might not be effective for quick initial prediction of surface settlements as it involves many input parameters i.e. 8 inputs. Furthermore, the input parameters such as earth pressure, advance rate, and grout pressure may not be available at the beginning of construction period. Hence, only soil related input parameters can be used to calculate the initial settlement predictions. In such circumstances, alternative network NN29 can be utilized as it requires only four inputs which are readily available from the site, namely cover (H), SPT1, SPT2, and moisture content (MC). Table 6.5 indicated that model NN29 generate results with high correlation coefficients for training, testing, and validation sets and relatively low error rates for testing and validation sets. Therefore, this model is acceptable for the initial predictions of surface settlement. The comparisons of measured and predicted settlements obtained using NN29 for training, testing and validation set are shown in Figures 6.11 to 6.13. The final weights and bias terms of model NN29 are presented in Table 6.32 and Table 6.33 respectively.

Table 6.32: Final weights of Model NN29

Hidden nodes	Weights				
	Cover (H)	SPT1	SPT2	MC	S_{max}
Hidden 1	-1.552	-0.743	0.508	4.053	3.268
Hidden 2	-0.943	-0.094	-0.141	0.415	0.785
Hidden 3	-0.162	1.228	-0.891	0.479	-1.124
Hidden 4	0.473	-0.170	-0.091	0.326	0.738
Hidden 5	-0.110	-0.037	-1.798	2.192	-1.251
Hidden 6	-1.030	1.458	-0.141	-1.401	-2.263
Hidden 7	-0.030	2.396	0.701	4.206	-3.496
Hidden 8	1.285	-0.036	-0.286	-0.138	1.109

Table 6.33: Bias terms of Model NN29

Hidden and Output nodes	Bias
Hidden 1	-1.981
Hidden 2	-1.038
Hidden 3	-0.436
Hidden 4	0.568
Hidden 5	-1.002
Hidden 6	1.128
Hidden 7	1.363
Hidden 8	-0.992
Output 1	-0.665

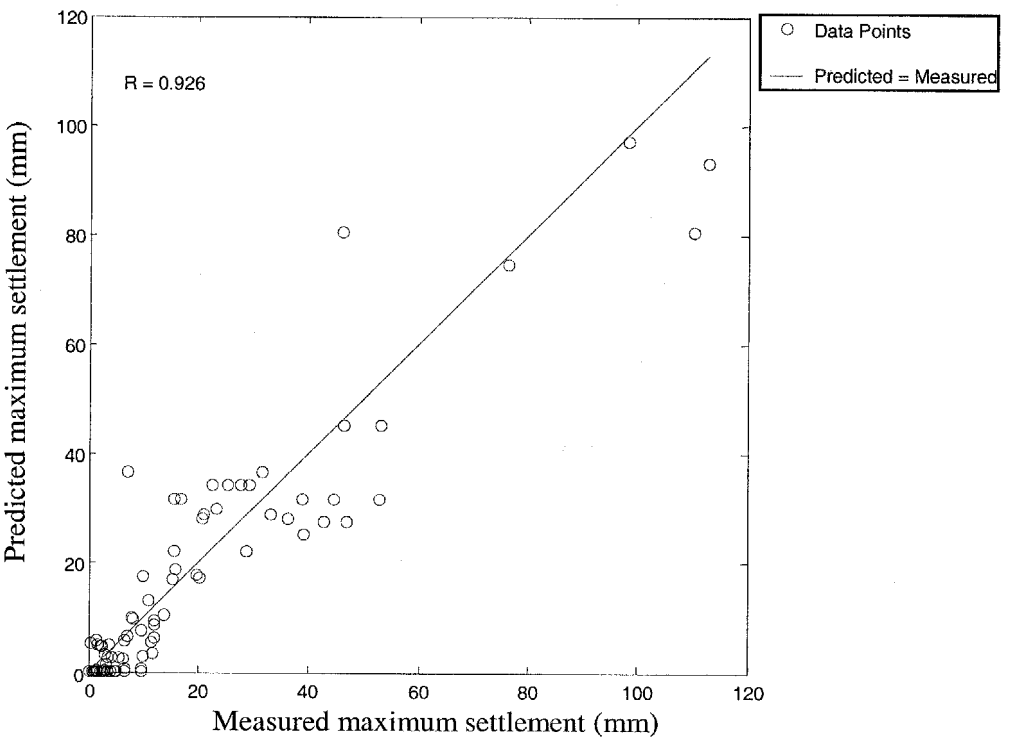


Figure 6.11: Comparison of measured and predicted maximum settlements for training set

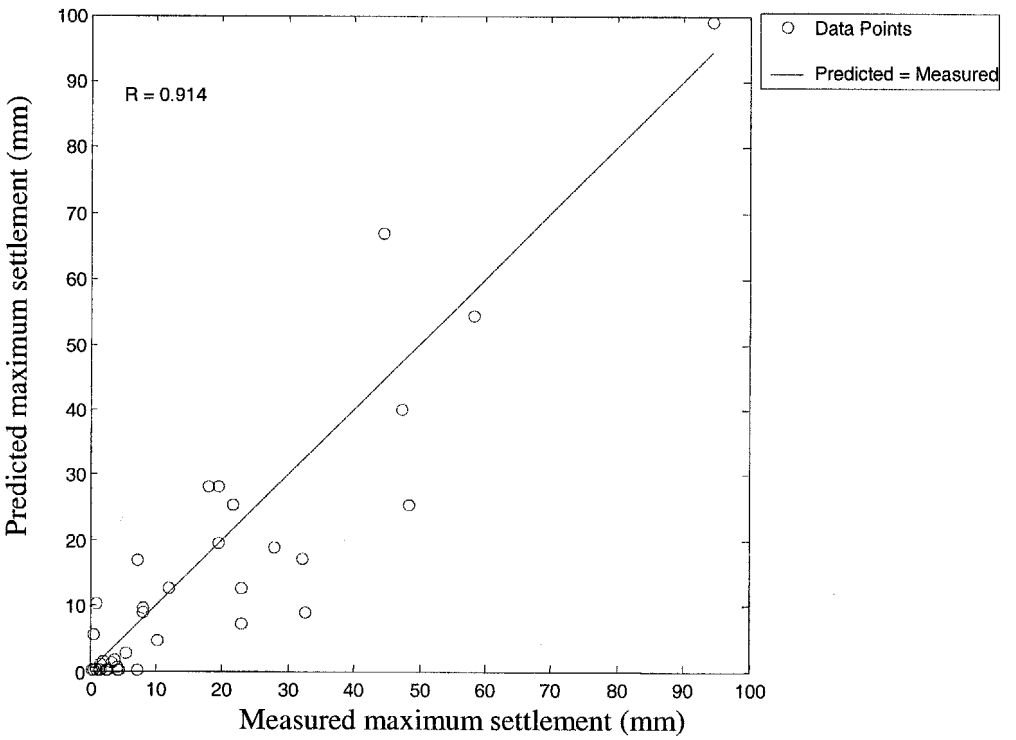


Figure 6.12: Comparison of measured and predicted maximum settlements for testing set

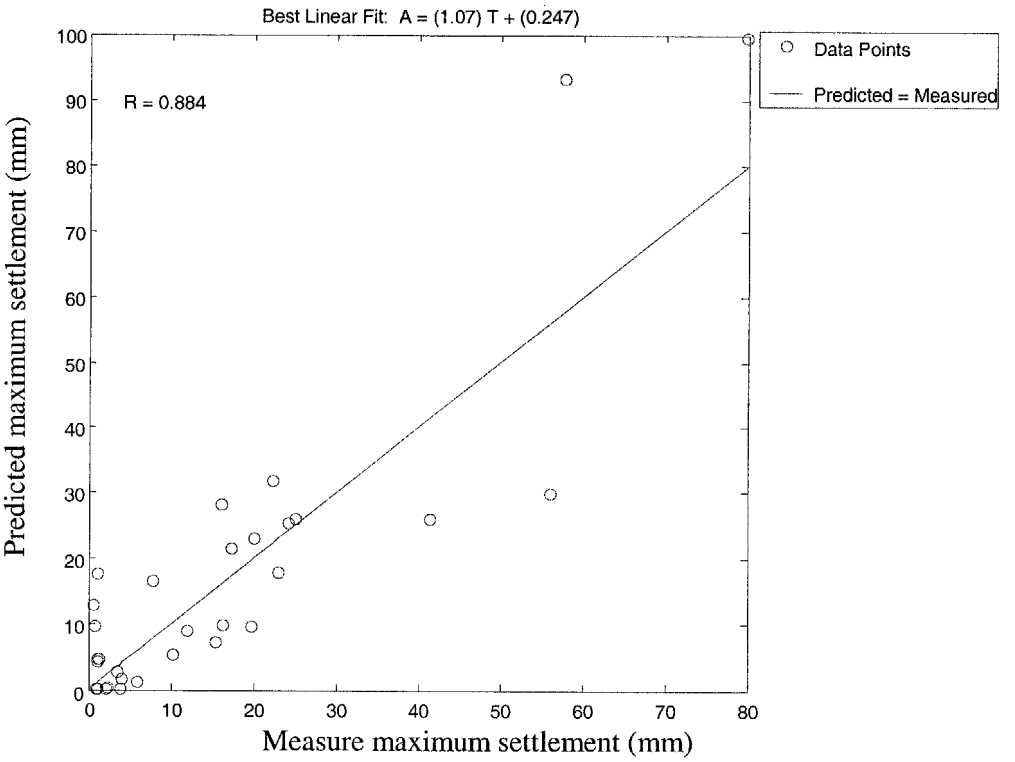


Figure 6.13: Comparison of measured and predicted maximum settlements for validation set

Chapter 7

Prediction of Maximum Surface Settlement (S_{max}) and Trough Width (i) by Multi Layer Perceptron

As explained in Chapter 4, settlement trough is characterized by two important parameters: the maximum surface settlement at the point above the tunnel centerline (S_{max}) and the width parameter (i) which is defined as the distance from the tunnel centerline to the inflection point of the trough. This chapter presents the development of neural network models to estimate both the maximum surface settlement and width parameter (i). In Chapter 6, the data set used to develop the networks is derived from the field reports database. However, the field reports did not provide information of volume loss which is an influential parameter for the prediction of both maximum surface settlement (S_{max}) and the trough width (i). Hence, for analysis in this chapter, the finite element software PLAXIS is utilized to generate the relevant input data for the training of the neural network model.

- **Overview of PLAXIS**

PLAXIS is a finite element program for geotechnical applications in which soil models are used to simulate the soil behavior. Its development started in 1987 at the Technical University of Delft as an initiative of the Dutch Department of Public Works and Water Management. Initially, it was developed to create an easy-to-use 2D finite element code for the analysis of river embankments on the soft soils of the lowlands of Holland. In later years, PLAXIS was extended to cover most other areas of geotechnical engineering. PLAXIS is used specifically for the analysis of deformation and stability in geotechnical engineering projects. It finds the applications in several studies such as submerged construction of an excavation, undrained river embankment, dry excavation using a tie back wall, settlement of circular footing on sand, and settlements due to tunnel construction. For the last application, PLAXIS requires a number of input

parameters which varies according to material model and material behavior. For the analysis of settlement, Mohr-Coulomb model is used to simulate the behavior of soil. This model involves five parameters, namely Young's modulus, E, Poisson's ratio, ν, cohesion, c, the friction angle, φ, and the dilatancy angle, ψ. PLAXIS provides a choice of three types of behavior for each soil model which incorporate the effect of pore water in the soil response: drained, undrained, and non-porous. Since we are interested in the short term settlements (settlements induced immediately after the passing of TBM machine), it is appropriate to set the material behavior to undrained. However, this setting requires that the effective model parameters should be entered, i.e. E' etc and not E_u etc. However, the field data, which is used to validate the neural network model, mostly provide the undrained parameters. Hence it is decided to use the Non-porous option instead to simulate the undrained behavior. In this setting, we can directly enter the undrained elastic properties $E = E_u$ and $\nu = \nu_u = 0.495$ in combination with the undrained strength properties $c = c_u$ and $\varphi = \varphi_u = 0°$. In this case a total stress analysis is performed and all pore pressures are set to zero.

There are two methods for numerical simulation using Plaxis namely method by relaxation factor and method by volume loss. For the simulation in this project, the method by volume loss has been adopted for the calculation of surface settlement trough. The input parameters used for PLAXIS in this project are presented in Table 7.1. For each parameter, several values are selected and they are used to generate the combinations comprising the six input parameters. The values are listed in Table 7.2.

Table 7.1: Input Parameters

No.	Inputs Parameters	Symbol	Units
1	Coefficient of lateral earth pressure	K_0	
2	Bulk unit weight	γ	kN/m^3
3	Cohesion	c	kN/m^2
4	Ratio of stiffness over cohesion	E/c	
5	Ratio of depth over diameter	H/D	
6	Volume Loss	VL	%

Table 7.2: Selected values for each input parameter

No.	Inputs Parameters	Values
1	K_0	0.5,1,2
2	γ	15,19,22
3	c	15,50,100,250,500
4	E/c	100,300,500
5	H/D	0.5,1,3,5
6	VL	0.1,1,3,10

The output of PLAXIS generation is the surface settlement curve from which the magnitude of horizontal and vertical settlements at each point on the surface can be obtained. However, the trough width parameter (i) of the settlement curve is not provided in the output report. In this case, Least-Square regression analysis was performed on the settlements measurements in order to obtain the width parameter (i) assuming Gauss settlement distribution. The tunnel diameter considered in the analysis is 6.0 m. The soil layer surrounding the tunnel is considered to have uniform properties described by the four input parameters (K_0, γ, c, E/c). From the initial mesh analysis carried out using the sample from field data, it is concluded that the adopted mesh in the numerical simulation is considered reasonable for shallow tunnels. The depth of 1D from tunnel invert is considered appropriate based on the result of sensitivity analysis carried out with different bottom boundary conditions. Figure 7.1 shows the geometry of the tunnel used in PLAXIS.

The total combinations which can be generated from parameter values in Table 7.2 amount to 2161. However, a large number of combinations can not be used since they produced either a collapsed soil body or bad surface settlement graph. This reduces the number of input combinations further to 1836. In the first analysis, network models are developed using three data sets i.e. training, testing and validation. In the second analysis, the input data is divided into training and testing sets only. The field data set is used as the validation set to examine the predictive capability of the optimum networks obtained in the first and second analysis.

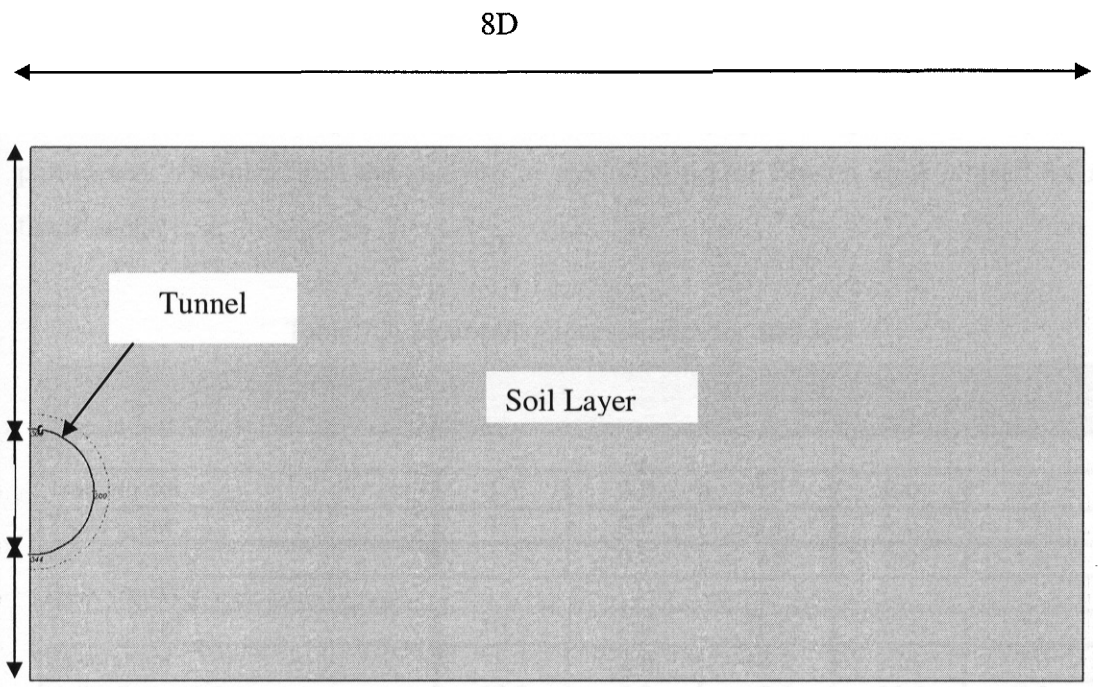


Figure 7.1: Geometry of the tunnel with indications of depth (H), diameter (D) and soil layer

7.1 Analysis using three sets of data: Training, Testing, and Validation

The input patterns are divided into three sets: training, testing and validation. 1031 patterns are used for training set, 441 patterns for testing set and 364 patterns for validation set. The three sets are divided in such a way that they are statistically consistent and thus represent the same statistical population. In order to achieve this, several random combinations of the training, testing and validation sets are tested until three statistically consistent data sets are obtained. The statistical properties of the training, testing and validation sets are shown in Table 7.3. It can be seen that, for each input variable, the statistical properties of the three sets are similar to each other. This is also confirmed by the results of null hypothesis tests in Table 7.4, which show that

he hypotheses of testing and validation sets for all input parameters passed the t-test nd the F-test.

The neural network tool available in MatLab is employed to develop the neural network models. Training is carried out using backpropagation algorithm and optimum parameters obtained from the analysis in section 6.4. One hidden layer is used for all he models.

Table 7.3: Input and output statistics for data sets

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
K_0					
Training set	1.1	0.6	0.5	2.0	1.5
Testing set	1.1	0.6	0.5	2.0	1.5
Validation set	1.1	0.6	0.5	2.0	1.5
Bulk Density					
Training set	18.6	2.9	15.0	22.0	7.0
Testing set	18.7	2.8	15.0	22.0	7.0
Validation set	18.8	2.9	15.0	22.0	7.0
Cohesion					
Training set	202.4	183.2	15.0	500.0	485.0
Testing set	193.4	172.2	15.0	500.0	485.0
Validation set	185.1	173.8	15.0	500.0	485.0
E/c					
Training set	295.3	164.3	100.0	500.0	400.0
Testing set	298.2	162.4	100.0	500.0	400.0
Validation set	297.3	163.6	100.0	500.0	400.0
H/D					
Training set	2.5	1.8	0.5	5.0	4.5
Testing set	2.5	1.7	0.5	5.0	4.5
Validation set	2.4	1.8	0.5	5.0	4.5
Volume Loss (%)					
Training set	3.4	3.9	0.1	10.0	9.9
Testing set	3.5	4.0	0.1	10.0	9.9
Validation set	3.3	3.8	0.1	10.0	9.9
Maximum Settlement (mm)					
Training set	16.4	27.2	0.3	212.6	212.2
Testing set	17.2	28.4	0.4	175.5	175.1
Validation set	17.6	29.3	0.5	211.0	210.5
Inflection point, i (m)					
Training set	15.4	7.9	4.6	66.9	62.3
Testing set	15.9	8.1	5.1	62.5	57.4
Validation set	15.3	7.7	5.1	59.6	54.5

Table 7.4: Null hypothesis tests

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
K_0								
Testing	-1.4	-2.0	2.0	Accept	1.0	0.9	1.2	Accept
Validation	-1.3	-2.0	2.0	Accept	1.1	0.8	1.2	Accept
Bulk Density								
Testing	-0.8	-2.0	2.0	Accept	1.0	0.9	1.2	Accept
Validation	-1.3	-2.0	2.0	Accept	1.0	0.8	1.2	Accept
Cohesion								
Testing	0.9	-2.0	2.0	Accept	1.1	0.9	1.2	Accept
Validation	1.6	-2.0	2.0	Accept	1.1	0.8	1.2	Accept
E/c								
Testing	-0.3	-2.0	2.0	Accept	1.0	0.9	1.2	Accept
Validation	-0.2	-2.0	2.0	Accept	1.0	0.8	1.2	Accept
H/D								
Testing	0.1	-2.0	2.0	Accept	1.0	0.9	1.2	Accept
Validation	0.6	-2.0	2.0	Accept	1.0	0.8	1.2	Accept
Volume Loss (%)								
Testing	-0.5	-2.0	2.0	Accept	0.9	0.9	1.2	Accept
Validation	0.6	-2.0	2.0	Accept	1.0	0.8	1.2	Accept
Maximum Settlement (mm)								
Testing	-0.6	-2.0	2.0	Accept	0.9	0.9	1.2	Accept
Validation	-0.7	-2.0	2.0	Accept	0.9	0.8	1.2	Accept
Inflection point, I (m)								
Testing	-1.1	-2.0	2.0	Accept	0.9	0.9	1.2	Accept
Validation	0.2	-2.0	2.0	Accept	1.1	0.8	1.2	Accept

7.1.1. Networks with two output neurons

In this analysis, a neural network model with two output neurons is trained to yield the predictions of maximum surface settlement (S_{max}) and trough width (i) simultaneously. Network of different hidden neurons are tested and the results in term of correlations coefficients and error rate are displayed in Table 7.5 and Table 7.6. Neural network model with 12 hidden neurons is considered optimal as it exhibits lowest error rate coupled with high correlation coefficient for both testing and validation sets.

Table 7.5: Performance of ANN model with two output neurons for S_{max} prediction

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN1	2	0.88	0.86	0.89	7.92	8.13	8.26
NN2	4	0.93	0.88	0.94	5.90	7.30	6.04
NN3	6	0.94	0.90	0.95	5.24	6.65	5.32
NN4	8	0.93	0.91	0.93	6.42	7.72	6.67
NN5	10	0.91	0.86	0.92	7.18	8.81	7.33
NN6	12	0.94	0.91	0.94	5.22	6.57	5.48
NN7	14	0.96	0.93	0.93	4.50	5.76	5.27
NN8	16	0.95	0.91	0.93	5.25	7.05	6.23

Table 7.6: Performance of ANN model with two output neurons for (i) prediction

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
NN1	2	0.92	0.87	0.91	1.99	2.26	2.04
NN2	4	0.94	0.88	0.93	1.63	2.13	1.92
NN3	6	0.96	0.91	0.94	1.36	1.75	1.74
NN4	8	0.96	0.90	0.92	1.35	1.86	1.94
NN5	10	0.96	0.88	0.93	1.48	2.11	2.04
NN6	12	0.97	0.92	0.94	1.33	1.71	1.73
NN7	14	0.96	0.89	0.93	1.28	1.87	1.85
NN8	16	0.97	0.88	0.94	1.20	2.08	1.74

- **Comparison of training algorithms**

In this section, faster training algorithms discussed in Chapter 2 are used to develop the network. The networks are trained using the same data sets as in the gradient descent method. The results of all the training algorithms including gradient descent method are tabulated in Table 7.7 and Table 7.8. Twelve neurons are used in the hidden layer as it is the optimum number in the case of the gradient descent method. The network has two neurons in the output layer meant for the simultaneous prediction of maximum

settlement (S_{max}) and inflection point (i). Table 7.7 and Table 7.8 show that the results of all faster training algorithms are better than the result of gradient descent method. The trained network produces more accurate predictions as indicated by higher correlation coefficient and lower error rate for the training, testing, and validation sets. This applies for both predictions of maximum settlement (S_{max}) and inflection point (i). Among the improved networks, the network trained using the One-Step Secant method (OSS) is considered to be optimum as its overall result is better than those of other networks. Figures 7.2 to 7.7 present the comparisons of measured against predicted S_{max} and (i) obtained using the OSS network with respect to the training, testing and validation set.

Table 7.7: Performance of ANN model using various training algorithms for S_{max} prediction

Algorithm	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
LM	14	0.96	0.92	0.91	5.61	7.59	7.31
GDM		0.96	0.93	0.93	4.50	5.76	5.27
BFG		0.99	0.96	0.96	2.94	4.18	4.33
RP		0.98	0.95	0.94	3.65	5.40	5.18
SCG		0.99	0.98	0.96	2.79	3.39	4.16
CGB		0.99	0.96	0.97	2.75	4.12	4.04
CGF		0.98	0.96	0.96	2.74	3.91	3.70
CGP		0.99	0.97	0.98	2.36	3.43	2.94
OSS		0.99	0.98	0.97	2.59	3.06	3.70
GDX		0.98	0.96	0.95	3.46	4.36	4.71

Table 7.8: Performance of ANN model using various training algorithms for (i) prediction

Algorithm	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
LM	14	0.98	0.94	0.96	1.24	1.74	1.51
GDM		0.96	0.89	0.93	1.28	1.87	1.85
BFG		0.99	0.96	0.98	0.88	1.21	1.03
RP		0.98	0.96	0.97	0.98	1.39	1.25
SCG		0.99	0.97	0.98	0.74	1.04	0.94
CGB		0.99	0.95	0.98	0.83	1.26	1.06
CGF		0.99	0.95	0.98	0.87	1.31	1.06
CGP		0.98	0.94	0.97	0.90	1.36	1.15
OSS		0.99	0.97	0.98	0.85	1.13	0.98
GDX		0.98	0.94	0.97	1.09	1.49	1.33

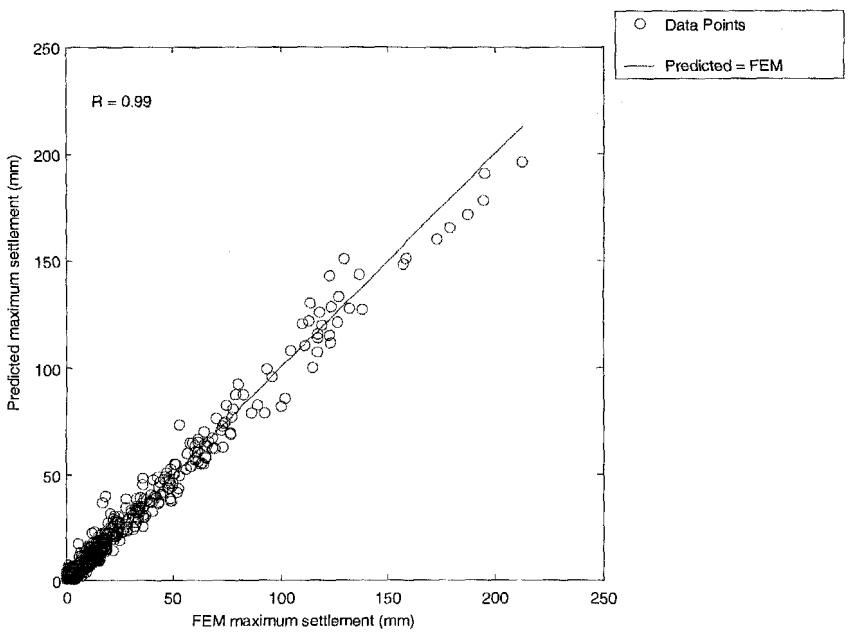


Figure 7.2: Predicted vs. FEM maximum settlements for training set

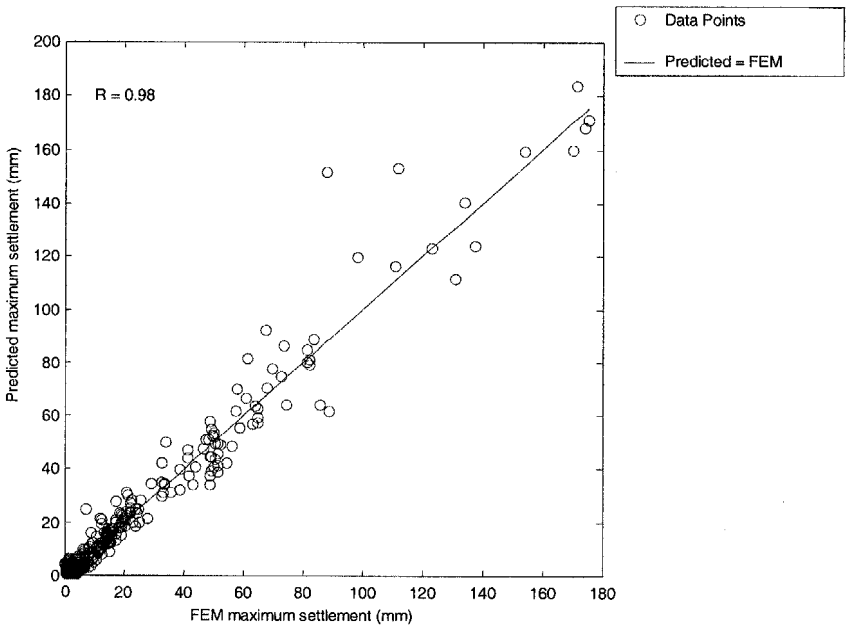


Figure 7.3: Predicted vs. FEM maximum settlements for testing set

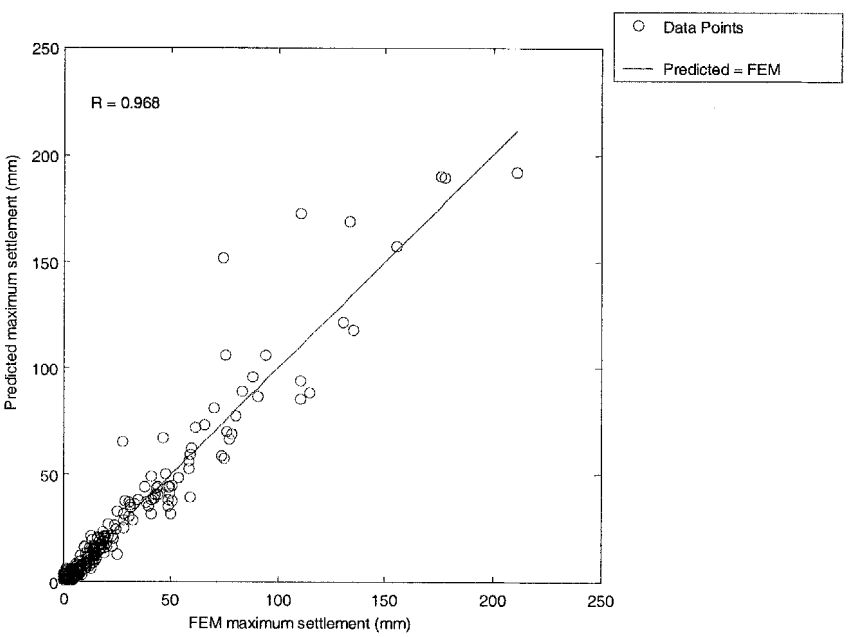


Figure 7.4: Predicted vs. FEM maximum settlements for validation set

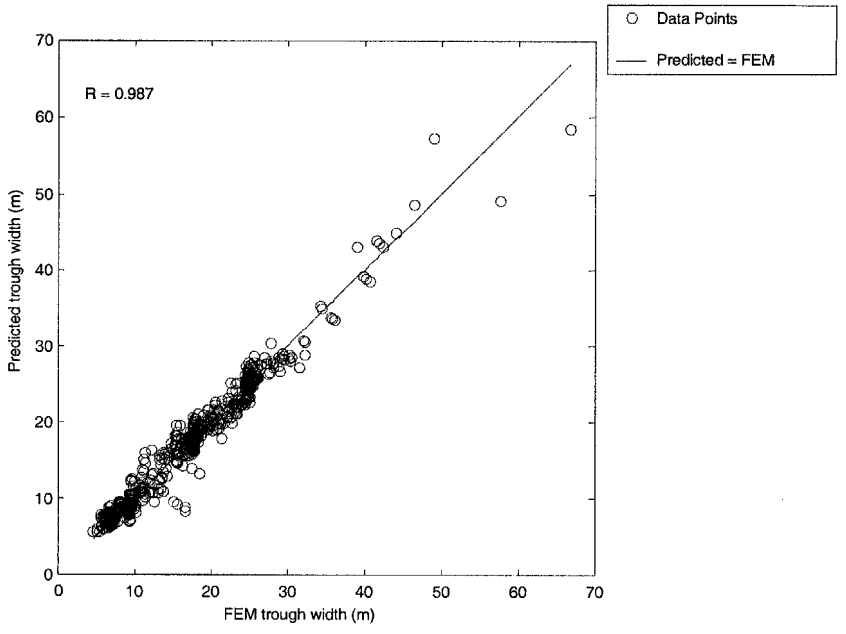


Figure 7.5: Predicted vs. FEM trough width for training set

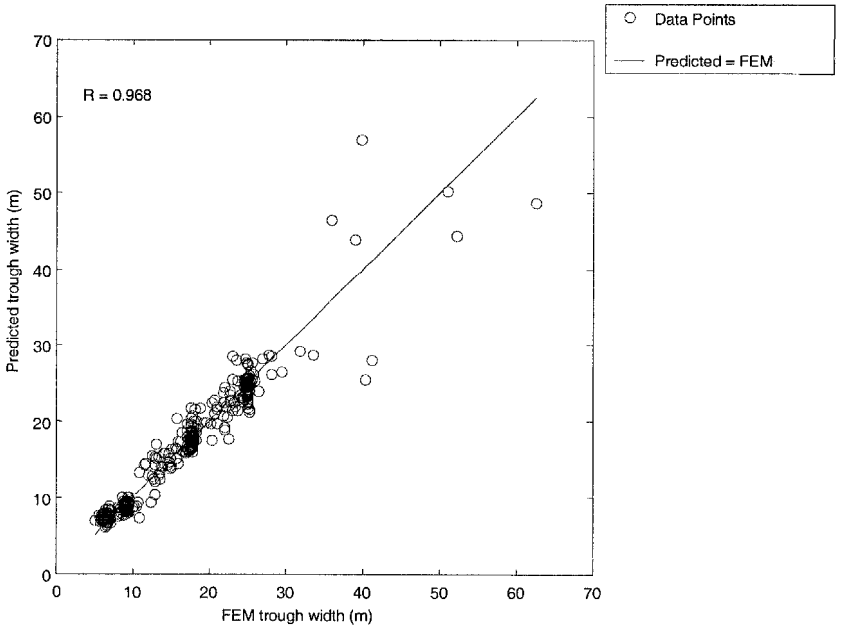


Figure 7.6: Predicted vs. FEM trough width for testing set

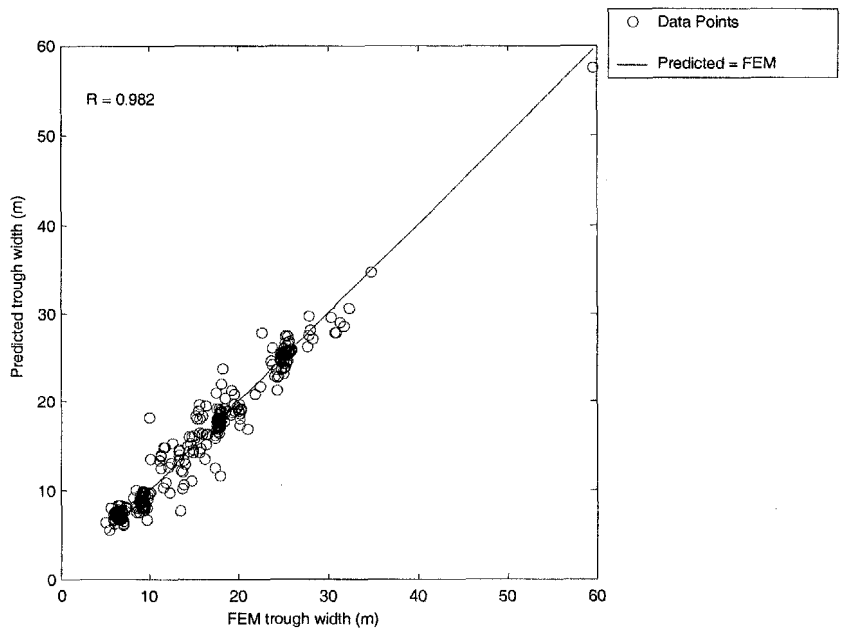


Figure 7.7: Predicted vs. FEM trough width for validation set

7.1.2. Networks with one output neuron

Two separate networks are developed whereby one network is used to predict S_{max} and the other to predict trough width (i). The results for the two networks tested with different hidden neurons are shown in Table 7.9 and 7.10. Network used for S_{max} prediction is labeled as ANNS followed by the hidden neuron number while that used for trough width (i) prediction is labeled as ANNi. Networks with 12 hidden neurons give optimal predictions of S_{max} as indicated by the maximum correlation coefficients and minimum error rates for both testing and validation sets. Network for (i) prediction is optimum when it used 6 neurons in the hidden layer. It is also observed that the two separate networks perform better than the single network used to predict S_{max} and (i) simultaneously. This is pointed out by higher correlation coefficients and lower error rates when the two separate networks are used. In this case, one network is independent

of another network. It can freely develop itself to higher degree of accuracy without having to take into account the conditions of another output. In a single network with combined S_{max} and (i) outputs, this is not the case as the training is carried out to produce a set of weights which will give optimal results for both outputs at once.

Table 7.9: Performance of ANN models for S_{max} prediction

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
ANNS	2	0.89	0.88	0.91	6.86	7.63	7.18
ANNS	4	0.93	0.91	0.94	5.21	5.63	5.23
ANNS	6	0.96	0.93	0.92	4.52	5.88	5.95
ANNS	8	0.96	0.94	0.94	4.73	6.00	5.61
ANNS	10	0.95	0.92	0.95	5.06	6.12	5.39
ANNS	12	0.97	0.92	0.94	4.21	5.85	5.44
ANNS	14	0.96	0.93	0.95	4.70	6.03	5.40

Table 7.10: Performance of ANN models for trough width (i) prediction

Model	No. of Hidden Neurodes	Correlation Coefficient			Error Rate (unscaled)		
		Training	Testing	Validation	Training	Testing	Validation
ANNi	2	0.96	0.90	0.92	1.44	1.87	1.90
ANNi	4	0.97	0.94	0.94	1.40	1.79	1.97
ANNi	6	0.97	0.92	0.96	1.18	1.55	1.48
ANNi	8	0.97	0.93	0.94	1.31	1.73	1.80
ANNi	10	0.96	0.89	0.92	1.43	1.95	2.02
ANNi	12	0.98	0.94	0.95	1.15	1.59	1.62

- **Comparison of training algorithms**

The two optimum networks for the predictions of S_{max} and (i) are improved further using the faster training algorithms. The results of all the training algorithms including gradient descent method are tabulated in Table 7.11 and Table 7.12. For speed

comparison, number of cycles required to reach the minimum error for each training algorithm is listed in the Tables as well. The network trained using the Polak-Ribiére Conjugate Gradient (CGP) is considered to be optimum for the prediction of S_{max} while network trained using Fletcher-Powell Conjugate Gradient (CGF) is optimal for trough width (i) prediction. Thus the two optimum networks are labeled as ANNS12CGP and ANNi6CGF. From the comparison of results, it is shown that the networks ANNS12CGP and ANNi6CGF performed relatively better than the improved OSS network used for simultaneous prediction of S_{max} and (i). This is indicated by higher coefficients of correlation and lower error rates for the training, testing, and validation sets when the two improved networks are used to predict S_{max} and (i) separately. Hence these two networks will be used in section 7.3 to provide the predictions of the field data. From the comparison of speed, it is deduced that training algorithm Levenberg-Marquardt exhibits the fastest convergence for both predictions of S_{max} and (i) as indicated by the lowest number of cycles required to reach minimum error. As expected, training algorithm Gradient-Descent with Momentum (GDM) requires the most number of cycles to reach minimum error. The optimum networks, ANNS12CGP and ANNi6CGF, also display fast convergence for the respective prediction. The variable learning rate algorithm (GDX) is usually much slower than the other faster methods as shown in the prediction of S_{max}. However for (i) prediction, it is faster than training algorithms Resilient Backpropagation (RP) and Scaled Conjugate Gradient (SCG). Figures 7.8 to 7.10 present the comparison of measured against predicted maximum settlements of the training, testing and validation set for ANNS12CGP, while the comparison of measured against predicted trough width (i) for ANNi6CGF are shown in Figures 7.11 to 7.13 .

Table 7.11: Performance of ANNS12 model using various training algorithms

S_{max}

Algorithm	Model	No of. cycles	Correlation Coefficient			Error Rate (unscaled)		
			Training	Testing	Validation	Training	Testing	Validation
LM	ANNS12	23	0.98	0.95	0.94	3.26	5.33	5.61
GDM		204245	0.97	0.92	0.94	4.21	5.85	5.44
BFG		2800	0.99	0.95	0.96	2.06	3.69	3.35
RP		4172	0.99	0.95	0.93	3.05	5.01	5.65
SCG		6887	0.99	0.96	0.97	1.94	3.69	3.56
CGB		3509	0.99	0.97	0.96	2.06	3.38	4.16
CGF		5627	0.99	0.95	0.96	2.07	3.77	3.39
CGP		1279	0.99	0.98	0.97	1.96	2.94	3.31
OSS		2350	0.99	0.95	0.97	1.74	3.36	3.14
GDX		68904	0.98	0.97	0.94	3.26	4.18	5.19

Table 7.12: Performance of ANNi6 model using various training algorithms

Trough width (i)

Algorithm	Model	No of. cycles	Correlation Coefficient			Error Rate (unscaled)		
			Training	Testing	Validation	Training	Testing	Validation
LM	ANNi6	34	0.98	0.92	0.97	0.95	1.42	1.12
GDM		89571	0.97	0.92	0.96	1.18	1.55	1.48
BFG		1895	0.98	0.96	0.96	1.20	1.40	1.46
RP		31982	0.99	0.97	0.98	0.87	1.15	1.02
SCG		17402	0.98	0.96	0.98	1.02	1.34	1.18
CGB		1869	0.98	0.93	0.96	1.16	1.58	1.46
CGF		2893	0.99	0.96	0.98	0.81	1.10	0.96
CGP		1129	0.99	0.96	0.98	0.86	1.12	0.98
OSS		3713	0.99	0.97	0.98	0.89	1.12	1.11
GDX		15391	0.98	0.97	0.98	1.01	1.22	1.18

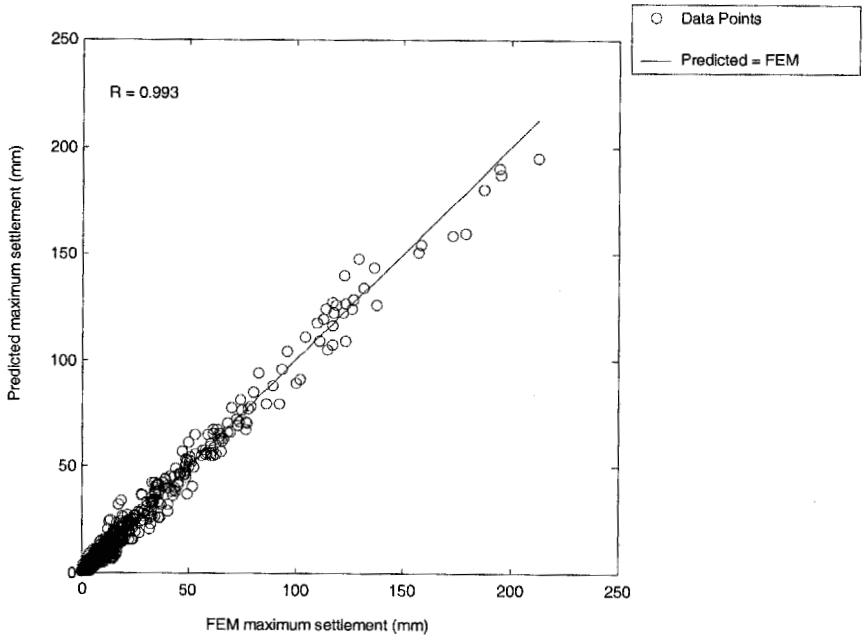


Figure 7.8: Predicted vs. FEM maximum settlements for training set (ANNS12CGP)

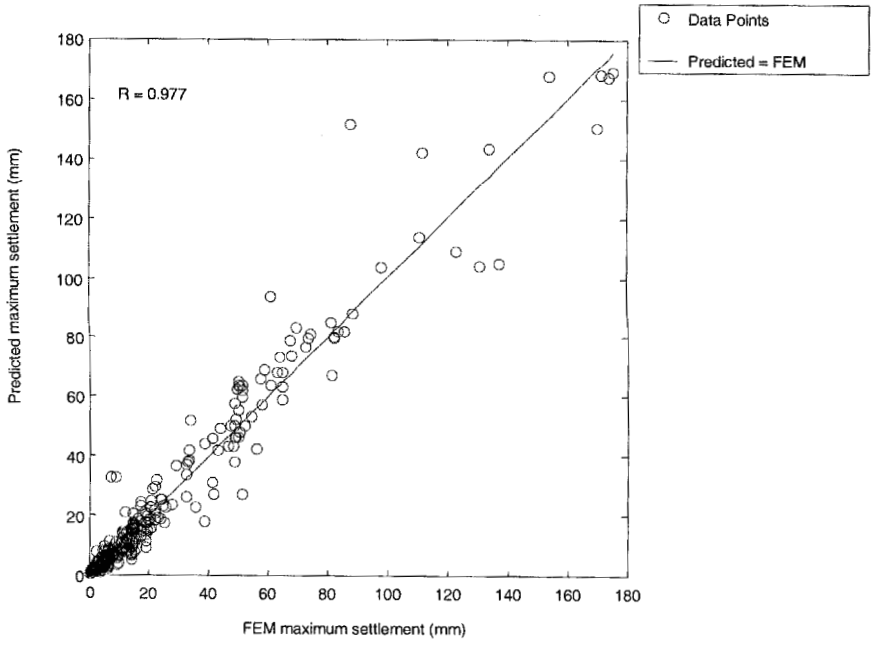


Figure 7.9: Predicted vs. FEM maximum settlements for testing set (ANNS12CGP)

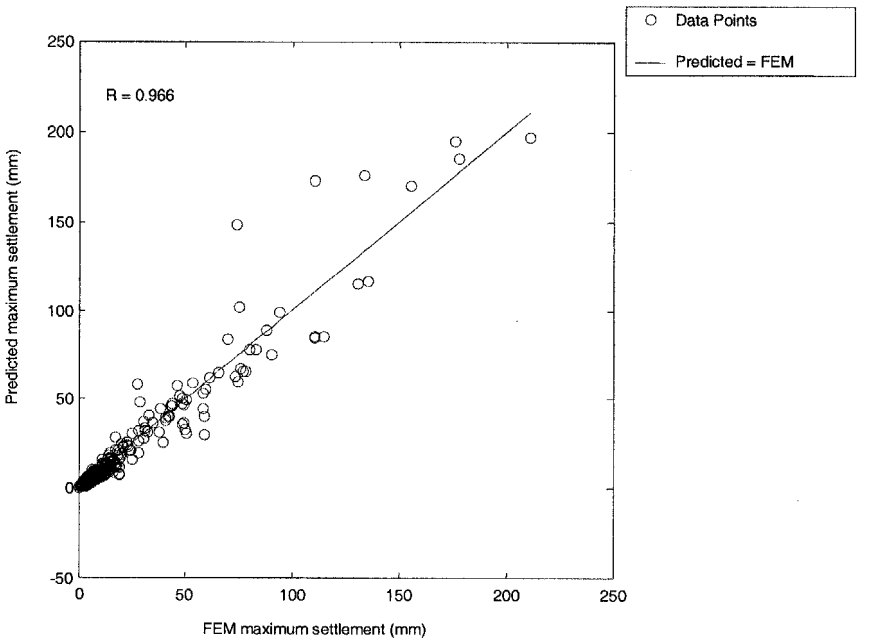


Figure 7.10: Predicted vs. FEM maximum settlements for validation set (ANNS12CGP)

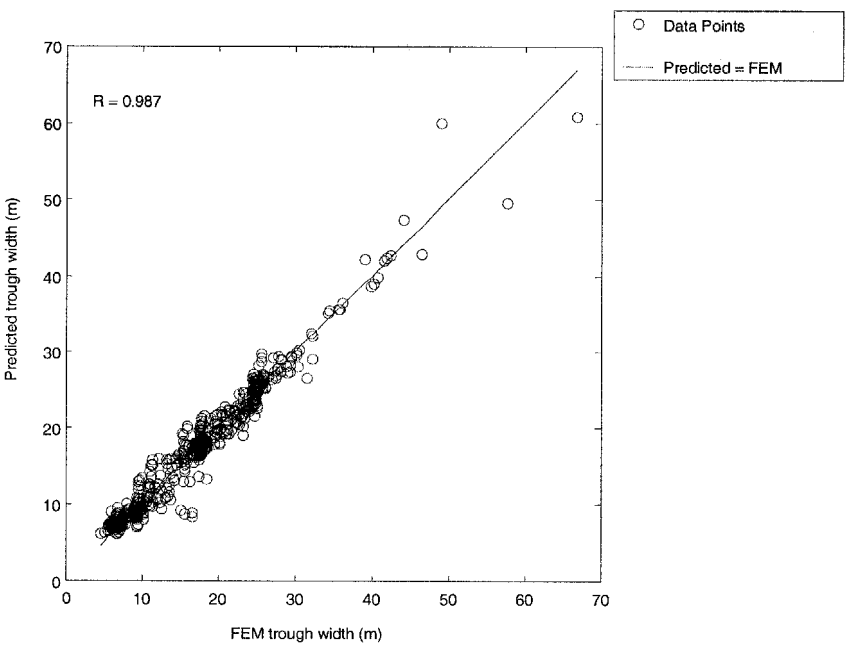


Figure 7.11: Predicted vs. FEM trough widths for training set (ANNi6CGF)

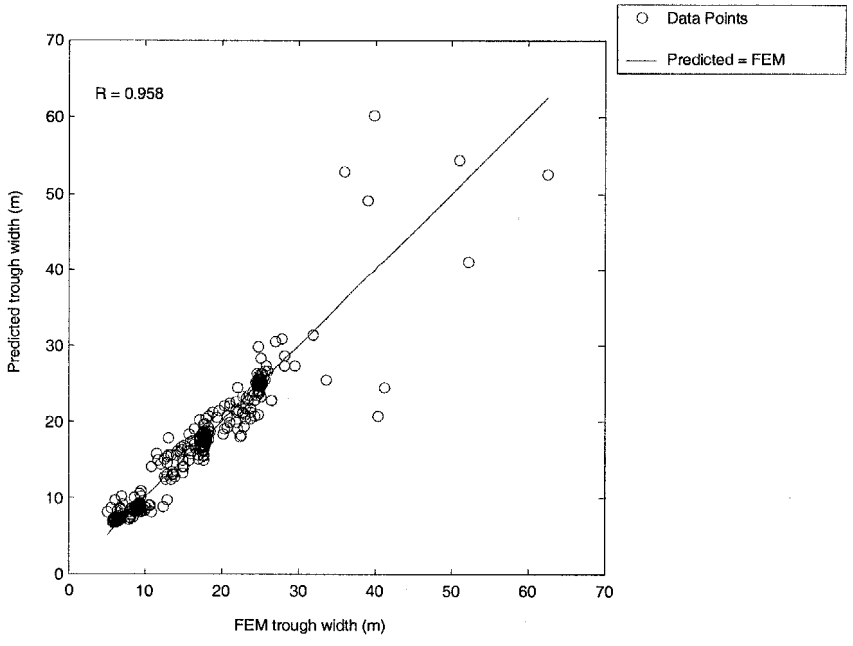


Figure 7.12: Predicted vs. FEM trough widths for testing set (ANNi6CGF)

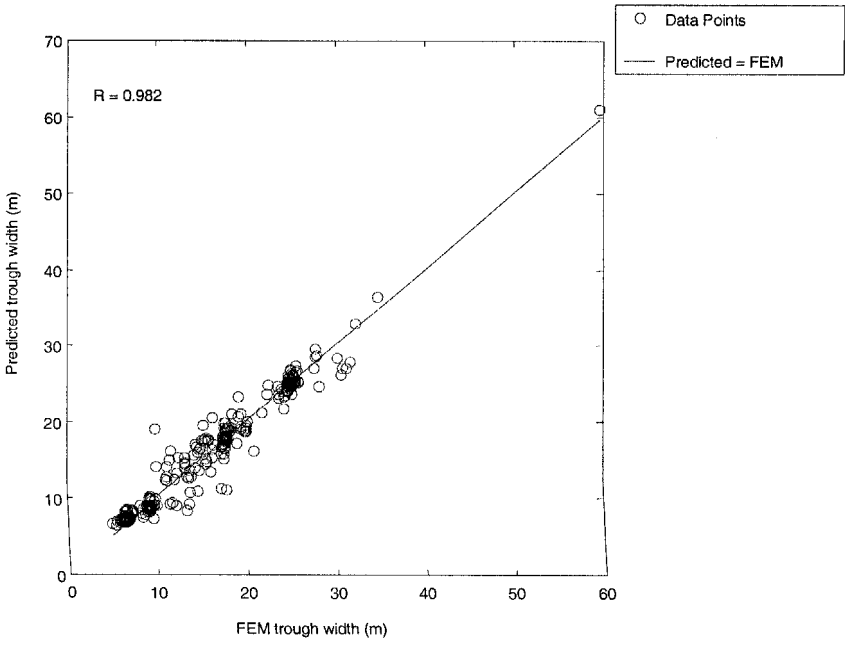


Figure 7.13: Predicted vs. FEM trough widths for validation set (ANNi6CGF)

7.2 Analysis using two sets of data: Training and Testing

The input patterns are divided into two sets: training and testing. 1395 patterns are used for training set and 441 patterns for testing set. The two sets are divided in such a way that they are statistically consistent and thus represent the same statistical population. In order to achieve this, several random combinations of the training and testing sets are tested until two statistically consistent data sets are obtained. The statistical properties of the training and testing sets are shown in Table 7.13. It can be seen that, for each input variable, the statistical properties of training and testing sets are similar to each other. This is also confirmed by the results of null hypothesis tests in Table 7.14, which show that the hypotheses of testing set for all input parameters passed the t-test and the F-test.

Table 7.13: Input and output statistics for data sets

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
K_0					
Training set	1.07	0.58	0.50	2.00	1.5
Testing set	1.10	0.60	0.50	2.00	1.5
Bulk Unit Weight (kN/m^3)					
Training set	18.64	2.92	15.00	22.00	7.0
Testing set	18.72	2.85	15.00	22.00	7.0
Cohesion (kN/m^2)					
Training set	197.87	180.87	15.00	500.00	485.0
Testing set	193.37	172.15	15.00	500.00	485.0
E/c					
Training set	295.84	164.09	100.00	500.00	400.0
Testing set	298.19	162.36	100.00	500.00	400.0
H/D					
Training set	2.49	1.78	0.50	5.00	4.5
Testing set	2.49	1.75	0.50	5.00	4.5
Volume Loss (%)					
Training set	3.36	3.89	0.10	10.00	9.9
Testing set	3.51	4.05	0.10	10.00	9.9
Maximum Settlement (mm)					
Training set	16.67	27.77	0.35	212.58	212.2
Testing set	17.23	28.37	0.37	175.50	175.1
Inflection point, i (m)					
Training set	15.42	7.84	4.62	66.91	62.3
Testing set	15.94	8.14	5.11	62.55	57.4

Table 7.14: Null hypothesis tests for testing set

Variable and data sets	t-value	Lower critical value	Upper critical value	t-test	F-value	Lower critical value	Upper critical value	F-test
K_0								
Testing	-1.05	-1.96	1.96	Accept	0.93	0.86	1.17	Accept
Bulk Density								
Testing	-0.48	-1.96	1.96	Accept	1.05	0.86	1.17	Accept
Cohesion								
Testing	0.46	-1.96	1.96	Accept	1.10	0.86	1.17	Accept
E/c								
Testing	-0.26	-1.96	1.96	Accept	1.02	0.86	1.17	Accept
H/D								
Testing	-0.06	-1.96	1.96	Accept	1.04	0.86	1.17	Accept
Volume Loss (%)								
Testing	-0.71	-1.96	1.96	Accept	0.92	0.86	1.17	Accept
Maximum Settlement (mm)								
Testing	-0.36	-1.96	1.96	Accept	0.96	0.86	1.17	Accept
Inflection point, I (m)								
Testing	-1.22	-1.96	1.96	Accept	0.93	0.86	1.17	Accept

7.2.1. Networks with two output neurons

In this analysis, a neural network model with two output neurons is trained to yield the predictions of maximum surface settlement (S_{max}) and trough width (i) simultaneously. Network of different hidden neurons are tested and the results in term of correlations coefficients and error rate are displayed in Tables 7.15 and 7.16. Neural network model with 12 hidden neurons is considered optimal as it exhibits lowest error rate coupled with high correlation coefficient for both training and testing sets.

Table 7.15: Performance of ANN model with two output neurons for S_{max} prediction

Model	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
NN1	2	0.89	0.87	8.10	8.07
NN2	4	0.92	0.89	6.61	7.51
NN3	6	0.92	0.88	6.52	7.96
NN4	8	0.95	0.91	4.98	5.93
NN5	10	0.94	0.91	5.28	6.16
NN6	12	0.95	0.91	4.85	5.91
NN7	14	0.95	0.89	4.70	6.37

Table 7.16: Performance of ANN model with two output neurons for (i) prediction

Model	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
NN1	2	0.92	0.87	1.98	2.22
NN2	4	0.93	0.88	1.70	1.98
NN3	6	0.95	0.90	1.54	1.87
NN4	8	0.95	0.92	1.56	1.74
NN5	10	0.97	0.95	1.39	1.56
NN6	12	0.96	0.92	1.39	1.65
NN7	14	0.97	0.92	1.32	1.71

• Comparison of training algorithms

As in the analysis of 3 data sets, the above optimum networks for the simultaneous predictions of S_{max} and (i) are improved further using the faster training algorithms. The results of all the algorithms training including gradient descent method are tabulated in Tables 7.17 and 7.18. The network trained using the BFGS Quasi Newton (BFG) method is considered to be optimum as its error rates for testing set are the lowest. Figures 7.14 to 7.17 present the comparisons of measured against predicted S_{max} and (i) obtained using the BFG network for the training and testing set.

Table 7.17: Performance of ANN model using various training algorithms for S_{max} prediction

Algorithm	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
LM	12	0.97	0.97	3.81	3.84
GDM		0.95	0.91	4.85	5.91
BFG		0.99	0.97	2.27	3.02
RP		0.98	0.97	2.38	3.43
SCG		0.98	0.97	2.66	3.20
CGB		0.98	0.97	3.21	3.58
CGF		0.97	0.95	3.51	4.36
CGP		0.98	0.97	2.51	3.47
OSS		0.98	0.97	3.34	3.63
GDX		0.98	0.96	3.25	4.03

Table 7.18: Performance of ANN model using various training algorithms for (i) prediction

Algorithm	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
LM	12	0.99	0.97	0.91	1.10
GDM		0.96	0.92	1.39	1.65
BFG		0.98	0.97	0.88	1.07
RP		0.98	0.97	1.16	1.29
SCG		0.98	0.96	1.02	1.20
CGB		0.98	0.97	0.95	1.18
CGF		0.98	0.96	1.13	1.33
CGP		0.98	0.96	1.06	1.21
OSS		0.98	0.97	1.08	1.29
GDX		0.98	0.96	1.10	1.33

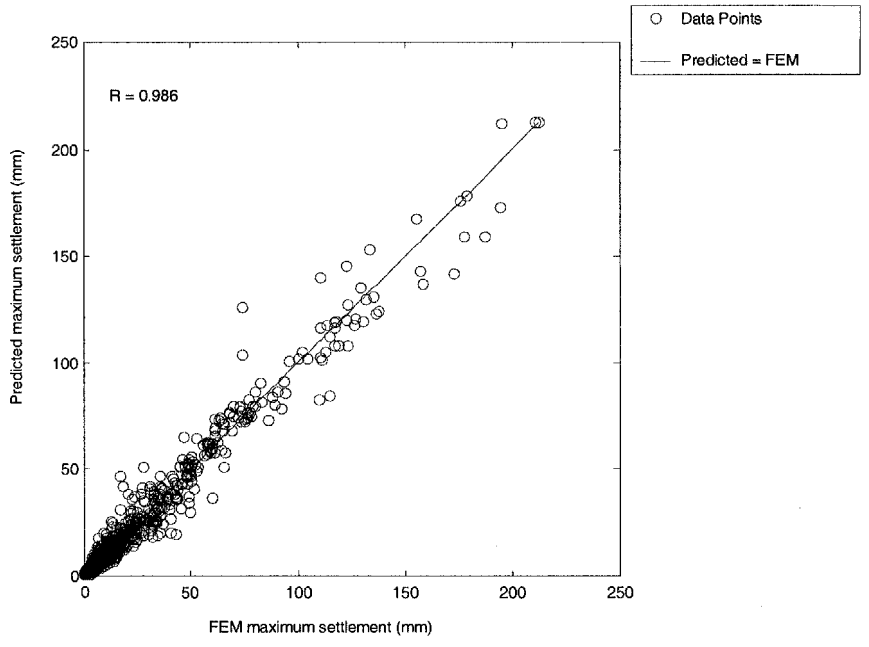


Figure 7.14: Predicted vs. FEM maximum settlements for training set

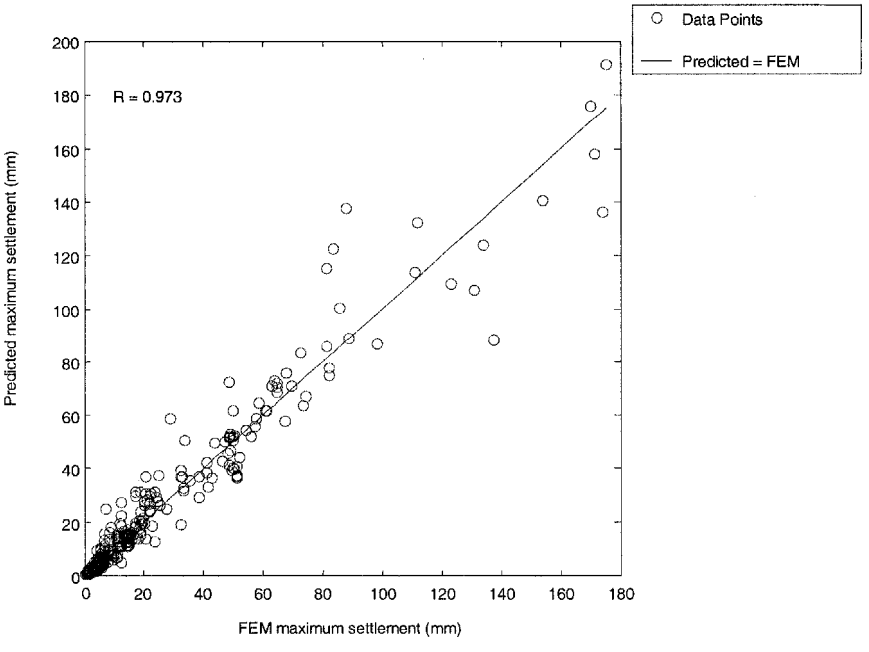


Figure 7.15: Predicted vs. FEM maximum settlements for testing set

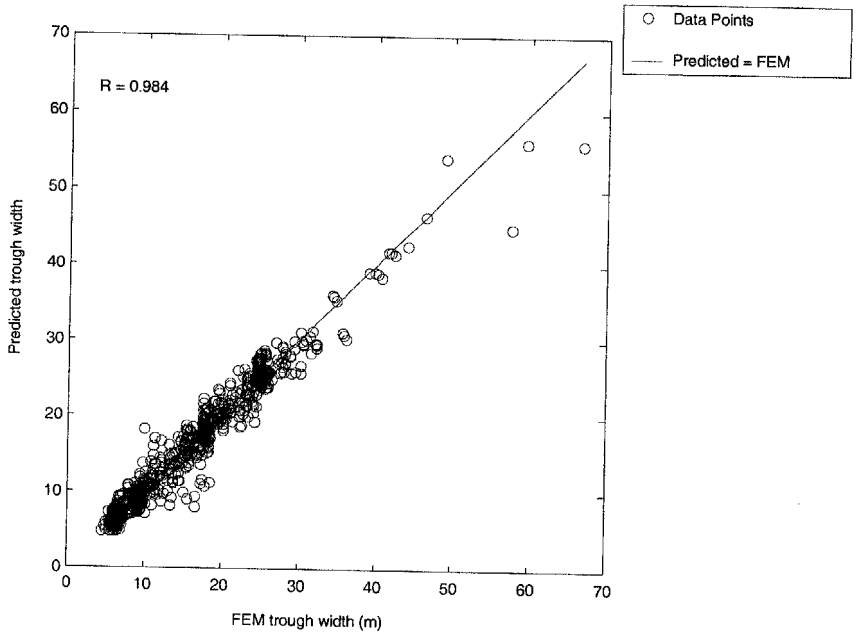


Figure 7.16: Predicted vs. FEM trough widths for training set

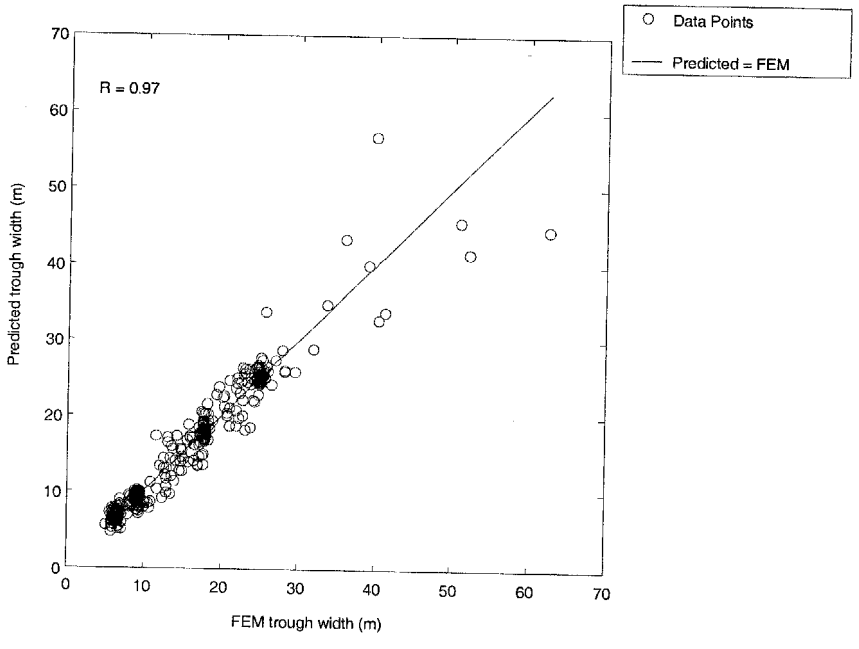


Figure 7.17: Predicted vs. FEM trough widths for testing set

7.2.2. Networks with one output neuron

Two separate networks are developed whereby one network is used to predict S_{max} and the other to predict trough width (i). The results for the two networks tested with different hidden neurons are shown in Table 7.19 and 7.20. Network used for S_{max} prediction is labeled as ANNS followed by the hidden neuron number while that used for (i) prediction is labeled as ANNi. Networks with 12 hidden neurons give optimal predictions of S_{max} as indicated by the maximum correlation coefficients and minimum error rates for both training and testing. Network for (i) prediction is optimum when it used 8 neurons in the hidden layer. From the comparison of results, it is also observed that the two separate networks perform better than the single network used to predict S_{max} and (i) simultaneously, as indicated by higher correlation coefficients and lower error rates when the two separate networks are used.

Table 7.19: Performance of ANN models for S_{max} prediction

Model	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
ANNS	2	0.90	0.93	7.69	6.85
ANNS	4	0.93	0.95	5.65	4.94
ANNS	6	0.95	0.96	4.89	4.77
ANNS	8	0.96	0.97	4.56	4.56
ANNS	10	0.95	0.96	4.60	4.45
ANNS	12	0.96	0.97	4.34	4.16
ANNS	14	0.95	0.96	5.31	5.14

Table 7.20: Performance of ANN models for trough width (i) prediction

Model	No. of Hidden Neurodes	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
ANNi	2	0.95	0.90	1.61	1.87
ANNi	4	0.96	0.93	1.52	1.73
ANNi	6	0.97	0.95	1.41	1.55
ANNi	8	0.98	0.96	1.25	1.40
ANNi	10	0.97	0.94	1.31	1.59
ANNi	12	0.98	0.96	1.19	1.41

- **Comparison of training algorithms**

The two optimum networks for the predictions of S_{max} and (i) are improved further using the faster training algorithms. The results of all the algorithms training including gradient descent method are tabulated in Table 7.21 and Table 7.22. The network trained using Fletcher-Powell Conjugate Gradient (CGF) is considered to be optimum for the prediction of S_{max} while the network trained using Levenberg-Marquardt (LM) algorithm is optimal for trough width (i) prediction. Thus the two optimum networks are labeled as ANNS12CGF and ANNi8LM. From the comparison of results, it is shown that the networks ANNS12CGF and ANNi8LM performed relatively better than the improved BFG network used for simultaneous prediction of S_{max} and (i). This is indicated by higher coefficients of correlation and lower error rates for the training and testing sets when the two improved networks are used to predict S_{max} and (i) separately. Hence these two networks will be validated in section 7.3 using the field data set. Figures 7.18 and 7.19 present the comparison of measured against predicted maximum settlements of the training and testing set for ANNS12CGF, while the comparison of measured against predicted trough width (i) for ANNi8LM are shown in Figures 7.20 and 7.21 .

Table 7.21: Performance of ANNS12 model using various training algorithms for S_{max} prediction

Algorithm	Model	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
LM	ANNS12	0.98	0.97	3.48	3.98
GDM		0.96	0.97	4.34	4.16
BFG		0.99	0.98	2.42	2.88
RP		0.99	0.98	2.53	2.98
SCG		0.98	0.98	3.11	2.97
CGB		0.99	0.98	2.08	2.61
CGF		0.99	0.98	2.01	2.47
CGP		0.98	0.97	3.41	3.67
OSS		0.99	0.98	2.87	3.43
GDX		0.98	0.98	3.00	3.04

Table 7.22: Performance of ANNi8 model using various training algorithms for (i) prediction

Algorithm	Model	Correlation Coefficient		Error Rate (unscaled)	
		Training	Testing	Training	Testing
LM	ANNi8	0.99	0.98	0.76	0.92
GDM		0.98	0.96	1.25	1.40
BFG		0.99	0.98	0.69	0.94
RP		0.99	0.97	0.86	1.05
SCG		0.98	0.97	0.95	1.08
CGB		0.99	0.97	0.77	1.03
CGF		0.99	0.97	0.92	1.06
CGP		0.99	0.97	0.80	0.97
OSS		0.99	0.97	0.87	1.12
GDX		0.98	0.96	1.13	1.30

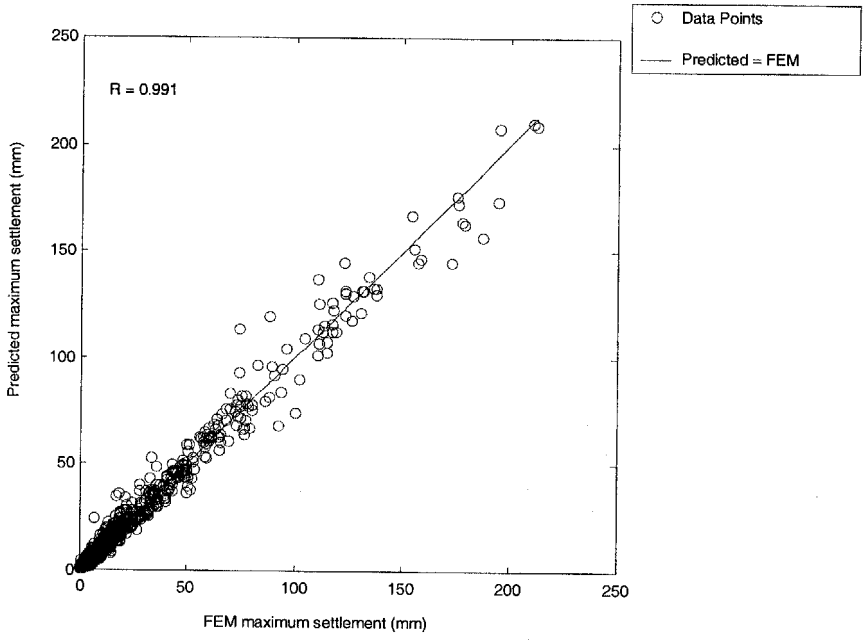


Figure 7.18: Predicted vs. FEM maximum settlements for training set (ANNS12CGF)

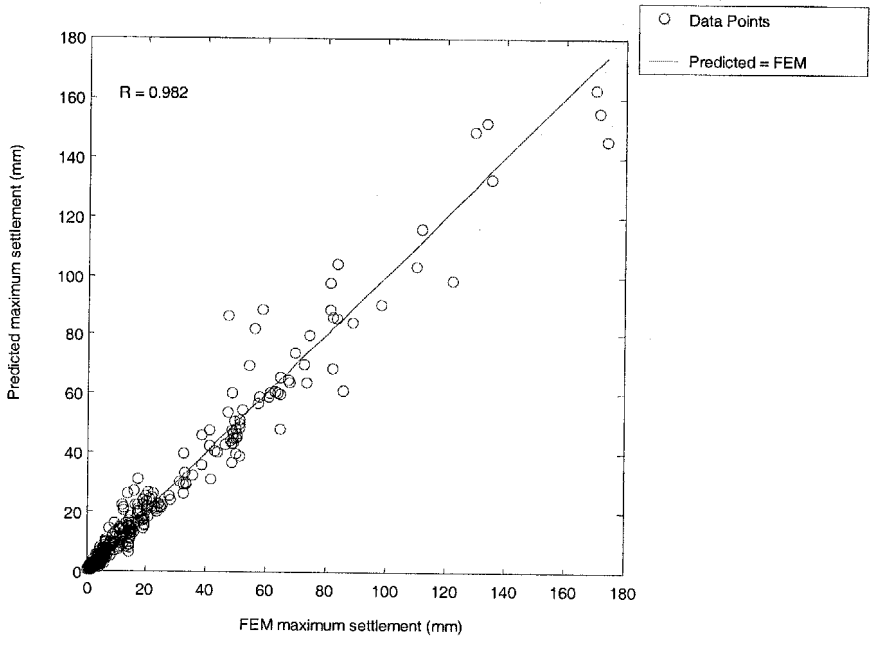


Figure 7.19: Predicted vs. FEM maximum settlements for testing set (ANNS12CGF)

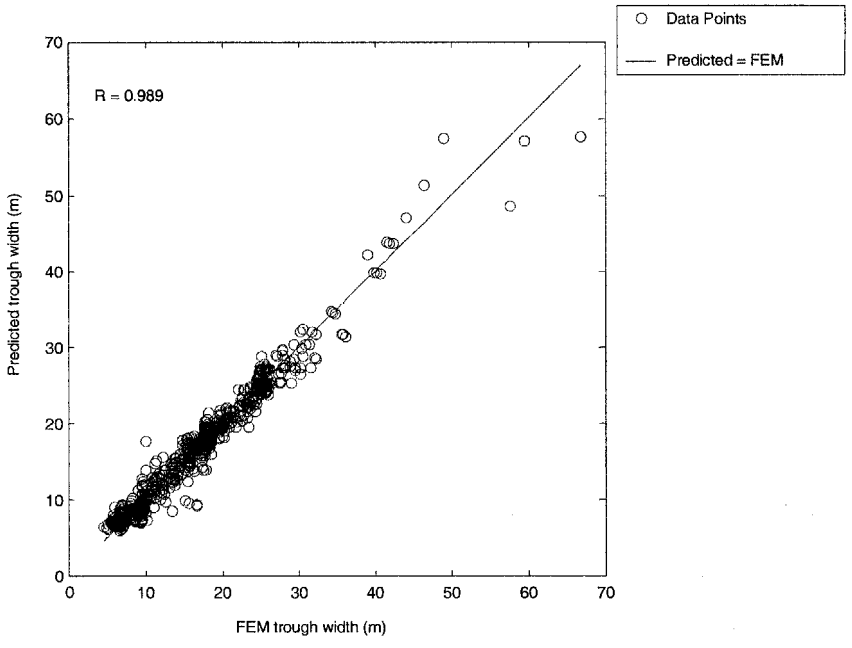


Figure 7.20: Predicted vs. FEM trough widths for training set (ANNi8LM)

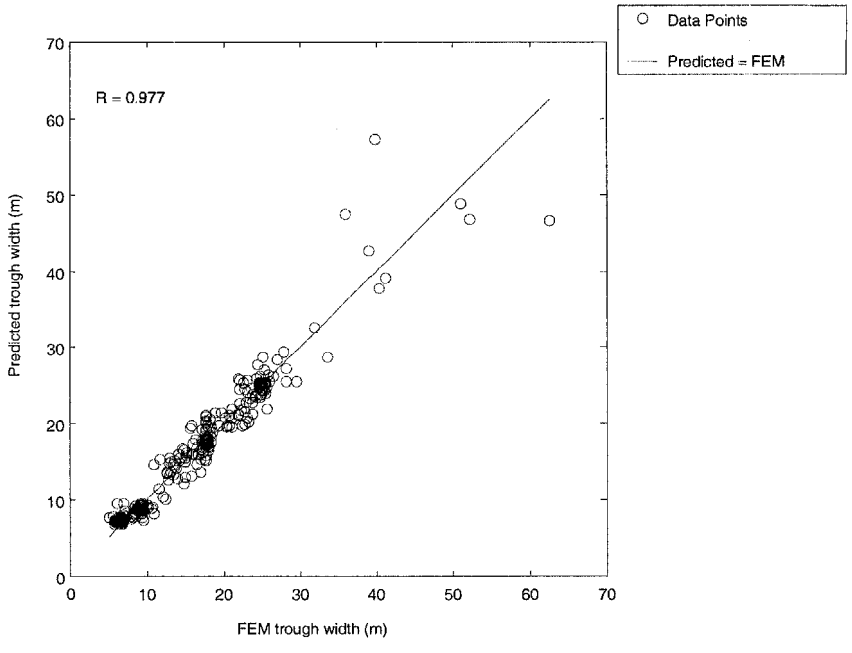


Figure 7.21: Predicted vs. FEM trough widths for testing set (ANNi8LM)

7.3 Validation of the optimum networks using the field data

The predictive ability of optimum neural models, obtained from analysis in section 7.1.2 and 7.2.2, is validated using a set of field data. The optimum networks are ANNS12CGP, ANNi6CGF, ANNS12CGF, and ANNi8LM. The optimum networks, trained using the gradient-descent method, in section 7.2.2 (ANNS12GDM and ANNi8GDM) are tested as well in this analysis for the purpose of comparison with the above networks. The field set comprises 15 patterns of input variables and the corresponding two outputs compiled from contracts C823 and C825 Circle Line project (Table 7.23). The limited number of patterns in the field set is due to insufficient good settlement troughs profile from the field data. The values of input variables are obtained from properties of the soil layer at the tunnel level. The field input and output variables lie within the range specified by the maximum and minimum values for the respective input and output given in Table 7.3. This ensures that the model performs correctly as neural network is designed specifically for interpolation purpose. The field input data is fed to ANNS12CGP, ANNS12CGF and ANNS12GDM to obtain predicted maximum settlements. Whereas the predicted trough widths (i) are obtained using networks ANNi6CGF, ANNi8LM, and ANNi8GDM. Table 7.24 and Table 7.25 show the correlation coefficients and error rates of the validation set for the prediction of S_{max} and (i) with respect to each network. The comparison of predicted settlements and measured settlements for ANNS12GDM, ANNS12CGP, and ANNS12CGF are shown in Figures 7.22 to 7.24. The comparison of predicted trough widths and measured trough widths for ANNi12GDM, ANNi6CGF, and ANNi8LM are shown in Figures 7.25 to 7.27.

Table 7.23: Field data set for validation of optimum networks

No	Instruments	K_0	g	c	E/c	H/D	VL	Smax (mm)	i (m)
1	L1094/01-05	0.69	18.22	89.27	144.84	2.13	9.71	137.84	7.90
2	L1095/01-13	0.89	18.54	134.62	103.55	2.13	6.19	53.36	13.02
3	L1101/01-13	0.63	15.54	32.25	191.01	2.67	2.73	26.67	11.49
4	L1107/01-05	0.63	17.82	35.00	147.71	2.53	2.11	20.78	11.41
5	L1109/01-12	0.70	16.86	26.20	201.53	2.40	2.74	23.91	12.87
6	L1111/01-05	0.63	15.73	18.53	300.05	2.20	1.78	21.53	9.28
7	L1103/02-82	0.63	16.10	43.20	187.27	2.73	2.30	21.87	10.56
8	L1103/07-87	0.63	17.10	44.73	196.51	2.87	1.98	23.24	9.57
9	L1097/01-13	0.75	18.42	61.58	353.69	2.27	3.73	39.71	10.52
10	L1113/01-05	0.63	15.80	19.00	318.95	2.00	0.78	7.41	11.35
11	L1100/01-05	0.80	17.38	37.00	175.14	2.60	3.02	18.21	18.62
12	GSM/01-30	0.63	16.94	39.60	135.35	2.73	2.70	20.06	11.26
13	L8037-L8048	0.90	20.77	620.00	193.55	3.83	0.98	5.14	21.17
14	L9817-L9823	0.80	19.27	45.00	453.33	4.33	2.02	16.59	24.16
15	L8114-L8121	0.75	21.38	454.23	194.06	3.33	0.14	1.60	17.50

Table 7.24: Performance of ANNS12GDM, ANNS12CGP and ANNS12CGF

Model	No. of Hidden Neurodes	Validation Set	
		CoC	Error Rate
ANNS12GDM	12	0.98	6.82
ANNS12CGP	12	0.96	7.14
ANNS12CGF	12	0.95	8.11

Table 7.25: Performance of ANNi8GDM, ANNi6CGF and ANNi8LM

Model	No. of Hidden Neurodes	Validation Set	
		CoC	Error Rate
ANNi8GDM	8	0.82	3.39
ANNi6CGF	6	0.87	2.94
ANNi8LM	8	0.82	2.11

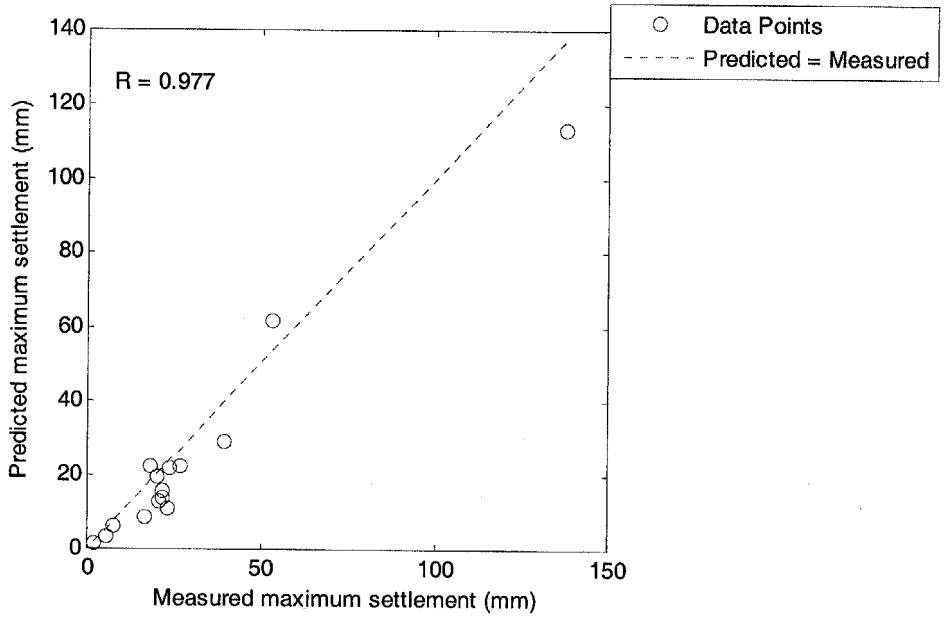


Figure 7.22: Measured vs. predicted maximum settlements for field data set (ANNS12GDM)

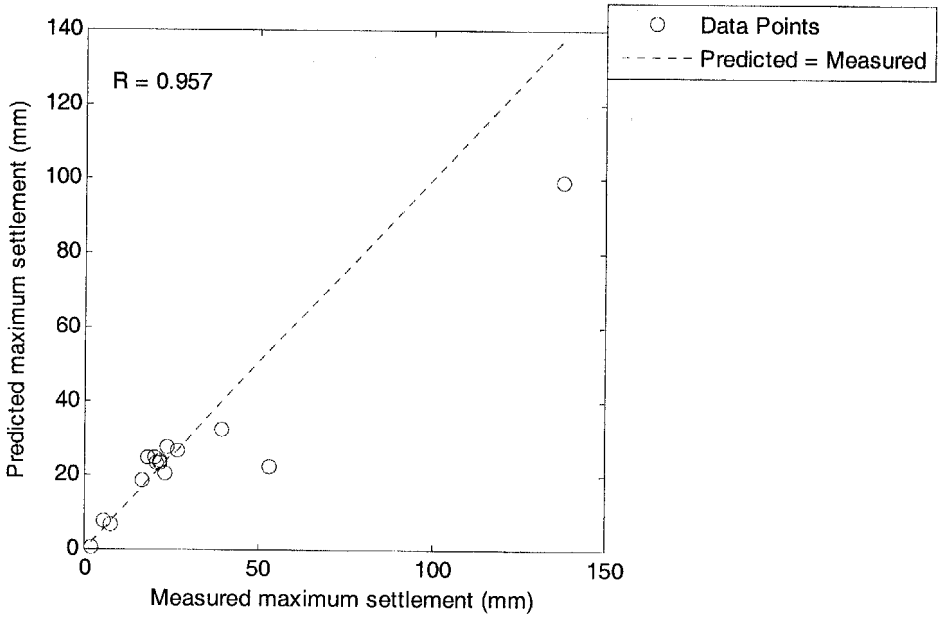


Figure 7.23: Measured vs. predicted maximum settlements for field data set (ANNS12CGP)

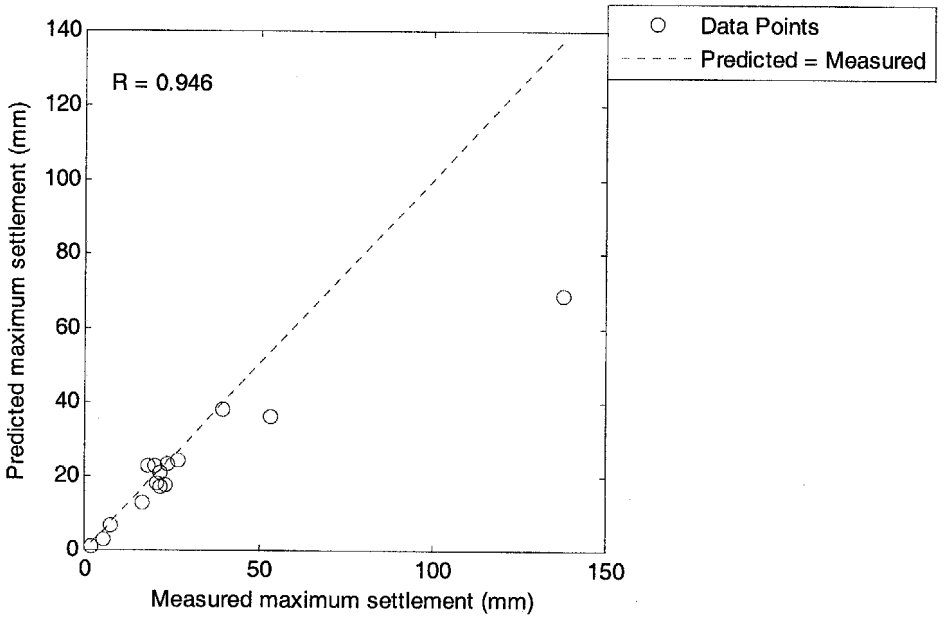


Figure 7.24: Measured vs. predicted maximum settlements for field data set (ANNS12CGF)

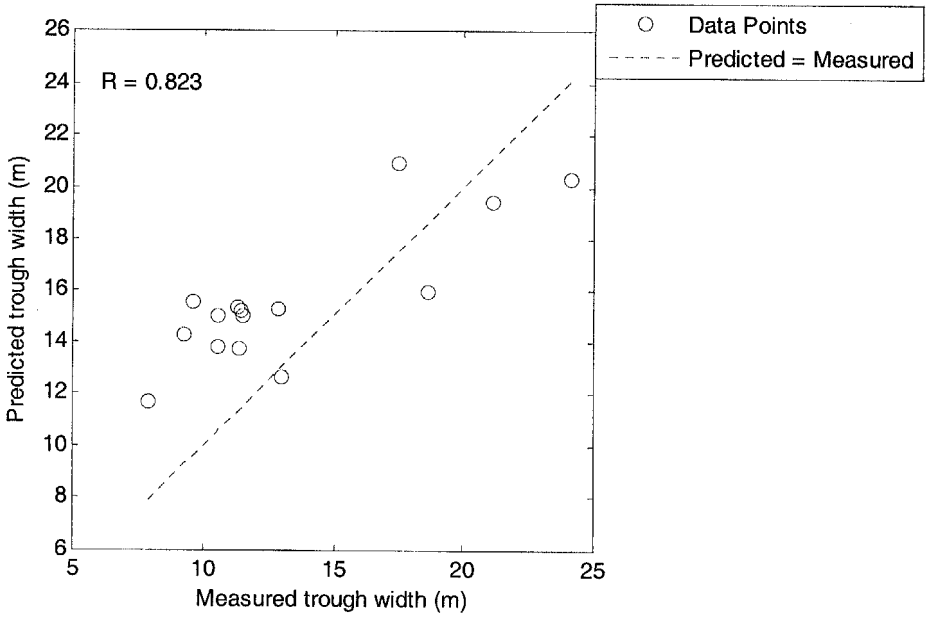


Figure 7.25: Measured vs. predicted trough widths for field data set (ANNi12GDM)

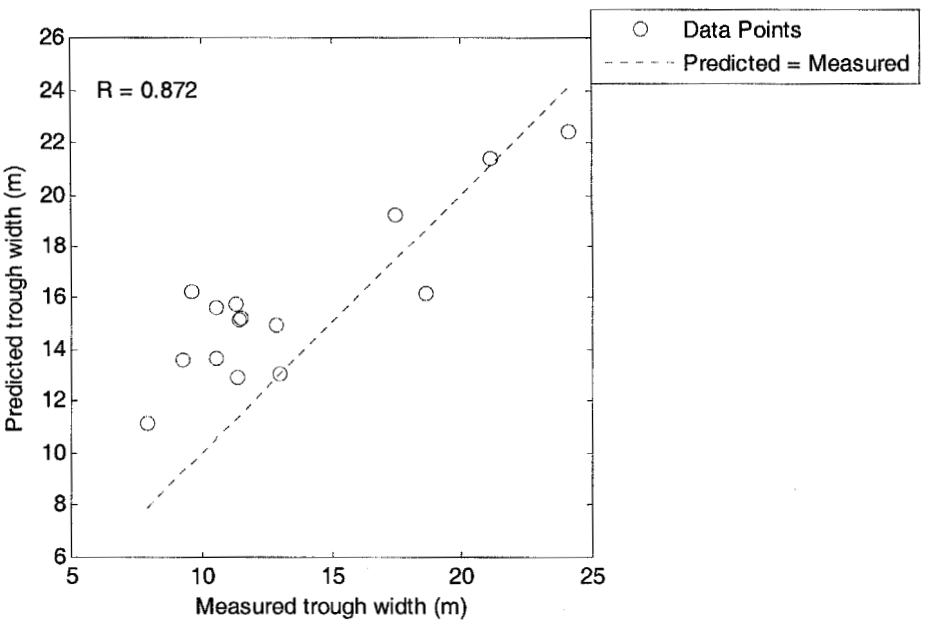


Figure 7.26: Measured vs. predicted trough widths for field data set (ANNi6CGF)

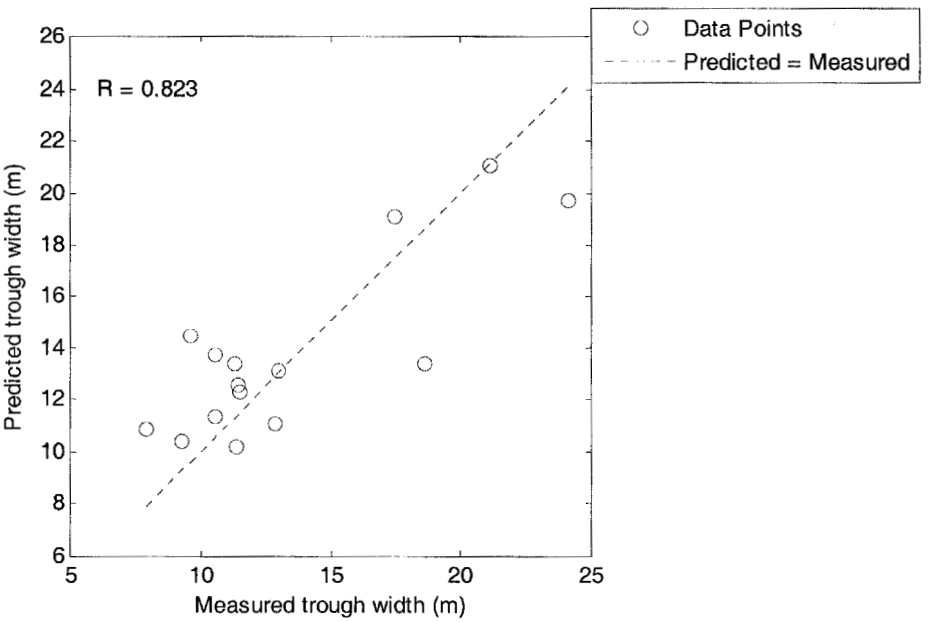


Figure 7.27: Measured vs. predicted trough widths for field data set (ANNi8LM)

The following conclusions can be drawn from the comparison of the networks' results. For the prediction of field maximum settlements, network trained using the gradient-descent algorithm ANNS12GDM perform comparatively better than the other two optimum networks ANNS12CGP and ANNS12CGF. This is indicated in Table 7.24 whereby the results of GDM network produced the highest correlation coefficient and lowest error rate with respect to field data set. For the prediction of field trough width, the results of network ANNi6CGF display the highest correlation coefficient, while the lowest error rate was obtained from the predictions of network ANNi8LM. Overall, network ANNi6CGF is considered optimum for the prediction of trough width. The final weights and bias terms of optimum models ANNS12GDM and ANNi6CGF are listed in Appendix E.

7.4 Sensitivity Analysis of the ANN Model Inputs

The neural models considered in this analysis are ANNS12CGP, ANNi6CGF, ANNS12CGF, and ANNi8LM. These are the optimum networks obtained previously in section 7.1.2 and 7.2.2 for the separate predictions of maximum settlement (S_{max}) and trough width (i). The relative importance of the six input variables is obtained from the connection weights of the network. By using Garson (1991) method, the relative importances of input parameters with respect to each optimum network are given in Table 7.26.

Table 7.26: Relative importance (%) of inputs for ANNS12CGP, ANNi6CGF, ANNS12CGF and ANNi8LM

Network	K_0	γ	c	E/c	H/D	VL
ANNS12CGP	9.51	3.30	36.51	11.74	14.11	24.83
ANNS12CGF	3.42	2.87	49.10	11.11	13.00	20.49
ANNi6CGF	21.97	2.15	40.53	1.25	28.24	5.86
ANNi8LM	27.82	0.77	36.98	2.83	16.75	14.85

Table 7.26 shows that cohesion (c) is the most important input variable for all the four networks. Its relative importance is significantly higher than those of other five inputs. Overall, volume loss (VL) is considered the second important output followed by the depth over diameter ratio (H/D), coefficient of earth pressure (K_0), ratio of stiffness over cohesion (E/*c*), and bulk density (γ).

It is observed that the comparison graphs for S_{max} and (i) in section 7.1 and 7.2 display similar trend, whereby the data points concentrated more on the lower region of settlement values (approximately 50 mm and below). This is perhaps contrary to the common expectation that small maximum settlement will correspond to large trough width and vice versa. This occurrence is related to the result of the sensitivity analysis above and is explained below.

Using the total 1836 patterns as the data, graphs were plotted with a particular input variable on the x-axis and the resulting maximum settlement and trough width on the y-axis. For each input variable, the graph with respect to maximum settlements is plotted separately from the graph corresponding to trough widths. These graphs are shown in Figures 7.28 to 7.33 for all the six input variables.

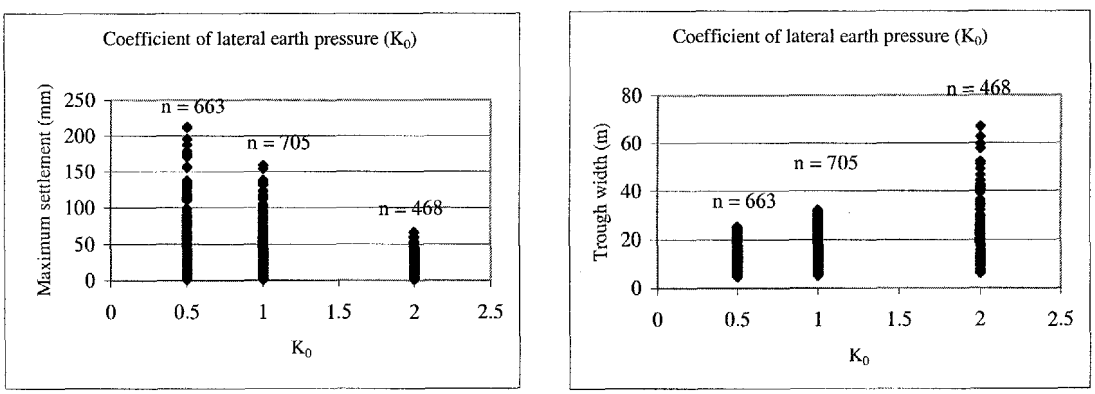


Figure 7.28: Coefficient of lateral earth pressure (K_0) vs. S_{max} and (i)

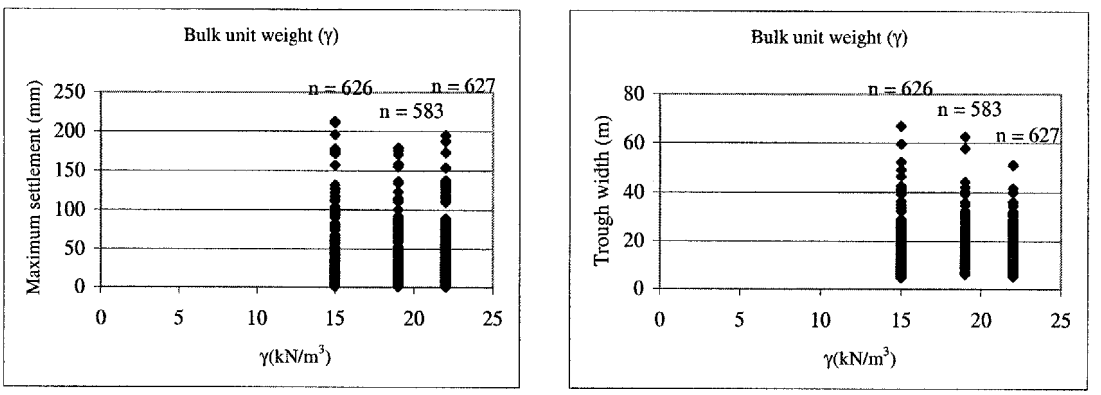


Figure 7.29: Bulk unit weight (γ) vs. S_{max} and (i)

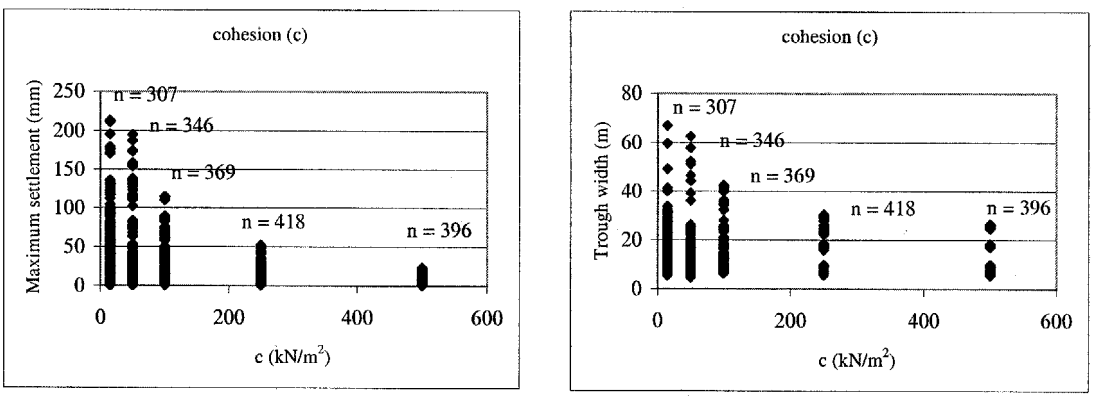


Figure 7.30: Cohesion (c) vs. S_{max} and (i)

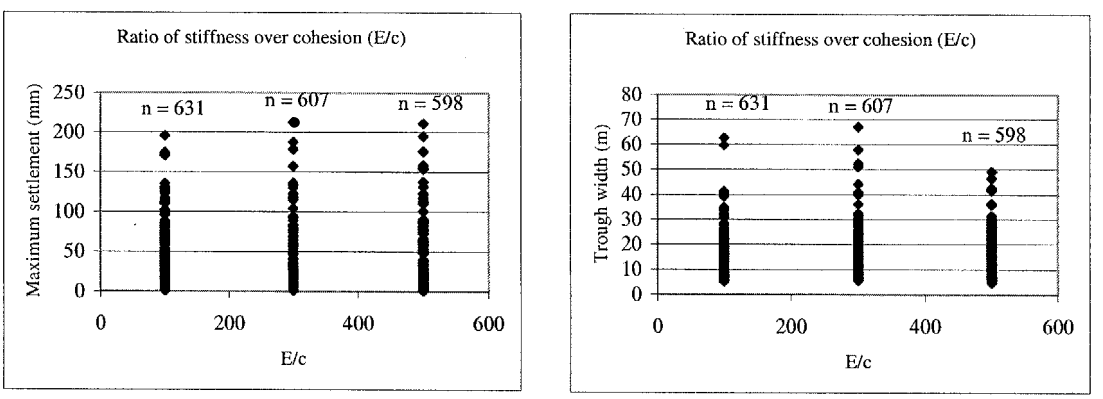


Figure 7.31: Ratio of stiffness over cohesion (E/c) vs. S_{max} and (i)

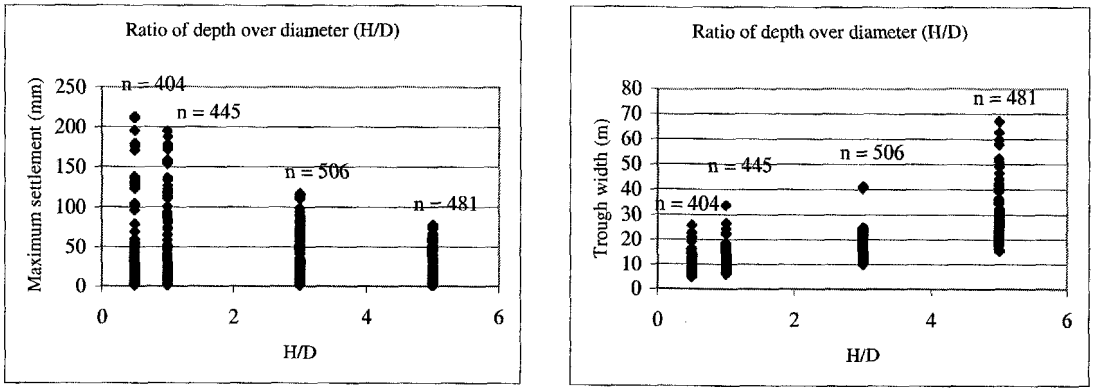


Figure 7.32: Ratio of depth over diameter (H/D) vs. S_{max} and (i)

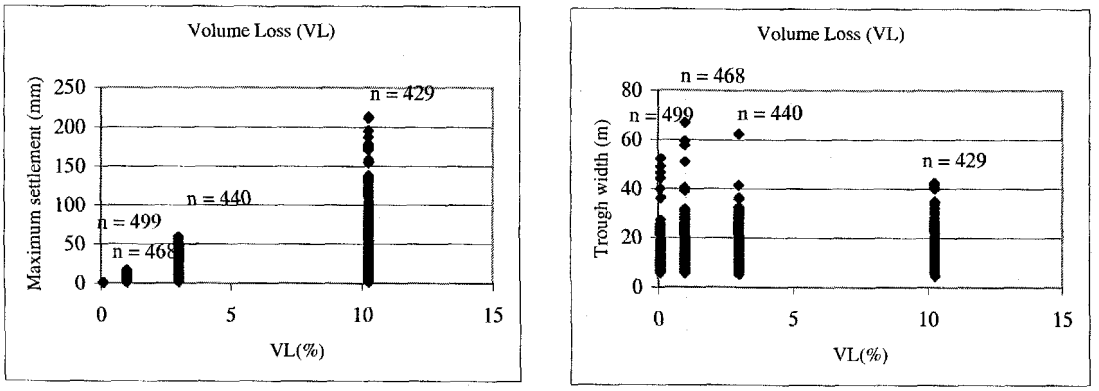


Figure 7.33: Volume loss (VL) vs. S_{max} and (i)

It is observed from the graphs above that, for some input variables, a trend exists which describe the relationship between maximum settlement (S_{max}) and the corresponding trough width (i). For coefficient of lateral earth pressure (K_0), the maximum settlement decreases and the corresponding trough width increases as K_0 value increases. The same trend is observed as well in the graphs of input variable ratio of depth over diameter (H/D). However the graphs of input variable volume loss display the opposite trend whereby the maximum settlement increase and the corresponding trough width decrease as volume loss increases. For input variables bulk unit weight and ratio of stiffness over cohesion (E/c), such distinct pattern is not observed in the graphs. In the

:ase of input variable cohesion, both the maximum settlement and the corresponding rough width decrease as cohesion values increases. Based on this pattern, it can be :tated as well that the trough with large maximum settlement has large trough width ınd vice versa.

Γhe optimum networks consider cohesion (c) as the most important input, and this will nfluence the process of predictions and the results. Consequently it is observed that the :omparison graphs of both S_{max} and trough width (i) display similar pattern following he trend for input variable cohesion explained above. The settlement data points in the :omparison graphs of both S_{max} and trough width (i) concentrate on the lower region of :ettlement values (approximately 50 mm and below).

hapter 8

onclusions and Recommendations

1 Conclusions

this research, the feasibility of using artificial neural networks (ANN) as an alternative ethod to predict the maximum ground surface settlement and settlement trough width ıe to tunneling was investigated. MatLab software was utilized for the network ıalyses.

wo main analyses were carried out:

Prediction of maximum surface settlement by ANN using the field data.

Prediction of maximum surface settlement and settlement trough width by ANN using e finite element data.

the first main analysis, a total of 158 case records of actual field measurements ɔtained from contracts NEL C705, Circle Line C823 and Marina Line C825 were used develop and verify the ANN models. Eleven input units considered to be influential on ttlement value were investigated. From the combinations of these input units, 32 ɛtwork models were developed and tested to obtain the optimal combination of input ırameters. The value of several network parameters including learning rates, momentum rms, transfer functions and initial weights were varied in order to study their effect on e network performance and obtain the optimum architecture of the network model. The fect of using other training algorithms on the performance of ANN model was also ɛamined. Four data division methods namely random, statistical consistent, Self-rganizing Map and fuzzy clustering were used to obtain optimal data sets for the aining, testing and validation. A sensitivity analysis was carried out on the ANN model study the relative importance of the factors that affect settlement. Finally, the ttlement predictions obtained using ANN model were compared with those obtained ɜing empirical method.

rom the analysis performed, the following conclusions can be made.

1. From 32 network models, ANN model with eight input parameters namely soil cover, advance rate, earth pressure, SPT1, SPT2, moisture content, stiffness and grout pressure is considered optimal.
2. It appears that the number of hidden nodes does not affect significantly the performance of ANN models. This can be attributed to the early-stopping method which was used as the stopping criterion. The initial set of weights used to train the network is important as well since the favorable initial weights will produce optimum final weights and vice versa. The ANN model with eight hidden neurons yielded optimum results for the prediction of ground surface settlements, and thus it was used in the network parameter analysis.
3. The study of various momentum terms and learning rates showed that the minimum error rate for testing set was achieved when the momentum and learning rate were 0.9 and 0.2. Consequently, the remaining analysis was carried out using 0.9 and 0.2 as the momentum and learning rate. These results also support the argument that in the initial phase of network analysis, it is recommended to set momentum and learning rate to these two values.
4. The results obtained using Backpropagation with gradient descent method was slightly better than those obtained using other algorithms. The major shortcoming of gradient descent method is its convergence which is much slower than the other algorithms. However, when early stopping is used as the stopping criterion, it is better to use algorithm that converges more slowly. If the algorithm which converges too quickly is used, there is a possibility that we may overshoot the point at which the error on the validation set is minimized.
5. It is essential to maintain the consistency of the statistics between the training, testing and validation sets in order to produce results which are representative of the available data set. This can be achieved by using one of the three approaches namely statistical consistent method, Self-Organizing Map and fuzzy clustering. In this project, it appears that the ANN model performed relatively better when it used data subsets obtained using statistical consistent method.

6. The sensitivity analysis indicated that every network assigns relatively different relative importance with respect to an input parameter. The most important input for a particular network might be the least important for other network and vice versa. Overall, moisture content is considered an essential input parameter as it displays relatively high importance whenever it is used as input of the network.
7. Optimum ANN model NN6 performed significantly better than the empirical method considered for the same validation set. The predictions obtained by ANN model NN6, were quite close to the measured maximum settlements; whilst the empirical method underpredicted the measured settlements. Hence, it is evident that ANN model provide more accurate settlement predictions than the empirical method. This confirms the feasibility of using artificial neural networks as an alternative method to predict the maximum settlements due to tunneling.
8. ANN model NN6 might not be effective for quick initial prediction of maximum surface settlements. This is because it requires complete data of the eight input parameters from the project, whereas some input such as earth pressure, advance rate, and grout pressure may not be available at the beginning of construction period. Hence it is proposed to use another model NN29 which requires only four inputs readily available from the site, namely cover (H), SPT1, SPT2, and moisture content (MC). The reliability of this alternative network is substantiated by high correlation coefficients for training, testing, and validation sets and relatively low error rates for testing and validation sets.

the second main analysis, neural network models were developed for the prediction of aximum settlements and trough width using the data generated from the finite element ftware PLAXIS. PLAXIS is commonly used for geotechnical applications in which soil odels are used to simulate the soil behavior. A total of 2161 patterns were generated om the combinations of six input parameters, namely coefficient of earth pressure (K_0), ılk density (γ), cohesion (c), ration of stiffness over cohesion (E/c), ratio of depth over ameter (H/D), and volume loss (VL). The main output of the program is the settlement ırve from which the settlement of the points along the soil surface can be obtained. The aussian distribution was fitted to the settlement points and the standard deviation of the

quation was used as the corresponding trough width (i). Patterns which produce failed esults or bad settlement graphs were discarded; hence leaving 1836 patterns for the nalysis. Two cases of network training are considered in the analysis. In the first case, etwork models are developed using three statistically consistent data sets i.e. training, esting and validation. In the second case, the input data is divided into training and esting sets only. For each case, two types of network models were developed. The first iodel has two output neurons where the predictions of S_{max} and trough width (i) can be arried out simultaneously. The second model is the networks with one output neuron /hich predict S_{max} and (i) separately. The use of faster training algorithms to improve the ccuracy of the network is investigated and the results are compared with the result of tandard gradient descent method. The optimum network models from the two cases of 'aining and the optimum network trained using gradient descent method are retained and alidated using a set of field data to examine the generalization ability. The field data /ere collected from contracts C823 and C825 Circle Line project. From the network nalysis, the following conclusions can be drawn:

1. Two separate networks performed better than the single network used to predict S_{max} and (i) simultaneously as shown by higher correlation coefficients and lower error rates when the two networks are used. As discussed earlier, the main reason for this is that each of the two independent networks has more freedom to improve its degree of accuracy without having to take into account the conditions of another output.
2. For the case of training with three data sets, the network of 12 hidden neuron trained using the Polak-Ribiére Conjugate Gradient (CGP) is considered to be optimum for the prediction of S_{max}. For trough width (i) prediction, network of six hidden neurons trained using Fletcher-Powell Conjugate Gradient (CGF) is the best. The two optimum networks are labeled as ANNS12CGP and ANNi6CGF. In case of training with two data sets, the network of 12 hidden neuron trained using Fletcher-Powell Conjugate Gradient (CGF) is considered to be optimum for the prediction of S_{max} while the network of 8 hidden neuron trained using Levenberg-

Marquardt (LM) algorithm is optimal for trough width (i) prediction. The two optimum networks are labeled as ANNS12CGF and ANNi8LM.

3. For the prediction of field maximum settlements, network trained using the Gradient-Descent Method (GDM) ANNS12GDM perform relatively better than the other two optimum networks ANNS12CGP and ANNS12CGF. The prediction results of GDM network showed the highest correlation coefficient and lowest error rate with respect to field data set. For the prediction of field trough width, the results of network ANNi6CGF display highest correlation coefficient, while the lowest error rate was obtained from the predictions of network ANNi8LM. Overall, the performance of network ANNi6CGF is considered better than the other two optimum networks ANNi12GDM and ANNi8LM. Hence, it is concluded that ANN models ANNS12GDM and ANNi6CGF are optimal and they can be used for the prediction of maximum settlements and trough widths respectively.
4. The sensitivity analysis indicated that the four optimum networks ANNS12CGP, ANNi6CGF, ANNS12CGF, and ANNi8LM consider cohesion (c) as the most important input. Its relative importance is significantly higher than those of remaining five inputs. In general, volume loss (VL) is the second important output followed by the depth over diameter ratio (H/D), coefficient of earth pressure (K_0), ratio of stiffness over cohesion (E/c), and bulk density (γ).

.2 Recommendations for Future Works

or the first main analysis, more case records are required in the network analysis to nprove the reliability of network model. Thus, it is recommended to collect more data in ıe future from various tunneling projects not only in Singapore but also in other parts of ıe world. Networks with different input parameters which showed satisfactory erformance in the testing of neural network models can be further investigated in order ɔ assess their actual predictive capability. Other factors related to Shield operation such

s pitching angle, thrust force, cutter torque, and shield position can be added in the list f input parameters and their effect on the network performance should be examined.

or the second main analysis, more field data are required to verify the predictive ability f the optimum networks. In this project, the limited amount of field data for validation et is due to shortage of good settlement trough curves from the site offices. Hence, good ettlement curves should be collected as many as possible from other tunneling projects urrently going on in Singapore. The good settlement troughs from other projects rovided in the literature can added to the field data set as well. The validation of NNS12GDM network performance using the field data indicated underestimation for irge maximum settlements around 120mm and above. The large maximum settlements re by and large associated with high volume loss. Among the volume losses used in LAXIS to generate the input data, only 10% volume loss is considered high. The emaining volume losses are below 5%. This means lack of input data for the large iaximum settlements. Hence, other volume loss from 5% to 10% should be included in ie input data generation as well in order to get better estimation of large maximum ettlements. In this project, the field measurements indicated small volume losses (around .1% to 3%) for most of the field data. Therefore, the values of volume losses selected or input data generation are within this range.

References

Anderson, D., and McNeil, G. (1992). "Artificial Neural Networks Technology." *http://www.dacs.dtic.mil/techs/neural/neural_ToC.html*. Data and Analysis Center for Software, Rome Laboratory, Rome, NY 13441.

Atkinson, J. H., and Potts, D. M. (1976). "Subsidence above shallow circular tunnels in soft ground". *Dept. of Engineering, University of Cambridge Report CUED/C-Soils/T.R.27.*

Attewell, P. B. (1977). "Ground movements caused by tunneling in soil". *Proc. Conf. on large Ground Movements and Structures*, Pentech Press, London, 812-948.

Attewell, P. B., and Farmer , I. W. (1974b). "Ground deformations resulting from shield tunneling in London clay". Canadian Geotech. J. 11, 380-395.

Attewell, P. B., and Woodman, J. P. (1982). "Predicting the dynamics of ground settlement and its derivatives caused by tunneling in soil". *Ground Engineering*, 15, (8), 13-22, 36.

Attewell, P. B., Yeates, J., and Selby, A.R. (1986). *Soil Movements Induced by Tunnelling and their Effects on Pipelines and Structures*, Chapman and Hall, New York.

Bowden, G. J., Maier, H. R., and Dandy, G. C. (2002). "Optimal division of data for neural network models in water resources applications." *Water Resources Research*, 38(2), 2.1-2.11.

Brown, M., An, P. C., Harris, C. J., and Wang, H. (1993). "How Biased is your Multi-Layer Perceptron?" World Congress on Neural Networks, 3, pp507-511.

Burd, H. J., Houlsby, G. T., Augarde, C. E., Liu, G. (2000). "Modelling tunneling-induced settlement of masonry buildings." *Proc. Institution of Civil Engineers, Geotechnical Engineering*, 143(1), pp 17-29.

Chern, S. G., Hu, R. F., Chang, Y. J., Tsai, I. F. (2001). "Fuzz-ART neural networks for predicting Chi-Chi Earthquake induced liquefaction in Yuan-Lin area." *J. Marine Science & Technology*, 10(1), pp 21-32.

Clough, G. W., and Schmidt, B. (1981). "Design and performance of excavations and tunnels in soft clay". *Soft Clay Engineering*, Elsevier Scientific Publishing Company, Amsterdam, 569-636.

Cording, E. J., and Hansmire, W. H. (1975)." Displacements around soft ground tunnels". *5th Pan American Conference on Soil Mechanics and Foundation Engineering*, Buenos Aires, 571–632.

Cybenko, G. (1989). "Approximation by Superposition of a Sigmoidal Function". *Mathematics of Control, Signals and Systems*, Vol 2, pp313-314.

Deere D. U., Peck, R. B., Monsees, J. E., and Schmidt, B. (1969). "Design of tunnel liners and support systems". *Final report by the Dept. of Civil Engineering, University of Illinois for the Office of High Speed Ground Transportation*, US Dept. of Transportation, Washington, Contract No. 3-0152.

Davalo, E., and Naim, P. (1991).*Neural Networks*, Macmillan Education, London.

Duan, Z. (2001). "Ground Movement Associated with Microtunneling." *Ph.D. Thesis*, Jackson State University , p195.

Eberhart, R. C., and Dobbins, R. W. (1990). *Neural Network PC Tools: A Practical Guide,* Academic Press, San Diego.

Fahlman, S. E. (1989). "Faster-learning variations on back-propagation: An empirical study." *Proc. the 1988 Connectionist Models Summer School*, San Mateo, CA, pp 38-51

Freeman, J. A., and Skapura, D. M. (1991). *Neural networks : algorithms, applications, and programming techniques* , Addison-Wesley Pub. Co., Reading, Massachusetts.

Fang, Y. S., Lin, J. S., and Su, C. S.(1994). "Estimation of ground settlement due to shield tunneling by the Peck-Fujita method." *Can. Geotech. J.*, 31(3), pp 431-443.

Garson, G. D. (1991). "Interpreting neural-network connection weights." *AI Expert*, 6(7), 47-51.

Glossop, N. H. (1978). "Soil deformations caused by soft-ground tunneling". *Ph.D. Thesis*, University of Durham, England.

Goh, A.T.C.(1994). "Nonlinear modeling in geotechnical engineering using neural networks." *Australian Civil Engineering Transactions*, CE36(4), pp 293-297.

Goh, A.T.C.(1995). "Seismic Liquefaction Potential Assessed by Neural Networks." *J. Geotech. Eng.*, 120(9), pp 1467-1480.

Goh, A.T.C. (1996). "Pile driving records reanalyzed using neural networks". *J. of Geotech. Engrg*, Vol 122, No. 6, pp 492-495.

Gunn, M. J. (1992). "Prediction of surface settlement profiles due to tunneling." *Proc. Wroth Memorial Symposium, Oxford, UK.* (Predictive Soil Mechanics p304).

Haack, A. (2000). "Political and social aspects of present and future tunnelling." Proc. Int. Conf. on Tunnels and Underground Structures, Singapore.

Hagan, M. T., Demuth, H., and Beale, M. (1996). *Neural Network Design,* PWS Publishing, Boston, MA.

Hansmire, W. H. (1975). "Field measurements of ground displacements about a tunnel in soil". *Ph.D. thesis*, University of Illinois at Urbana-Champaign.

Hebb, D.O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. John Wiley & Sons, New York.

Hefny, A. M. (2000). *NNGeo*. Artificial Neural Network Software, Internal Communication, Nanyang Technological University.

Henseler, J. (1995). "Back Propagation" *Artificial Neural Networks: An Introduction to ANN Theory and Practice*, Lecture Notes in Computer Science, Springer, 931, pp 11-36

Hoi, L. M. (1999). "Pile Driving Analysis and Axial Capacity Prediction using Artificial Neural Network."*Research Projects in Civil Engineering*, Faculty of Science and Technology University of Macau.

Hornik, K. (1993). "Some new results on neural network approximation." *Neural Networks*, 6, pp 1069-1072.

Hornik, K., Stinchcombe, M., and White, H. (1989). "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2, pp 359-366

Hoyaux, B., and Ladanyi, B.(1970). "Gravitational stress field around a tunnel in soft ground." *Can. Geotech. J.*, 7, pp 54-61.

In-Mo, L., and Jeong-Hark, L. (1996). "Prediction of Pile Bearing Capacity Using Artificial Neural Networks." *Computers and Geotechnics*, 18 (3), pp. 189-200.

Izumi, C., Khatri, N. N., Norrish, A., Davies, R. (2000)." Stability and settlement due to bored tunneling for LTA, NEL". *Proc. Int. Conf. on Tunnels and Underground Structures*, Singapore, 555-560.

Javadi, A.A., Brown, M.J., and Hyde, A.F.L. (2001). "Prediction of load bearing capacity of piles from Statnamic pile testing data using neural networks." *Proc. 5th Int. Conf. On Deep Foundation Practice*, Singapore, pp. 235-242.

Jordan, M. (1995). "Why the logistic function? A tutorial discussion on probabilities and neural networks". *Technical Report 9503*, Computational Cognitive Science, MIT, Cambridge.

Kaufman, L., and Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*, John Wiley & Sons, New York.

Kim, C.Y., Bae, G.J., Hong, S.W., Park, C.H., Moon, H.K., Shin, H.S. (2001)." Neural network based prediction of ground surface settlements due to tunneling." *Computers and Geotechnics*, 28(6-7), pp 517-47.

Kordos, M., and Duch, W. (2003). "On some factors influencing MLP error surface." *Proc. The Seventh International Conference on Artificial Intelligence and Soft Computing (ICAISC),* Zakopane.

Kohonen, T. (1995). *Self-organizing maps*, Springer-Verlag, Berlin Heildelberg.

Lathauwer, W. D (1992). "Urban co-ordinated underground works: A way to solve many problems". *Proc. Int. Congress Towards New Worlds in Tunnelling*, Acapulco, Vol. 1, pp. 3-10.

le Cun, Y. (1988). "A theoretical framework for Backpropagation". *Proc. Connectionist Models Summer School*, San Mateo (CA), pp. 21-28. 36

Leach, G. (1985)." Pipeline response to tunneling". Unpublished paper presented to the North of England Gas Association, January 1985.

Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S. (1993). "Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function." *Neural Networks,* 6(6), pp 861-867.

Levine, D. M., Berenson, M. L., and Stephan, D.(1999). *Statistics for managers using Microsoft Excel*, Prentice Hall, New Jersey.

Lippmann, R. P. (1992). "An introduction to computing with neural network." *Artificial Neural Networks : Concepts and Theory*, IEEE Computer Society Press Tutorial, Los Alamitos, California, 13-45

Mair, R. J., Taylor, R. N., and Bracegirdle, A. (1993). "Subsurface settlement profiles above tunnels in Clays." *Geotechnique*, 43(2), pp 315-320.

Maier, R. M., and Dandy, G. C. (1998). "The effect of internal parameters and geometry on the performance of back-propagation neural networks: An empirical study." *Environmental Modelling & Software*, 13, pp 193-209.

Masters, T. (1993). *Practical neural network recipes in C++*, Academic Press, San Diego, California.

McCulloch, W. S., and W. H. Pitts (1943). "A logical calculus of ideas immanent in nervous activity." *Bulletin of Mathematical Biophysics*, 5, pp. 115-133.

McKinley, J. D.(1996). "Extracting pattern from scattered data -applicability of artificial neural networks to the interpretation of bearing capacity data".*Technical Reports*, Geotechnical Research Group, Department of Engineering, University of Cambridge, http://wwwciv.eng.cam.ac.uk/geotech_new/publications/TR/TR299.pdf

McLauchlan, P. F. (2002). "Levenberg-Marquardt minimization". http://www.ee.surrey.ac.uk/Research/VSSP/RAVL/share/doc/RAVL/LevenbergMarquardt/

Meyerhof, G. G. (1965). "Shallow foundations." *J. Soil Mech. & Found. Div.*, 91(SM2), 21-31.

Moss, N. A. (2000). "Tunnel linings for Contract C705, Singapore Northeast Line". *Proc. Int. Conf. on Tunnels and Underground Structures*, Singapore, 219-225.

Nawari, N.O., Liang, R., and Nusairat, J.(1999)."Artificial intelligence techniques for the design and analysis of deep foundations." *Electronic J. Geotech. Eng.*, http://www.ejge.com/1999/Ppr9909/Ppr9909.htm.

O'Reilly, M. P., and New, B. M.(1982). "Settlements above tunnels in the United Kingdom – their magnitude and prediction". *Proc. Tunneling '82*, London, 137-181.

Osborne, N., Noren, C., Li, G.J., Chinniah, R., and Jonsson, P. (2004). "Design and Construction of MRT Project Contract 825 of CCL1 in Singapore". *Proc. Underground Singapore 2003 and Workshop*, Singapore, 75-82.

Ou, C. Y., and Cherng, J. C. (1995). "Effect of tail void closure on ground movement during shield tunneling in sandy soil". *Geotechnical Engineering, Southeast Asian Geotechnical Society (SEAGS)*, 26(1), 17-32.

Ou, C. Y., Hwang, R. N., and Lai, W. J. (1998). "Surface settlement during shield tunneling at CH218 in Taipei". *Canadian Geotech. J.* 35, 159-167.

Peck, R. B. (1969). "Deep excavations and tunneling in soft ground". *State of the art report, in 7th Int. Conf. On Soil Mechanics and Foundation Engrg,* Mexico City, 225-290.

Randolph, M. F., and Wroth, C. P. (1978). "Analysis of deformation of vertically loaded piles."*J. Geotech. Eng.*, 104(12), 1465–1488.

Rausche, F., Moses, F., and Goble, G. G.(1972). "Soil resistance prediction from pile dynamics."*J. Soil Mech. and Found. Div.*, ASCE, **98**(9), 917–937.

Riedmiller, M., and Braun, H. (1994). "RPROP -- Description and Implementation Details". *Technical Report*. Universitat Karlsruhe.

Robin, W. L. P. (2001). "Prediction of ground settlement due to Tunnelling using Artificial Neural Networks." *Final Year Report*, Nanyang Technological University.

Rosenblatt, F. (1962). *Principles of Neurodynamics*, Spartan Press, Washington D.C.

Rowe, R. K., and Lee K. M. (1992). "An evaluation of simplified techniques for estimating three dimensional undrained ground movements due to tunneling in soft soils." *Canadian Geotech. J.*, 29, pp 39-52.

Rumelhart, D. E., Durbin, R., Golden, R., and Chauvin,Y.(1995)." Backpropagation: The Basic Theory." *Backpropagation: Theory,Architectures, and Applications*, D. E. Rumelhart and Y. Chauvin, eds.,Lawrence Erlbaum Associates, New Jersey.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning internal representation by error propagation." *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, eds., MIT Press, Cambridge.

Sarle, W. S.(2002). *Old Neural Net FAQ*, ftp://ftp.sas.com/pub/neural/FAQ.html.

Saxena, S. C. (19--). *Tunnel Engineering: (A Text-Books for Civil Engineering Students)*, Dhanpat Rai and Sons, Nai Sarak, Delhi.

Schmertmann, J. H. (1970). "Static cone to compute static settlement over sand."*J. Soil Mech. & Found. Div.*, 96(SM3), 7302–1043.

Schmertmann, J. H., Hartman, J. P., and Brown, P. B.(1978). "Improved strain influence factor diagrams."*J.Geotech. Eng.*, 104(GT8), 1011–1043.

Schmidt, B. (1969)."Settlements and ground movements associated with tunneling in soil". *Ph.D. Thesis*, University of Illinois.

Schultze, E., and Sherif, G. (1973). "Prediction of settlements from evaluated settlement observations for sand". *Proc. 8th Int. Conf. Soil Mechan. & Found. Eng.*, 1, 225-230.

Seed, H.B., Idriss, I.M., and Arango, I.(1983). "Evaluation of Liquefaction Potential Using Field Performance Data." *J. Geotech. Eng.*,109, pp458-482.

Shahin, M. A. (2003). "Use of artificial neural networks for predicting settlement of shallow foundations on cohesionless soils." *Ph.D. Thesis*, University of Adelaide.

Shi, J., Ortigao, J.A.R., Bai, J (1998). "Modular neural networks for predicting settlements during tunneling." *J. Geotech. and Geoenv. Eng.*, 124(5), pp 389-395.

Shirlaw, J. N., Ong, J. C. W., Rosser, H. B., Osborne, N. H., Tan, C. G., and Heslop, P. J. E. (2001). "Immediate settlements due to tunneling for the North East Line." *Proc. Underground Singapore 200*, Singapore, session 3.

Sivakugan, N., Eckersley, J. D., and Li, H. (1998). "Settlement predictions using neural networks." *Australian Civil Engineering Transactions*, CE40, pp49-52.

Smith, A. E. (1997). *Pittnet Neural Network Educational Software*. Department of Industrial Engineering. University of Pittsburgh.

Sun, J. (2000). "Soil disturbance and stability control of shield tunneling in urban environment". *Proc. Int. Conf. on Tunnels and Underground Structures*, Singapore, 561-566.

Sundararajan, N., and Saratchandran, P. (1998). *Parallel architectures for artificial neural networks: paradigms and implementations*. IEEE Computer Society, Los Alamitos, California.

Teh, C.I., Wong, K.S., Goh, A.T.C., Jaritngam, S.(1997)."Prediction of pile capacity using neural networks". *J. Computing in Civil Engineering*, 11(2), pp 129-138.

Terzaghi, K., and Peck, R. (1967). *Soil mechanics in engineering practice*, John Wiley & Sons, New York.

Tokar, S. A., and Johnson, P. A. (1999). "Rainfall-runoff modeling using artificial neural networks." *J. Hydrologic Eng.*, 4(3), pp 232-239.

Toll, D.(1996). "Artificial Intelligence Systems for Geotechnical Engineering with Specific Reference to the Ground Improvement." *Proc. 10th European Young Geotechnical Engineers Conference*, Izmir, Turkey.

Tsoukalas, L. H., and Uhrig, R. E. (1997). *Fuzzy and Neural Approaches in Engineering*, John Wiley & Sons Inc, New York, USA.

Tveter,D.R.(1998).Backpropagator'sReview,http://www.dontveter.com/bpr

Ural, D. N., and Saka, H. (1998). "Liquefaction Assessment by Artificial Neural Networks." *Elect. J.Geotech.Eng.*,http://www.ejge.com/1998/Ppr9803/Ppr9803.htm

Uriel, A. O., and Sagaseta, C. (1989). "Selection of design parameters for underground construction." *Proc. 12th ICSMFE*, Rio De Janeiro, Brazil.

Von Neumann, J.(2004). The Artificial Reputation of John von Neumann's Intelligence, http://www.21stcenturysciencetech.com/articles/von_Neumann.html.

Weijters, A. J. M. M., and Hoppenbrouwers, G. A. J. (1995). " Backpropagation Networks for Grapheme-Phoneme Conversion: a Non-Technical Introduction." *Artificial Neural Networks: An Introduction to ANN Theory and Practice*, Lecture Notes in Computer Science, Springer, 931, pp 11-36

Werbos, P. J. (1947). *The roots of backpropagation : from ordered derivatives to neural networks and political forecasting*, J. Wiley & Sons, New York.

Widrow, B., and Hoff, M. E. (1960). "Adaptive switching circuits."*IRE WESTCON Connection Record*, 4, pp 96-104.

Appendix A.

Pile driving data (Goh, 1995) for testing the accuracy of various softwares

Training Data		
S_t	σ'_v	f_s
26	96	27
15	102	26
23	54	14
23	87	26
17	49	12
13	37	11
15	32	9
10	33	12
12	39	10
15	19	8
19	146	29
57	109	24
19	38	17
36	82	28
22	89	22
45	60	23
30	44	38
31	142	30.7
104	448	109.2
162	718	162
38	162	30
80	354	44
67	273	47.6
170	651	192.1
45	153	29.3
52	148	21.8
45	112	42.3
129.5	51.6	76.7
39	105	39.8
16	33	9.9
30	59	23.4
165	297	80.9
52	91	30.7
61	99	34.2
110	80	53.9
208	105	91.5
144	54	73.4
100	87	55
137	112	64.4
335	115	154.1
120.5	43	84.6
35.4	121.4	30.4
48.8	108	27.1
112.8	158.2	53
24.4	102.8	23.5

Testing Data		
S_t	σ'_v	f_s
9	27	9
29	52	18
29	67	16
27	57	18
15	42	13
31	147	28.8
28.2	149.6	30.5
60	223	31.2
21	44	13.4
52	91	27
21	27	13
185	244	88.8
53	66	27.6
33	51	32
64	152	37.8
115	141	59.8
22	81	20
19	147	30
72	80	35
96	110	54.7

A lix B

D Used for ANN Models

Tr Set

No.	(	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress	SttImnt(mm)
1		32.0	179.0	52.24	88.93	100.00	19.56	13.90	2.30	120.00	400.0	1.1
2		29.0	208.0	56.55	89.78	100.00	20.33	13.90	2.30	120.00	300.0	2.3
3		39.0	196.0	53.21	91.02	100.00	19.69	13.83	2.88	120.00	300.0	2.6
4		29.0	146.0	61.32	100.00	100.00	20.56	13.54	4.55	120.00	400.0	2.8
5		30.0	357.0	1.31	6.56	26.51	15.50	24.40	3.78	17.27	400.0	11.8
6		34.0	170.0	1.36	71.21	100.00	17.21	13.60	5.10	100.65	400.0	12
7		34.0	323.0	2.67	20.97	55.72	16.44	19.53	3.91	35.83	198.8	10
8		33.0	344.0	2.36	11.03	29.79	15.50	24.40	4.07	22.46	400.0	10.9
9		40.4	199.2	2.93	3.79	3.21	16.31	60.77	1.37	8.79	38.1	23.3
10		13.0	196.0	0.88	4.20	12.00	15.71	48.71	0.52	12.75	13.5	46.2
11		26.0	233.7	1.04	6.14	14.00	15.64	40.20	0.43	13.94	34.0	53.2
12		27.0	242.3	2.38	9.42	24.50	16.37	40.70	1.08	17.68	32.0	7.2
13		30.0	147.0	49.24	88.33	100.00	19.56	13.90	2.30	120.00	400.0	1.5
14		28.0	131.0	59.05	100.00	100.00	20.56	13.70	4.90	120.00	400.0	1.6
15		18.0	166.0	24.29	94.29	97.17	19.57	15.27	2.94	120.00	400.0	4.3
16		23.0	70.0	8.44	67.50	79.07	19.14	15.93	3.37	98.77	190.0	7
17		27.0	52.0	7.14	82.00	83.00	18.60	16.00	2.40	120.00	150.0	3.7
18		26.0	61.0	6.96	82.71	93.96	19.17	16.97	1.88	120.00	50.0	1.5
19		25.0	60.0	10.18	66.58	73.92	19.14	16.51	2.94	98.30	200.0	6.6
20		39.0	360.0	2.35	23.17	52.29	16.42	18.09	3.38	36.76	400.0	8
21		24.0	244.0	3.65	16.18	99.95	15.50	58.75	2.35	35.89	200.0	52.9
22		25.0	265.0	3.63	16.16	99.93	16.30	58.73	2.33	35.87	200.0	16.9
23		26.0	162.0	57.91	100.00	100.00	20.56	13.67	4.90	120.00	400.0	4.1
24		28.0	157.0	65.48	100.00	100.00	20.78	13.25	3.90	120.00	400.0	2.4

25	35.0	227.3	3.78	12.11	33.18	17.11	41.84	1.93	20.21	33.0	47
26	41.0	249.3	3.33	6.76	16.27	16.74	45.36	2.69	10.79	31.5	25.6
27	35.0	217.0	2.56	4.39	11.00	16.24	51.77	3.05	6.48	33.0	36.4
28	39.0	341.0	2.78	6.06	0.54	16.35	61.35	1.00	7.44	51.0	44.7
29	43.0	237.0	2.93	10.40	1.36	16.46	50.78	2.67	5.26	46.0	39.1
30	43.0	225.3	3.25	11.36	2.29	16.54	44.47	3.82	5.17	39.0	33.3
31	37.1	217.2	2.78	6.06	0.54	16.35	61.35	1.00	7.44	37.8	15.6
32	15.3	93.6	80.27	100.00	100.00	20.73	6.08	2.56	120.00	350.0	2.1
33	9.8	109.7	76.13	100.00	100.00	20.63	6.81	2.80	120.00	351.0	6.5
34	41.0	228.7	3.25	11.36	2.29	16.54	44.47	3.82	5.17	35.0	21.2
35	40.0	214.2	3.45	11.42	2.83	16.60	41.17	4.56	5.13	38.2	19.9
36	47.9	235.6	52.51	100.00	100.00	19.88	17.82	3.40	120.00	300.0	4.8
37	30.0	152.0	49.24	88.33	100.00	19.56	13.90	2.30	120.00	300.0	3
38	30.0	70.0	9.72	66.83	75.29	19.14	16.35	3.05	98.43	300.0	1.8
39	46.5	99.2	25.67	91.98	100.00	18.77	14.94	3.55	120.00	350.0	3.6
40	46.4	98.6	23.76	91.92	100.00	18.81	16.15	3.37	120.00	321.0	5.5
41	51.0	83.5	22.28	91.88	100.00	18.83	17.09	3.23	120.00	320.0	6.4
42	29.0	235.0	3.26	26.62	100.00	15.50	50.26	2.91	48.16	400.0	20.9
43	45.0	252.0	3.23	0.00	0.00	16.64	65.80	2.83	6.06	38.0	15.4
44	22.0	320.0	2.50	44.33	100.00	16.72	35.70	3.78	69.01	400.0	76.4
45	31.0	350.0	3.40	28.67	32.00	16.83	18.70	2.93	38.40	500.0	13.9
46	38.0	350.0	1.85	20.54	62.00	16.44	17.80	3.60	35.98	400.0	8.2
47	20.0	177.0	66.45	100.00	100.00	20.67	13.60	4.80	120.00	400.0	0.7
48	27.0	227.0	62.58	99.41	100.00	20.44	15.18	2.39	120.00	400.0	0.8
49	31.0	139.0	64.44	94.32	100.00	20.56	13.45	2.55	120.00	400.0	2.6
50	45.0	223.7	3.24	8.45	2.09	16.58	47.89	3.95	5.28	45.0	29.4
51	29.0	193.0	61.60	99.57	100.00	20.44	14.76	2.23	120.00	400.0	0.8
52	26.0	188.0	73.69	100.00	100.00	20.44	15.76	3.24	120.00	400.0	0.8
53	25.0	191.1	2.38	9.42	24.50	16.37	40.70	1.08	17.68	37.3	31.7
54	26.3	217.4	3.78	12.11	33.18	17.11	41.84	1.93	20.21	34.4	42.9
55	27.7	216.6	3.33	6.76	16.27	16.74	45.36	2.69	10.79	33.0	27.9

56	29.0	187.0	64.10	100.00	100.00	20.67	13.60	3.87	120.00	350.0	2.4
57	49.7	162.0	46.94	100.00	100.00	19.65	12.33	3.23	120.00	290.0	6.5
58	50.0	164.3	49.45	100.00	100.00	19.80	12.33	3.13	120.00	300.0	9.8
59	26.0	178.0	62.21	100.00	100.00	20.44	13.60	3.12	120.00	400.0	1.3
60	47.8	88.3	21.22	91.84	100.00	18.85	17.76	3.13	120.00	354.0	9.9
61	17.8	120.4	74.93	100.00	100.00	20.59	7.36	2.83	120.00	318.0	9.8
62	21.2	96.7	70.33	100.00	100.00	20.44	9.51	2.95	120.00	300.0	4.1
63	20.5	100.4	65.90	100.00	100.00	20.30	11.58	3.06	120.00	323.0	3.5
64	15.0	117.4	61.89	100.00	100.00	20.18	13.45	3.16	120.00	319.0	3
65	43.1	220.2	3.24	8.45	2.09	16.58	47.89	3.95	5.28	33.1	22.7
66	45.4	241.1	3.05	5.71	1.42	16.56	54.09	3.38	5.42	31.2	15.9
67	28.0	168.0	51.07	100.00	100.00	20.56	13.49	4.90	120.00	300.0	1.3
68	22.0	210.0	32.87	97.36	100.00	19.63	14.38	3.72	120.00	400.0	3.3
69	46.1	88.5	18.54	88.51	100.00	18.57	17.61	2.93	115.40	323.0	2.9
70	17.9	123.0	80.33	100.00	100.00	20.68	7.79	2.36	120.00	370.0	2.7
71	23.1	173.9	0.84	4.70	12.00	15.63	45.48	0.44	12.93	28.0	112.9
72	25.0	203.0	65.20	99.00	100.00	20.11	16.30	2.80	120.00	400.0	0.2
73	29.0	152.0	69.41	98.11	100.00	20.78	13.02	2.78	120.00	400.0	2.2
74	37.0	370.0	4.35	17.41	8.92	15.40	19.85	2.52	20.10	500.0	11.5
75	9.5	108.7	6.86	23.97	30.78	17.86	19.47	2.76	35.46	158.0	98.5
76	17.0	11.0	10.61	72.41	79.32	19.00	17.50	1.95	106.70	100.0	0.4
77	15.0	99.2	80.33	100.00	100.00	20.68	7.79	2.36	120.00	335.0	3.4
78	43.0	244.0	2.66	0.23	0.06	16.53	66.48	2.25	5.69	38.5	28.8
79	29.0	348.0	0.66	3.83	24.50	15.50	24.40	3.60	14.10	400.0	12
80	27.0	315.0	4.59	14.59	3.15	15.40	20.14	2.42	15.53	370.0	12
81	38.0	334.7	2.92	3.50	6.50	16.43	48.72	3.16	5.00	30.5	39
82	18.1	200.9	0.88	4.20	12.00	15.71	48.71	0.52	12.75	1.8	110.4
83	47.6	225.4	2.66	0.23	0.06	16.53	66.48	2.25	5.69	38.3	15.6
84	52.1	240.6	2.83	2.74	0.68	16.55	60.80	2.77	5.56	34.7	20.2
85	23.0	73.7	3.22	0.40	4.00	16.20	57.00	5.25	8.46	35.0	9.7

86	8.5	19.0	142.0	9.76	78.33	100.00	17.80	16.80	0.50	112.50	400.0	2.1
87	30.0	14.6	107.9	67.97	100.00	100.00	20.03	9.78	2.39	120.00	350.0	1.1
88	12.2	26.0	43.0	10.09	62.52	59.01	18.80	17.35	2.01	92.09	700.0	2.4
89	12.8	23.7	187.1	1.04	6.14	14.00	15.64	40.20	0.43	13.94	27.7	46.4
90	24.0	20.3	119.6	80.26	100.00	100.00	20.73	5.95	2.58	120.00	362.0	5.1

Maximum	30.0	52.1	370.0	80.33	100.00	100.00	20.78	66.48	5.25	120.00	700.0	112.9
Minimum	8.5	9.5	11.0	0.66	0.00	0.00	15.40	5.95	0.43	5.00	1.8	0.2
Average	16.9	30.6	189.0	26.59	54.28	62.44	18.17	27.51	2.83	70.56	247.2	16.8
StdDev	3.9	10.4	83.6	28.80	42.18	42.45	1.92	18.19	1.09	51.29	166.1	22.9

Testing set

No.	Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress	SttImnt(mm)
1	11.6	43.0	243.3	3.81	0.00	0.00	16.75	64.60	3.46	6.42	36.0	22.9
2	16.4	38.0	224.0	2.61	9.44	0.43	16.39	57.09	1.53	5.36	56.0	19.5
3	12.0	42.8	261.2	3.23	0.00	0.00	16.64	65.80	2.83	6.06	37.7	7.3
4	11.6	42.2	258.1	3.81	0.00	0.00	16.75	64.60	3.46	6.42	37.7	11.9
5	22.0	26.5	155.0	44.74	100.00	100.00	18.87	11.10	2.03	120.00	270.0	3.8
6	15.1	27.0	340.0	4.08	15.34	60.42	15.50	43.21	2.40	27.33	400.0	19.5
7	25.0	12.3	133.0	20.95	58.15	70.26	18.13	21.41	2.76	78.00	180.0	58.2
8	25.0	16.3	127.8	23.64	55.71	64.48	18.27	22.52	2.09	73.80	320.0	47.2
9	19.7	31.0	197.0	63.74	100.00	100.00	20.67	13.60	3.73	120.00	350.0	2.5
10	16.4	35.1	207.6	2.61	9.44	0.43	16.39	57.09	1.53	5.36	36.9	18.1
11	16.0	40.7	210.1	2.93	10.40	1.36	16.46	50.78	2.67	5.26	41.1	21.7
12	22.0	11.8	113.8	70.33	100.00	100.00	20.44	9.51	2.95	120.00	320.0	4.1
13	23.0	12.6	115.7	67.09	100.00	100.00	20.34	11.02	3.03	120.00	325.0	4.3
14	22.0	10.5	112.0	59.67	100.00	100.00	20.10	14.48	3.22	120.00	347.0	4.5
15	22.0	41.4	204.5	53.53	100.00	100.00	19.91	17.34	3.37	120.00	328.0	1.0

16	18.0	50.0	109.7	28.20	87.51	100.00	19.24	16.09	4.59	108.94	336.0	5.6
17	18.3	22.0	347.0	6.50	92.63	100.00	18.39	11.70	1.90	120.00	400.0	1.0
18	19.2	30.0	198.0	60.17	91.06	100.00	20.33	13.81	2.35	120.00	300.0	1.7
19	11.2	48.0	251.0	4.44	0.00	0.00	16.87	63.30	4.14	6.81	36.5	32.7
20	11.2	23.0	73.7	3.64	0.28	2.77	16.42	58.85	4.96	7.98	35.0	23.0
21	16.4	41.0	227.3	2.71	4.21	7.32	16.24	57.57	2.05	8.09	42.0	48.3
22	19.7	26.0	162.0	63.38	100.00	100.00	20.67	13.60	3.59	120.00	400.0	1.3
23	29.0	15.2	135.5	70.42	100.00	100.00	20.14	10.16	2.29	120.00	331.0	1.3
24	29.0	15.1	127.3	75.57	100.00	100.00	20.36	10.95	2.09	120.00	324.0	2.8
25	15.0	29.0	340.0	4.64	14.03	2.00	15.50	20.20	2.40	14.61	500.0	10.3
26	11.0	31.0	70.0	6.86	83.10	100.00	19.17	17.50	1.60	120.00	300.0	0.6
27	15.7	18.0	62.0	2.99	70.36	95.20	17.21	14.14	4.70	100.22	200.0	8.1
28	17.6	25.0	205.0	57.66	86.96	88.00	19.45	16.64	2.76	106.34	500.0	0.5
29	25.0	13.0	140.0	17.87	60.93	76.87	17.97	20.13	3.52	82.80	250.0	44.5
30	18.0	17.0	233.0	38.49	99.30	100.00	19.63	13.52	4.58	120.00	300.0	3.4
31	17.4	25.0	323.0	4.54	14.26	12.37	15.50	24.28	2.40	16.87	300.0	8.1
32	13.2	49.0	228.3	2.83	2.74	0.68	16.55	60.80	2.77	5.56	42.5	32.3
33	14.0	45.0	214.7	3.05	5.71	1.42	16.56	54.09	3.38	5.42	43.0	28.0
34	23.0	48.0	157.3	49.28	100.00	100.00	19.68	12.12	2.92	120.00	273.0	4.2
35	22.0	41.1	190.7	46.69	100.00	100.00	19.22	11.54	2.41	120.00	294.0	1.7
36	23.0	37.0	162.7	45.83	100.00	100.00	19.07	11.35	2.24	120.00	272.0	2.1
37	16.8	27.0	166.0	59.97	90.91	100.00	20.04	13.83	2.34	120.00	400.0	7.3
38	20.3	28.0	159.0	65.20	99.00	100.00	20.11	16.30	2.80	120.00	350.0	0.3
39	25.0	11.0	123.3	6.78	20.35	26.61	17.75	21.45	3.52	30.71	172.0	94.4

Maximum	29.0	50.0	347.0	75.57	100.00	100.00	20.67	65.80	4.96	120.00	500.0	94.4
Minimum	11.0	10.5	62.0	2.61	0.00	0.00	15.50	9.51	1.53	5.26	35.0	0.3
Average	18.7	29.4	187.4	29.60	58.51	61.81	18.30	28.41	2.91	73.55	243.2	15.6
Std.Dev.	4.9	12.4	73.6	27.08	42.91	45.13	1.71	20.68	0.88	51.93	144.5	20.2

Validation set

No.	Cover	Adv.Rt.	EP	SPT1	SPT2	SPT3	BD	MC	GWL	E	GrPress	SttImnt(mm)
1	11.2	40.8	267.5	4.44	0.00	0.00	16.87	63.30	4.14	6.81	43.0	12.0
2	16.4	32.8	233.2	2.71	4.21	7.32	16.24	57.57	2.05	8.09	39.4	24.2
3	17.3	23.0	338.0	4.40	14.59	27.07	15.50	30.07	2.40	20.07	300.0	19.7
4	12.0	46.3	171.0	11.67	78.67	100.00	17.72	16.47	2.41	101.78	344.0	1.2
5	12.0	47.5	173.7	9.44	75.47	100.00	17.44	16.09	2.24	97.34	326.0	1.3
6	11.0	43.9	162.0	7.68	68.62	95.85	17.08	15.40	2.37	93.56	328.0	1.1
7	17.2	35.0	251.0	2.93	3.79	3.21	16.31	60.77	1.37	8.79	32.0	56.1
8	16.0	34.0	202.7	2.45	4.68	12.05	16.17	53.87	2.84	7.29	33.0	41.4
9	12.8	16.0	221.3	0.84	4.70	12.00	15.63	45.48	0.44	12.93	30.0	57.9
10	20.0	33.3	178.9	39.04	100.00	100.00	19.20	12.32	3.54	120.00	330.0	4.1
11	21.0	32.0	180.7	43.43	100.00	100.00	19.45	12.33	3.37	120.00	283.0	6.0
12	19.4	28.0	196.0	51.02	88.68	100.00	19.56	13.90	2.30	120.00	350.0	2.4
13	17.7	30.0	338.0	2.81	25.56	43.47	16.44	18.36	3.18	37.48	350.0	7.8
14	29.0	14.9	142.6	73.00	100.00	100.00	20.25	10.55	2.19	120.00	317.0	2.2
15	17.2	22.0	316.0	4.27	14.90	40.64	16.30	35.42	2.40	23.02	400.0	16.3
16	14.9	35.0	330.0	3.88	23.04	20.46	16.74	19.28	2.72	29.25	400.0	15.5
17	17.9	24.0	230.0	77.39	100.00	100.00	20.44	16.30	3.60	120.00	500.0	1.1
18	14.4	24.0	340.0	3.70	16.23	100.00	15.50	58.80	2.40	35.94	400.0	20.1
19	17.7	26.0	240.0	60.40	99.76	100.00	20.44	14.25	2.04	120.00	500.0	0.9
20	17.8	21.0	190.0	58.87	100.00	100.00	20.44	13.60	1.80	120.00	400.0	1.2
21	28.0	12.4	116.3	78.03	100.00	100.00	20.46	11.32	2.00	120.00	306.0	3.8
22	15.2	24.0	340.0	3.97	15.59	71.73	16.78	47.67	2.40	29.79	300.0	17.4
23	13.0	30.0	90.0	7.86	67.80	80.78	19.14	15.74	3.51	98.92	200.0	10.2
24	21.0	41.0	181.0	36.36	96.51	100.00	19.38	12.55	4.44	120.00	286.0	3.5

25	25.0	10.4	117.4	6.71	17.63	23.48	17.66	22.94	4.10	27.15	178.0	79.9
26	15.2	33.8	211.9	2.92	3.50	6.50	16.43	48.72	3.16	5.00	33.4	22.3
27	15.6	42.6	211.9	2.56	4.39	11.00	16.24	51.77	3.05	6.48	34.2	16.1
28	16.0	39.5	214.1	2.45	4.68	12.05	16.17	53.87	2.84	7.29	32.5	25.1
29	14.8	47.0	218.3	3.45	11.42	2.83	16.60	41.17	4.56	5.13	40.0	23.1

Maximum	29.0	47.5	340.0	78.03	100.00	100.00	20.46	63.30	4.56	120.00	500.0	79.9
Minimum	11.0	10.4	90.0	0.84	0.00	0.00	15.50	10.55	0.44	5.00	30.0	0.9
Average	17.1	30.7	220.8	20.99	46.36	57.60	17.67	30.69	2.75	60.07	245.4	17.0
Std.Dev.	4.5	10.5	70.9	26.48	41.70	42.49	1.74	18.91	0.92	49.99	158.0	19.5

Appendix C

Membership Values of Fuzzy Clustering

Number of clusters: 8

Patterns 1 through 8

	1	2	3	4	5	6	7	8
1	0.0008	0.0089	0.004	0.0032	0.0038	0.0055	0.0137	0.1386
2	0.0138	0.0848	0.047	0.0847	0.1092	0.0093	0.1627	0.0628
3	0.0021	0.024	0.012	0.009	0.0111	0.0059	0.0365	0.0721
4	0.0008	0.009	0.0041	0.0032	0.0039	0.0052	0.0139	0.1179
5	0.003	0.029	0.012	0.0098	0.0113	0.0984	0.0544	0.3305
6	0.0031	0.0191	0.0081	0.01	0.0116	0.8481	0.0574	0.1088
7	0.0129	0.75	0.877	0.048	0.0571	0.013	0.1644	0.1071
8	0.9635	0.0752	0.0359	0.8321	0.7921	0.0145	0.497	0.0622

Patterns 9 through 16

	9	10	11	12	13	14	15	16
1	0.0026	0.0276	0.2197	0.2364	0.2889	0.4003	0.6372	0.0039
2	0.0047	0.0737	0.0032	0.0069	0.005	0.0043	0.005	0.1059
3	0.0029	0.0372	0.0077	0.0166	0.0113	0.0092	0.0106	0.0116
4	0.0024	0.0264	0.7564	0.7126	0.6735	0.5677	0.3253	0.004
5	0.0506	0.2292	0.004	0.0083	0.0067	0.006	0.0071	0.0123
6	0.9226	0.3524	0.0024	0.0052	0.0041	0.0036	0.0042	0.0125
7	0.0067	0.1141	0.0041	0.0087	0.0065	0.0056	0.0066	0.0599
8	0.0075	0.1394	0.0024	0.0053	0.004	0.0034	0.004	0.7899

Patterns 17 through 24

	17	18	19	20	21	22	23	24
1	0.0232	0.0066	0.0055	0.0023	0.0521	0.0302	0.136	0.0075
2	0.1339	0.222	0.1006	0.006	0.1751	0.0518	0.0853	0.0226
3	0.0429	0.0205	0.0155	0.9775	0.0949	0.7734	0.3821	0.9203
4	0.0231	0.0068	0.0056	0.0027	0.053	0.0359	0.1703	0.0087
5	0.0927	0.0184	0.0195	0.0018	0.1107	0.0182	0.0475	0.0061
6	0.1432	0.0186	0.0203	0.0012	0.1434	0.0129	0.0334	0.0043
7	0.1301	0.0919	0.0834	0.0058	0.1415	0.0513	0.0925	0.0206
8	0.4109	0.6153	0.7496	0.0027	0.2293	0.0261	0.053	0.0098

	Patterns 25	through 32					
	25	26	27	28	29	30	31	32
1	0.0042	0.025	0.1623	0.1599	0.0058	0.0091	0.0066	0.0011
2	0.0076	0.0967	0.0818	0.0752	0.0122	0.01	0.3524	0.0269
3	0.0046	0.0428	0.1192	0.1007	0.007	0.0078	0.0216	0.0032
4	0.0039	0.0246	0.1556	0.1436	0.0055	0.0083	0.0068	0.0012
5	0.076	0.2162	0.1955	0.2373	0.1164	0.8778	0.0189	0.0037
6	0.8802	0.2807	0.0814	0.088	0.8158	0.058	0.0162	0.0039
7	0.011	0.1296	0.135	0.1285	0.0172	0.0164	0.2682	0.0181
8	0.0124	0.1844	0.0692	0.0668	0.0202	0.0126	0.3093	0.9419

	Patterns 33	through 40						
	33	34	35	36	37	38	39	40
1	0.3284	0.8391	0.4919	0.09	0.2115	0.4642	0.5701	0.7226
2	0.003	0.0018	0.027	0.0005	0.0047	0.031	0.0033	0.0019
3	0.0067	0.0037	0.0431	0.0012	0.011	0.0479	0.0073	0.004
4	0.6492	0.1472	0.294	0.9059	0.7538	0.2837	0.405	0.263
5	0.004	0.0027	0.0526	0.0007	0.0059	0.0653	0.0046	0.0028
6	0.0024	0.0016	0.0309	0.0004	0.0036	0.0373	0.0027	0.0016
7	0.0039	0.0024	0.0367	0.0007	0.0059	0.0428	0.0044	0.0025
8	0.0023	0.0015	0.0238	0.0004	0.0036	0.0277	0.0026	0.0015

	Patterns 41	through 48						
	41	42	43	44	45	46	47	48
1	0.409	0.3639	0.4741	0.2877	0.7199	0.0132	0.006	0.0007
2	0.0017	0.0026	0.0018	0.0017	0.0016	0.5218	0.0703	0.0148
3	0.0038	0.0059	0.004	0.004	0.0035	0.0609	0.0174	0.0019
4	0.5782	0.6167	0.5122	0.6993	0.2678	0.014	0.0061	0.0007
5	0.0023	0.0034	0.0025	0.0023	0.0023	0.0272	0.0178	0.0023
6	0.0014	0.0021	0.0015	0.0014	0.0014	0.0243	0.0121	0.0024
7	0.0022	0.0034	0.0024	0.0023	0.0021	0.166	0.8165	0.0109
8	0.0013	0.002	0.0014	0.0013	0.0013	0.1725	0.0539	0.9663

	Patterns 49	through 56						
	49	50	51	52	53	54	55	56
1	0.0009	0.002	0.0049	0.9313	0.8748	0.2352	0.2351	0.4185
2	0.0211	0.0504	0.0625	0.0008	0.0017	0.0567	0.0575	0.0045
3	0.0025	0.0056	0.0115	0.0016	0.0035	0.1866	0.1896	0.0097
4	0.0009	0.002	0.005	0.063	0.1121	0.3434	0.3378	0.548
5	0.003	0.0063	0.0204	0.0011	0.0026	0.0455	0.0459	0.0062
6	0.0031	0.0065	0.0219	0.0007	0.0016	0.0312	0.0315	0.0037
7	0.0145	0.0312	0.0725	0.001	0.0023	0.0621	0.063	0.0059
8	0.954	0.8958	0.8013	0.0006	0.0014	0.0392	0.0397	0.0035

Patterns 57 through 64

	57	58	59	60	61	62	63	64
1	0.5824	0.0038	0.0063	0.0034	0.0066	0.0129	0.014	0.0155
2	0.0066	0.7658	0.5916	0.8222	0.7339	0.2039	0.1789	0.2216
3	0.0133	0.0152	0.0255	0.0145	0.0324	0.0409	0.0461	0.0551
4	0.3677	0.004	0.0066	0.0036	0.007	0.0131	0.0143	0.016
5	0.01	0.0088	0.0141	0.0073	0.0122	0.0447	0.0466	0.0486
6	0.0059	0.0072	0.0111	0.0059	0.0096	0.036	0.0342	0.0364
7	0.0087	0.123	0.2568	0.093	0.1382	0.4087	0.5072	0.4353
8	0.0054	0.0722	0.088	0.05	0.0601	0.24	0.1588	0.1715

Patterns 65 through 72

	65	66	67	68	69	70	71	72
1	0.113	0.1258	0.1248	0.3328	0.2978	0.184	0.1607	0.1723
2	0.0971	0.0981	0.0931	0.0234	0.0285	0.0061	0.0044	0.0022
3	0.3693	0.326	0.3553	0.0513	0.0686	0.0154	0.0113	0.005
4	0.1396	0.1563	0.1593	0.4992	0.4967	0.771	0.8064	0.8117
5	0.0667	0.0714	0.063	0.0281	0.0317	0.007	0.0051	0.0028
6	0.0416	0.0448	0.0407	0.0182	0.0204	0.0044	0.0031	0.0017
7	0.1128	0.1152	0.1055	0.0287	0.0346	0.0076	0.0056	0.0028
8	0.0598	0.0624	0.0583	0.0183	0.0216	0.0046	0.0033	0.0017

Patterns 73 through 80

	73	74	75	76	77	78	79	80
1	0.1338	0.1833	0.3304	0.2151	0.4586	0.2498	0.2257	0.6009
2	0.001	0.0015	0.0026	0.0013	0.0026	0.0025	0.002	0.0019
3	0.0024	0.0034	0.006	0.003	0.0059	0.006	0.0046	0.0042
4	0.8586	0.8057	0.65	0.7751	0.5218	0.7313	0.7595	0.3846
5	0.0013	0.0019	0.0034	0.0017	0.0035	0.0033	0.0026	0.0027
6	0.0008	0.0012	0.0021	0.001	0.0021	0.002	0.0015	0.0016
7	0.0013	0.0019	0.0034	0.0017	0.0034	0.0033	0.0026	0.0025
8	0.0008	0.0011	0.002	0.001	0.002	0.002	0.0015	0.0015

Patterns 81 through 88

	81	82	83	84	85	86	87	88
1	0.2733	0.0204	0.0038	0.0054	0.0042	0.0053	0.004	0.0112
2	0.0024	0.1162	0.0444	0.1279	0.0518	0.1001	0.0872	0.1591
3	0.0055	0.0438	0.0136	0.0175	0.0154	0.022	0.0153	0.0515
4	0.7088	0.0202	0.0039	0.0056	0.0043	0.0055	0.0041	0.0117
5	0.0031	0.0776	0.0092	0.017	0.01	0.0117	0.0093	0.0212
6	0.0019	0.0494	0.0062	0.0132	0.0066	0.0082	0.0067	0.0145
7	0.0031	0.5385	0.8943	0.6963	0.8808	0.8077	0.8365	0.6725
8	0.0018	0.134	0.0246	0.1171	0.0269	0.0395	0.0369	0.0583

Patterns 89 through 96

	89	90	91	92	93	94	95	96
1	0.003	0.0099	0.0122	0.0392	0.0691	0.0236	0.004	0.0061
2	0.0362	0.133	0.1838	0.193	0.0953	0.3263	0.8436	0.762
3	0.0097	0.0444	0.0615	0.2604	0.4996	0.1022	0.0182	0.0258
4	0.0031	0.0103	0.0129	0.043	0.0821	0.0251	0.0043	0.0064
5	0.0083	0.0193	0.0227	0.0587	0.0479	0.0554	0.0079	0.0123
6	0.0056	0.013	0.0154	0.0368	0.0293	0.0418	0.0068	0.0109
7	0.91	0.719	0.6304	0.2809	0.1252	0.2692	0.0617	0.0814
8	0.0241	0.0511	0.0611	0.088	0.0514	0.1563	0.0535	0.095

Patterns 97 through 104

	97	98	99	100	101	102	103	104
1	0.0042	0.0023	0.0012	0.0023	0.012	0.01	0.0105	0.0117
2	0.802	0.9	0.9445	0.8845	0.1389	0.5666	0.5683	0.5639
3	0.017	0.0108	0.0055	0.0098	0.0315	0.0475	0.0479	0.066
4	0.0044	0.0024	0.0012	0.0024	0.0122	0.0105	0.0111	0.0124
5	0.0092	0.0045	0.0024	0.0051	0.0426	0.0217	0.0225	0.0222
6	0.0081	0.0037	0.002	0.0044	0.0322	0.0182	0.0198	0.0182
7	0.0696	0.0492	0.0271	0.0442	0.5517	0.1954	0.1626	0.1975
8	0.0854	0.0272	0.0161	0.0472	0.1789	0.1302	0.1574	0.1079

Patterns 105 through 112

	105	106	107	108	109	110	111	112
1	0.0152	0.9189	0.8631	0.0207	0.0216	0.0024	0.0237	0.0112
2	0.5213	0.0009	0.0023	0.0489	0.05	0.0046	0.0484	0.1011
3	0.0905	0.0018	0.0045	0.0267	0.0272	0.0027	0.0281	0.0235
4	0.0162	0.0743	0.1196	0.0199	0.0208	0.0023	0.0227	0.0111
5	0.0267	0.0013	0.0035	0.2469	0.1469	0.0465	0.1562	0.0531
6	0.0221	0.0008	0.0021	0.4811	0.5837	0.9275	0.5806	0.0592
7	0.1904	0.0012	0.003	0.0712	0.0606	0.0066	0.0589	0.1411
8	0.1177	0.0007	0.0019	0.0846	0.0893	0.0075	0.0814	0.5997

Patterns 113 through 120

	113	114	115	116	117	118	119	120
1	0.0053	0.0107	0.0089	0.0055	0.0243	0.0011	0.0024	0.0084
2	0.1636	0.0119	0.0253	0.1976	0.381	0.0201	0.0595	0.0893
3	0.0159	0.0093	0.9048	0.0243	0.1954	0.0028	0.0093	0.0194
4	0.0054	0.0098	0.0102	0.0057	0.0264	0.0011	0.0025	0.0084
5	0.0154	0.8559	0.0074	0.0123	0.0377	0.0038	0.006	0.0387
6	0.0156	0.0677	0.005	0.009	0.0299	0.0039	0.0043	0.0423
7	0.0765	0.0197	0.0273	0.6937	0.1926	0.0166	0.8917	0.1185
8	0.7023	0.0151	0.0111	0.0519	0.1127	0.9507	0.0243	0.675

Patterns 121 through 128

	121	122	123	124	125	126	127	128
1	0.0144	0.007	0.0059	0.0033	0.0085	0.0022	0.0017	0.3899
2	0.5251	0.5547	0.6601	0.8448	0.7105	0.9137	0.9185	0.0018
3	0.084	0.0243	0.0216	0.0151	0.0504	0.011	0.0081	0.004
4	0.0154	0.0073	0.0062	0.0035	0.0091	0.0023	0.0018	0.5967
5	0.0265	0.0171	0.014	0.0066	0.0142	0.0041	0.0034	0.0024
6	0.022	0.0158	0.0126	0.0053	0.0113	0.0034	0.0027	0.0014
7	0.1926	0.1152	0.1031	0.0816	0.1335	0.0388	0.043	0.0023
8	0.12	0.2587	0.1765	0.0399	0.0624	0.0245	0.0207	0.0014

Patterns 129 through 136

	129	130	131	132	133	134	135	136
1	0.8733	0.8847	0.6931	0.78	0.8208	0.7697	0.8345	0.0108
2	0.0013	0.0013	0.0027	0.005	0.0035	0.0057	0.0009	0.0203
3	0.0027	0.0027	0.0058	0.0096	0.0069	0.0108	0.0019	0.0125
4	0.1169	0.1055	0.2869	0.1821	0.1523	0.1868	0.1587	0.0103
5	0.0019	0.0019	0.0037	0.0079	0.0056	0.0093	0.0013	0.4005
6	0.0011	0.0011	0.0022	0.0046	0.0033	0.0054	0.0008	0.4842
7	0.0017	0.0017	0.0035	0.0067	0.0047	0.0077	0.0012	0.0303
8	0.0011	0.001	0.0021	0.0041	0.0029	0.0047	0.0007	0.031

Patterns 137 through 144

	137	138	139	140	141	142	143	144
1	0.0165	0.0247	0.027	0.0025	0.0029	0.0005	0.0091	0.0081
2	0.0406	0.0614	0.0731	0.0364	0.0459	0.009	0.1581	0.1799
3	0.0212	0.0321	0.0364	0.0061	0.0072	0.0013	0.0242	0.0226
4	0.0157	0.0234	0.0258	0.0025	0.0029	0.0005	0.0092	0.0082
5	0.2205	0.2769	0.229	0.0097	0.0108	0.0017	0.0333	0.0276
6	0.5475	0.3696	0.3555	0.0103	0.0114	0.0018	0.0281	0.0233
7	0.0625	0.1009	0.1134	0.0389	0.0441	0.0078	0.3708	0.358
8	0.0755	0.1109	0.1398	0.8936	0.8748	0.9774	0.3672	0.3723

Patterns 145 through 152

	145	146	147	148	149	150	151	152
1	0.0089	0.0082	0.0117	0.0055	0.0019	0.0817	0.007	0.0236
2	0.1509	0.0138	0.2137	0.0095	0.0039	0.0889	0.0133	0.1282
3	0.024	0.0088	0.0356	0.0059	0.0023	0.5263	0.0079	0.0428
4	0.009	0.0077	0.012	0.0052	0.0018	0.1	0.0066	0.0234
5	0.0343	0.6729	0.0373	0.0999	0.0393	0.0389	0.1293	0.0997
6	0.0285	0.246	0.0381	0.8462	0.9387	0.0285	0.7954	0.1618
7	0.3904	0.0221	0.1439	0.0132	0.0056	0.0852	0.0188	0.13
8	0.3541	0.0205	0.5077	0.0147	0.0065	0.0506	0.0216	0.3905

Patterns 153 through 160

	153	154	155	156	157	158	159	160
1	0.0019	0.0225	0.0222	0.0077	0.0212	0.0032	0.0198	0.0091
2	0.0298	0.0503	0.4083	0.0229	0.1357	0.0061	0.1192	0.0107
3	0.0048	0.0281	0.1761	0.919	0.0415	0.0036	0.045	0.0083
4	0.002	0.0217	0.0239	0.0089	0.0212	0.003	0.0198	0.0083
5	0.0075	0.1509	0.0341	0.0062	0.0839	0.0646	0.0785	0.8775
6	0.0079	0.5797	0.0273	0.0044	0.1287	0.901	0.0487	0.055
7	0.0303	0.0599	0.1958	0.021	0.1235	0.0086	0.5356	0.0177
8	0.9157	0.0869	0.1123	0.01	0.4444	0.0098	0.1335	0.0134

Pattern 161

	161
1	0.0032
2	0.8498
3	0.0131
4	0.0033
5	0.0069
6	0.006
7	0.0543
8	0.0633

Appendix D

Calculation of settlement using empirical method

1) Tunnel geometry

Excavated tunnel diameter	D =	6
Excavated tunnel radius	r =	3
Depth of tunnel axis	z_0 =	18.94
Cover to tunnel crown	C =	15.94
Volume excavated per metre r	V =	28.28

2) Subsurface soil conditions

Using the values reccomended by Attewell (1981)

Depth from ground (m)		Layer thickness (m)	Description		K	n
from	to					
0	2	2	fill; debris	Sand	0.63	0.97
2	8	6	fine silty	Sand	0.63	0.97
8	12	4	soft marine	Clay	1	1
12	18.94	6.94	fine silty	Sand	0.63	0.97
total		18.94				

Representative K: 0.71

Representative n: 0.98

where, by linear interpolation,

$K = (CK_c + SK_s) / (C + S)$

$n = (Cn_c + Sn_s) / (C + S)$

K_c = K of Clay K_s = K of Sand

n_c = n of Clay n_s = n of Sand

C = total thickness of Clay layer

S = total thickness of Sand layer

3) Settlement trough prediction

Volume Loss

Assumed volume loss, $V_S(\%)$ = 3

Estimation of Point of Inflection, i Attewell (1977)

$(i / r) = K (z_0 / 2r)^n$

Location of i = 6.527 m

Estimation of Maximum Settlement over crown, W_{max}

$V_S(\%) = \{[i * (2\pi)^{0.5} * W_{max}] / V\} \times 100$

Maximum settlement (over crown of each tunnel), W_{max} = 51.84 mm

Appendix E

Final weights and bias terms for optimum ANN models ANNS12GDM and ANNi6CGF

Final weights of ANNS12GDM

Hidden nodes	Weights						
	K_0	γ	c	E/c	H/D	VL	S_{max}
Hidden 1	0.866	0.392	0.647	-0.104	0.150	-0.292	0.894
Hidden 2	-0.405	-0.268	3.868	0.184	-1.083	1.239	-2.495
Hidden 3	-0.400	-0.176	3.481	0.174	-1.255	1.552	1.914
Hidden 4	-0.750	0.372	-0.213	0.094	1.036	0.220	0.317
Hidden 5	-0.079	0.147	-1.418	1.143	0.376	-0.131	-0.606
Hidden 6	0.608	0.173	0.441	-0.081	-0.309	0.095	-1.763
Hidden 7	-0.812	-0.195	0.686	0.312	1.261	-0.827	-0.391
Hidden 8	0.064	-0.058	0.313	0.048	-0.494	2.049	1.331
Hidden 9	-1.031	0.571	0.122	0.638	-0.752	0.416	0.103
Hidden 10	-0.675	0.521	0.220	0.066	0.443	0.384	-0.254
Hidden 11	0.788	1.143	-0.482	-0.275	-0.194	-0.649	0.005
Hidden 12	0.521	-0.479	-1.024	-0.141	1.103	-0.659	-0.046

Bias terms of ANNS12GDM

Hidden and Output nodes	Bias
Hidden 1	0.474
Hidden 2	0.441
Hidden 3	0.399
Hidden 4	1.029
Hidden 5	0.984
Hidden 6	0.571
Hidden 7	0.443
Hidden 8	1.908
Hidden 9	0.453
Hidden 10	0.574
Hidden 11	-0.619
Hidden 12	-0.298
Output 1	-1.872

Final weights of ANNi6CGF

Hidden nodes	Weights						
	K_0	γ	c	E/c	H/D	VL	i
Hidden 1	1.894	0.074	-2.293	-0.020	0.756	0.715	-2.117
Hidden 2	0.006	0.002	0.016	0.000	0.539	0.006	3.951
Hidden 3	-2.240	0.123	-2.470	0.244	0.306	0.923	-0.188
Hidden 4	-1.258	-0.095	2.024	0.024	-1.102	-0.118	-4.997
Hidden 5	-1.605	-0.181	2.586	0.110	-1.363	0.152	1.755
Hidden 6	-0.472	-0.239	4.169	0.046	0.567	-0.109	-2.245

Bias terms of ANNi6CGF

Hidden and Output nodes	Bias
Hidden 1	-5.490
Hidden 2	1.292
Hidden 3	-4.158
Hidden 4	4.267
Hidden 5	4.701
Hidden 6	5.518
Output 1	-0.938

Appendix F

MatLab Neural Network Program Code

```
Training_Input = xlsread('Data.xls','trainI');
Training_Target = xlsread('Data.xls','trainT');
Testing_Input = xlsread('Data.xls','testI');
Testing_Target = xlsread('Data.xls','testT');
Valid_Input = xlsread('Data.xls','validI');
Valid_Target = xlsread('Data.xls','validT');


[Training_Inputn,minTraining_Input,maxTraining_Input,Training_Targetn,minTrainin
g_Target,maxTraining_Target] = premnmx(Training_Input,Training_Target);



val.P =Testing_Input ;
val.T = Testing_Target;
test.P = Valid_Input ;
test.T = Valid_Target ;

net=newff(minmax(Training_Inputn),[2,1],{'tansig','tansig'},'traingdm');
net.initFcn =  'initlay';
net.layers{1}.initFcn = 'initwb';
net.layers{2}.initFcn = 'initwb';
net.biases{1}.initFcn = 'rands';
net.inputWeights{1,1}.initFcn = 'rands';
net.biases{2}.initFcn = 'rands';
net.layerWeights{2,1}.initFcn = 'rands';
net = init(net);
wts1 = net.IW{1,1};
wts2 = net.LW{2,1};
bias1 = net.b{1};
bias2 = net.b{2};


net.trainParam.show = 1000;
net.trainParam.lr = 0.2;
net.trainParam.mc = 0.9;
net.trainParam.epochs = 40000;
net.trainParam.goal = 1e-4;
net.trainParam.max_fail = 30000;
net.adaptParam.passes = 100;
```

```
[net,tr]=train(net,Training_Inputn,Training_Targetn,[],[],val,test);


network1_outputs = sim(net,Training_Inputn);
network1_outputsTe = sim(net,Testing_Input);
network1_outputsV = sim(net,Valid_Input);

Training_Output = postmnmx(network1_outputs,minTraining_Target,maxTraining_Target);
Error_Training = Training_Target-Training_Output;
perf = mae(Error_Training)

Testing_Output = postmnmx(network1_outputsTe,minTraining_Target,maxTraining_Target);
Testing_Targetn = postmnmx(Testing_Target,minTraining_Target,maxTraining_Target);
Error_Testing =Testing_Targetn-Testing_Output;
perf=mae(Error_Testing)

Valid_Output = postmnmx(network1_outputsV,minTraining_Target,maxTraining_Target);
Valid_Targetn = postmnmx(Valid_Target,minTraining_Target,maxTraining_Target);
Error_Valid = Valid_Targetn-Valid_Output;
perf=mae(Error_Valid)

[m,b,r] = postreg(Training_Output,Training_Target)
[m,b,r] = postreg(Testing_Output,Testing_Targetn)
[m,b,r] = postreg(Valid_Output,Valid_Targetn)

Fwts1 = net.IW{1,1};
Fwts2 = net.LW{2,1};
Fbias1 = net.b{1};
Fbias2 = net.b{2};
```