USE OF ARTIFICIAL NEURAL NETWORKS FOR PREDICTING SETTLEMENT OF SHALLOW FOUNDATIONS ON COHESIONLESS SOILS

By

Mohamed A. Shahin
B.Sc., M.Sc., MIEAust

THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSPHY

The University of Adelaide
School of Civil and Environmental Engineering

January, 2003

To my wife Fifi

and my parents Amin and Laila

Contents

Abstract

The research presented in this thesis focuses on the settlement prediction of shallow foundations on cohesionless soils using artificial neural techniques. The problem of estimating the settlement of shallow foundations on cohesionless soils is very complex and not yet entirely understood. Over the years, many methods have been developed to predict the settlement of shallow foundations on cohesionless soils. However, methods for such predictions that have the required degree of accuracy and consistency have not yet been developed. Accurate prediction of settlement is essential since settlement, rather than bearing capacity, generally controls the design process of shallow foundations. In this research, artificial neural networks (ANNs) are used in an attempt to obtain more accurate settlement prediction. ANNs are numerical modelling techniques that are inspired by the functioning of the human brain and nerve system. ANNs use the data alone to determine the structure of the model as well as the unknown model parameters. ANNs have been applied successfully to many problems in the field of geotechnical engineering and some of their applications are demonstrated in this thesis.

A large database comprising a total of 189 case records is used to develop and verify the ANN models. Five parameters are considered to have the most significant impact on the settlement of shallow foundations on cohesionless soils and are thus used as the ANN model inputs. These include the 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 model output is the average measured settlement of the foundation, considered in its final state. Two types of ANNs are used for the development of ANN models. The first type is multi-layer perceptrons (MLPs) that are trained using the back-propagation algorithm, whereas the second type are B-spline neurofuzzy networks that are trained with the adaptive spline modelling of observation data (ASMOD) algorithm. In relation to the multi-layer perceptrons, the feasibility of ANNs for predicting the settlement of shallow foundations on cohesionless soils is investigated. A number of issues in relation to ANN construction, optimisation and validation are also investigated and guidelines for improving ANN performance are developed. The issue of data division and its impact on ANN model performance is investigated in some detail by examining four different data division methods, namely, random data division; data division to ensure statistical consistency of the subsets needed for ANN model development; data division using self-organising maps (SOMs) and a new data division method using fuzzy clustering. The success or otherwise of ANNs for settlement prediction of shallow foundations on cohesionless soils is illustrated and compared with three of the most commonly used settlement prediction methods. A hand-calculation design formula for settlement prediction of shallow foundations on cohesionless soils that is based on a more accurate settlement prediction from ANN model is presented. It was found that ANNs have the ability to predict the settlement of shallow foundations on cohesionless soils with a high degree of accuracy and outperform traditional methods. It was also found that the new data division method that is based on fuzzy clustering is suitable approach for data division. In relation to the neurofuzzy models, the ability of ANNs to provide a better understanding of the relationship between settlement and the factors affecting settlement is investigated. It was found that neurofuzzy networks have the ability to provide a transparent understanding of the relationship between settlement and the factors affecting it.

Settlement analysis is often affected by considerable levels of uncertainty that are usually ignored by traditional methods. In this research, ANNs are linked with Monte Carlo simulation to provide a stochastic solution for settlement prediction that takes into account the uncertainties associated with settlement analysis. A set of stochastic design charts that provide the designer with the level of risk associated with predicted settlements are developed and provided.

Publications

The following publications have been prepared as a result of this research:

Refereed Journal Papers

1. Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2002). "Predicting settlements of shallow foundations using neural networks." *Journal of Geotechnical & Geoenvironmental Engineering*, ASCE, 128(9), 785-793.
2. Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2002). "Artificial neural network–based settlement prediction formula for shallow foundations on granular soils." *Australian Geomechanics*, 37(4), 45-52.
3. Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2001). "Artificial neural network applications in geotechnical engineering." *Australian Geomechanics*, 36(1), 49-62.
4. Shahin, M. A., Jaksa, M. B., and Maier, H. R. "Neural network based stochastic design charts for settlement prediction." Submitted to the *Canadian Geotechnical Journal*.
5. Shahin, M. A., Maier, H. R., and Jaksa, M. B. "Data division for developing neural networks applied to geotechnical engineering." Submitted to the *Journal of Computing in Civil Engineering*, ASCE.
6. Shahin, M. A., Maier, H. R., and Jaksa, M. B. "Investigation into the robustness of artificial neural network models: A case study." Submitted to the *Journal of Neural Computing and Applications*.

Refereed Conference Papers

1. Shahin, M. A., Jaksa, M. B., and Holger, R. Maier (2003). "Neurofuzzy networks applied to settlement of shallow foundations on granular soils." ICASP9, *9th International Conference on Applications of Statistics and Probability in Civil Engineering*, University of California, Berkeley, (accepted).
2. Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2003). "Neural and neurofuzzy techniques applied to modelling settlement of shallow foundations on granular soils." MODSIM 2003, *International Congress on Modelling and Simulation*, Townsville, Queensland, July 14-17, (abstract accepted).

Departmental Reports

1. Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2000). "Predicting the settlement of shallow foundations on cohesionless soils using back-propagation neural networks." *Research Report No. R167*, The University of Adelaide, Adelaide.
2. Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2000). "Evolutionary data division methods for developing artificial neural network models for geotechnical engineering." *Research Report No. R171*, The University of Adelaide, Adelaide.

Statement of Originality

This work contains no material which has been accepted for the award of any other degree or diploma at any university or other tertiary institution and, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED:. DATE:...27/2/2003.....

Acknowledgments

In the name of Allah, the most gracious, the most merciful. I wish to express my deep gratitude and sincere appreciation to my supervisors, Dr. Mark Jaksa and Dr. Holger Maier, of the School of Civil and Environmental Engineering, the University of Adelaide, for their guidance, continual support and invaluable technical assistance throughout this research. In addition, their advice, encouragement and friendship have been of great value to me, and this work would not be possible without their contribution.

I wish to thank the School of Civil and Environmental Engineering, the University of Adelaide and in particular, the computer officers (Dr. Stephen Carr and Mr. Paul Carter) and office staff (Ms Diane Keable, Mrs. Josie Peluso and Mrs. Julia Ormsby) for their assistance. My special thanks are also extended to my fellow postgraduate students for their friendship and company.

The fund given to this research project by the University of Adelaide Scholarship (UAS) is sincerely acknowledged. My sincere acknowledgment is also extended to Dr. William Kaggwa for the proof reading of this thesis.

Finally, I am indebted to my family and in particular, my wife, Fifi, for her love, support, patience and strength during the period of this research, and my parents, Amin and Laila, for their encouragement and sacrifice.

List of Figures

List of Tables

Notation

A	fuzzy set or membership function
$a(i)$	average dissimilarity of data point i to all data points in cluster A
a_i	output from pth basis functions
$b(i)$	the smallest average dissimilarity of point i to all points in any cluster E different from A
B	footing width
c	rule confidence
C	objective function of fuzzy clustering
C_1	correction factor for embedment
C_2	correction factor for creep
$C_{y_j d_j}$	covariance between model output (y_j) and desired output (d_j)
$\bar{d}$	mean of desired output (d_j)
d_{ij}	given distance between data point i and j
d_j	desired (historical) actual output of node j
D_f	depth of footing embedment
D_j	Euclidean distance for node j
E	global error function
E_s	Young's modulus of soil
f_0	function bias
$f_{(1,2,\ldots,n)}$	function of order n
$f(I_j)$	transfer function of node j

$f'(I_j)$	derivative of transfer function of node j
F	settlement coefficient
I	number of input variables
I_j	activation level of node j
I_z	strain influence factor
k	settlement ratio (predicted settlement/measured settlement)
K	performance measure for B-spline neurofuzzy network
l	number of fuzzy clusters
L	footing length
n	number of nodes or data points
N	average SPT blow count
$N_{corrected}$	corrected SPT blow count
p	size of B-spline neurofuzzy model
$P_{N/E}$	probability of non-exceedance
q	footing net applied pressure
r	correlation coefficient
$s(i)$	silhouette value of data point i
$\bar{s}(l)$	average silhouette width of the entire data set
S_c	calculated settlement
S_m	average measured final settlement
S_p	predicted settlement
$u_{i\upsilon}$	unknown membership function of data point i to cluster υ;
$u_A(x)$	membership of x in fuzzy set A
w_i	connection weight associated with a_i for B-spline neurofuzzy network

w_{ji}	connection weight between nodes i and j
x	input variable
$\overline{x}$	average of sample data
x_i	input from node i
x_{max}	maximum value of input variable x
x_{min}	minimum value of input variable x
x_n	scaled value of input variable x
y	model output for B-spline neurofuzzy network
$\hat{y}$	desired output for B-spline neurofuzzy network
$\overline{y}$	average of model output y_j
y_j	predicted output of node j
α	shape parameter for log-logistic distribution
α_1	the first shape parameter for beta-general distribution
α_2	the second shape parameter for beta-general distribution
β	scale parameter for log-logistic and inverse-gamma distributions or decay constant for exponential distribution
γ	location parameter for log-logistic distribution
δ_j	error value between predicted and desired output for node j
Δw_{ji}	weight increment from node i to node j
Δ_{z_i}	thickness of ith soil layer
ε_z	vertical strain
η	learning rate
θ_j	threshold for node j

λ	shape parameter for inverse-gaussian distribution
μ	momentum term
σ_{d_j}	standard deviation of desired output y_j
σ_{y_j}	standard deviation of model output y_j

Abbreviations

AIC	Akaike's information criterion
ANNs	Artificial neural networks
ANOVA	The analysis of variance
ASMOD	The adaptive spline modelling of observation data
BIC	Bayesian information criterion
CDF	Cumulative distribution function
COV	Coefficient of variation
CPT	Cone penetration test
EN	Root mean squared percentage error
FEM	Finite element method
FPE	Final prediction error
LMS	Least mean squared
MAE	Mean absolute error
MLPs	Multi-layer perceptrons
MSE	Mean squared error
N/A	Not applicable
NLMS	Normalised least mean squared
NNNs	Natural neural networks
PDF	Probability density function
PE	Processing element
RMSE	Root mean squared error
SCANN	Site characterisation using artificial neural networks
SOMs	Self-organising maps
SPT	Standard penetration test

Chapter 1

Introduction

1.1 Introduction

The settlement of shallow foundations is usually divided into three components (Fang 1991): (a) immediate, or distortion, settlement, (b) consolidation settlement and (c) secondary compression settlement. Immediate settlement occurs with load application during, or immediately after, the erection of a structure. It is primarily a consequence of soil-grain distortion and reorientation. Consolidation settlement, on the other hand, is time-dependent and generally takes months to years to occur and is due to the dissipation of pore water pressure over time. Secondary compression settlement occurs as a result of soil creep, which is viscous flow under loading with no changes in effective stress. The total settlement of a foundation is the sum of the above three components. For cohesionless soils, only the immediate settlement is of concern, whereas consolidation and secondary compression settlements are the primary factors associated with cohesive soils.

It is generally understood that sand deposits are much more heterogeneous than the clay deposits and, as a result, differential settlements are likely to be higher in sand deposits than in clay profiles (Maugeri et al. 1998). Because cohesionless soils exhibit high degrees of permeability, settlement occurs in a short time; immediately after load application (Coduto 1994). Such quick settlement causes relatively rapid deformation of superstructures, which results in an inability to remedy damage and to avoid further deformation. Furthermore, excessive settlement occasionally leads to structural failure (Sowers 1970).

The two major criteria that control the design of shallow foundations are the bearing capacity of the footing and settlement of the foundations. However, settlement usually controls the design process, rather than bearing capacity, especially when the width of footing exceeds 1 metre (3-4 ft) (Schmertmann 1970). As a consequence, settlement prediction is a major concern and is an essential criterion in the design process of shallow foundations.

The prediction of settlement of shallow foundations on cohesionless soils is very complex and not yet entirely understood. This can be attributed to the fact that settlement is governed by many factors that are uncertain and difficult to quantify. Among these factors are the distribution of applied stress (Holzlohner 1984), the stress-strain properties of the soil, soil compressibility and the difficulty in obtaining undisturbed samples of cohesionless soils (Moorhouse 1972) for laboratory testing.

The geotechnical literature contains many methods, both theoretical and experimental, to predict settlement of shallow foundations on cohesionless soils. Due to the difficulty of obtaining undisturbed samples for cohesionless soils, many settlement prediction methods have focussed on correlations with in-situ tests, such as the standard penetration test (SPT), cone penetration test (CPT), dilatometer test, plate load test, pressuremeter test and screw plate load test. However, most of the available methods simplify the problem by incorporating several assumptions associated with the factors that affect settlement. Consequently, consistent and accurate prediction of settlement has yet to be achieved by the use of a variety of methods ranging from purely empirical to complex non-linear finite elements (Poulos 1999). Comparative studies of the available methods (e.g. Jorden 1977; Jeyapalan and Boehm 1986; Gifford et al. 1987; Tan and Duncan 1991; Wahls 1997) indicate inconsistent prediction of the magnitude of settlements. As a result, alternative methods are needed, which can overcome the limitations of the existing methods and provide more accurate settlement prediction.

The intention of this study is to apply an alternative approach, which is based on the data alone, in an attempt to obtain more accurate settlement prediction. The approach has been successfully applied to many problems including those of a geotechnical engineering nature and is known as artificial neural networks (ANNs). ANNs are a form of artificial intelligence, which, by means of their architecture, attempt to simulate the biological structure of the human brain and nervous system. ANNs have the ability to model the non-linear relationship between a set of input variables and the corresponding outputs without need for predefined mathematical equations. The ANN modelling philosophy is similar to most available methods for settlement prediction in the sense that both are attempting to capture the relationship between a set of model inputs and corresponding outputs. However, unlike most available methods, ANNs do not need prior knowledge about the nature of the relationship between the model inputs and corresponding outputs. ANNs use the data alone to determine the structure of the model as well as the unknown model parameters. This enables ANNs to overcome the limitations of the existing methods.

Settlement analysis, as in many geotechnical engineering problems, is affected by a considerable level of uncertainty associated with the factors that influence settlement. Most available methods for settlement prediction of shallow foundations on cohesionless soils disregard this uncertainty in their analysis and simulation. In order to provide practical design tools, ANNs will be linked with Monte Carlo simulation to incorporate the uncertainties associated with the factors that affect settlement prediction. Such a probabilistic approach is useful in the sense that it can overcome the limitations of the deterministic techniques and provide the geotechnical practitioner with some guidance about the level of risk (i.e. degree of uncertainty) that is associated with the predicted settlement.

1.2 Objectives and Scope of the Research

The overall objectives of this research are:

1. To explore the use of ANN models for predicting the settlement of shallow foundations on cohesionless soils and to compare their performance with some of the most commonly used traditional methods;

2. To introduce new data division methods for use in the development and verification of ANN models for settlement prediction;

3. To introduce a new validation approach for ANN models by carrying out a parametric study with a set of hypothetical data;

4. To provide a mathematical equation and produce a set of design charts for settlement prediction of shallow foundations on cohesionless soils based on the ANN technique;

5. To provide a better understanding of the relationships between the ANN model inputs and outputs for settlement prediction in the form of a set of fuzzy rules by applying the neurofuzzy technique;

6. To investigate the influence of including the uncertainty associated with the factors affecting settlement on the magnitude of settlement prediction and to produce a set of stochastic design charts for routine use in practice that provide the geotechnical practitioner with some guidance regarding the level of risk associated with predicted settlements; and

7. To assess the benefits and limitations of the proposed techniques as a practical tool for settlement prediction in comparison with more traditional methods.

Original contributions provided in this thesis and research are detailed in §8.3.

1.3 Layout of the Thesis

In the following chapter (Chapter 2), the structure and operation of ANNs are described. Categories for different classifications of ANNs are presented. Issues related to the development of ANN models are demonstrated and discussed. In Chapter 3, the major ANN applications in the field of geotechnical engineering are reviewed to illustrate the relative success or otherwise of ANNs in this field.

In Chapter 4, the main causes of settlement of shallow foundations are presented and the main factors that govern settlement prediction of shallow foundations on cohesionless soils are discussed. The different methods of settlement prediction of shallow foundations on cohesionless soils, are categorised and the more successful are highlighted.

In Chapter 5, the modelling methodology of multi-layer perceptrons that are trained with the back-propagation algorithm for predicting the settlement of shallow foundations on cohesionless soils is described and a method that tests the robustness of ANN models is introduced. Different data division methods for the development of ANN models are presented and a new approach for data division based on fuzzy clustering is introduced and evaluated. The effect of input data transformation on the performance of ANN models is examined. The relative importance of the factors affecting settlement is investigated. A comparison of the results obtained using ANNs and some of the most commonly used traditional methods is presented. A simple practical equation and a series of design charts for settlement prediction, based on ANNs, are developed.

In Chapter 6, the neurofuzzy technique is applied to assist with providing a better understanding of the relationships between settlements and the factors affecting them. The technique helps to produce a set of fuzzy rules that govern these relationships.

In Chapter 7, a practical stochastic approach is proposed for settlement prediction of shallow foundations on cohesionless soils that includes the uncertainty associated with the factors affecting settlement. The approach is based on linking the Monte Carlo technique with predicted settlements from the ANN model developed in Chapter 5. The results of the proposed stochastic approach are presented in the form of cumulative probability distribution design charts from which the probability that certain settlement predictions are exceeded can be readily obtained. The effect of varying the uncertainty associated with the factors affecting settlement on the magnitude of predicted settlement is examined. A set of stochastic design charts for settlement prediction are developed and provided for routine use in practice.

In the final chapter (Chapter 8), the research work is summarised and conclusions are presented. Recommendations for future work are also given.

Chapter 2

Artificial Neural Networks

2.1 Introduction

Artificial neural networks (ANNs) are a form of computing that attempt to simulate the operation of the human brain and nervous system. Although the concept of artificial neurons was first introduced in 1943 (McCulloch and Pitts 1943), research into applications of ANNs has blossomed since the introduction of the back-propagation training algorithm for feed-forward ANNs in 1986 (Rumelhart et al. 1986; McClelland and Rumelhart 1988). ANNs may thus be considered a relatively new tool in the field of prediction and forecasting. Recently, ANNs have been applied successfully to a wide range of areas including classification, estimation, prediction and functions synthesis (Moselhi et al. 1992). Moreover, ANNs have also been used successfully in predicting business failure, speech production and recognition, pattern recognition, medical diagnosis and treatment, control problems (Fausett 1994) and many fields of engineering, including geotechnical engineering, as will be examined in Chapter 3.

ANNs learn 'by example' in which an actual measured set of input variables and the corresponding outputs are presented to determine the rules that govern the relationship between the variables. Consequently, ANNs are well suited to model complex problems where the relationship between the variables is unknown (Hubick 1992) and when non-linearity is suspected (Maier 1995).

The aim of this chapter is to detail the more important features associated with ANNs and particularly those aspects addressed in the present research. The chapter begins with a brief description of natural neural networks and follows with an overview of the structure and operation of ANNs. The classification of different ANN types is presented and finally, the salient features of ANN model development are described and discussed.

2.2 Natural Neural Networks

The structure and operation of natural neural networks (NNNs) have been described by many authors (e.g. Hertz et al. 1991; Zurada 1992; Masters 1993; Fausett 1994). NNNs, of which the brain is an example, consist of billions of densely interconnected nerve cells called neurons. Each neuron receives the combined output signals (information) of many other neurons through synaptic gaps by input transmission paths called dendrites (Figure 2.1). The transmitted signals are electrochemical, which means that they are electronic impulses that transmit across the synaptic gaps to the dendrites by means of a chemical process (Fausett 1994). Consequently, the connection between neurons is chemical and the strength of this connection is modified by the action of the chemical transmitters and as the brain learns. The dendrites collect the incoming signals and send them to the cell body, or the soma, of the neuron. The soma sums the incoming signals and, if the charge of these signals is strong enough, the neuron is activated and produces an output signal; otherwise the neuron remains inactive. The output signal is then transmitted to the neighbouring neurons through an output structure called the axon. The axon of a neuron divides and connects to dendrites of the neighbouring neurons through junctions called synapses. The way neural networks receive, process and transmit the electrochemical signals, as well as the action of the chemical transmitters, comprise the basic memory mechanism and communication system of the human brain.

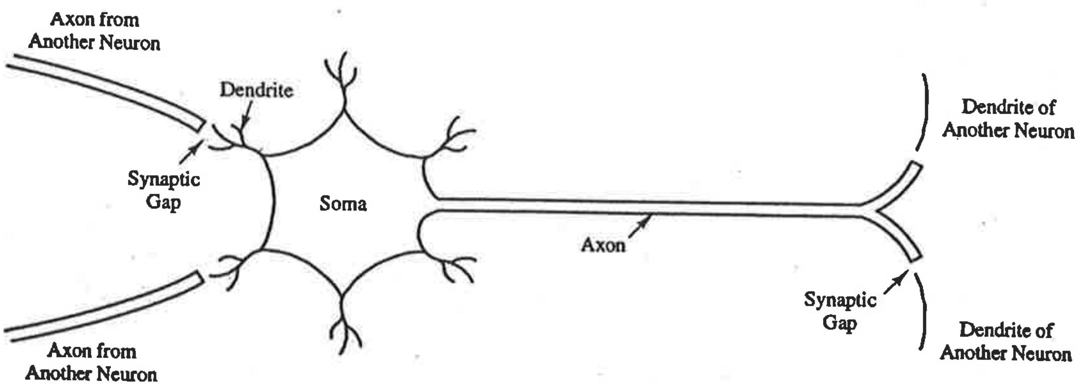


Figure 2.1: Typical structure of biological neuron (Fausett 1994)

2.3 Structure and Operation of Artificial Neural Networks

Artificial neural networks (ANNs) attempt to mimic some of the behaviour of the basic biological and chemical processes of NNNs. Many authors have described the structure and operation of ANNs (e.g. Hecht-Nielsen 1990; Maren et al. 1990; Zurada 1992; Fausett 1994; Ripley 1996). Briefly, ANNs consist of a number of artificial neurons variously known as 'processing elements' (PEs), 'nodes' or 'units', representing the neurons in NNNs. Processing elements in ANNs are usually arranged in layers: an input layer, an output layer and one or more intermediate layers called hidden layers (Figure 2.2).

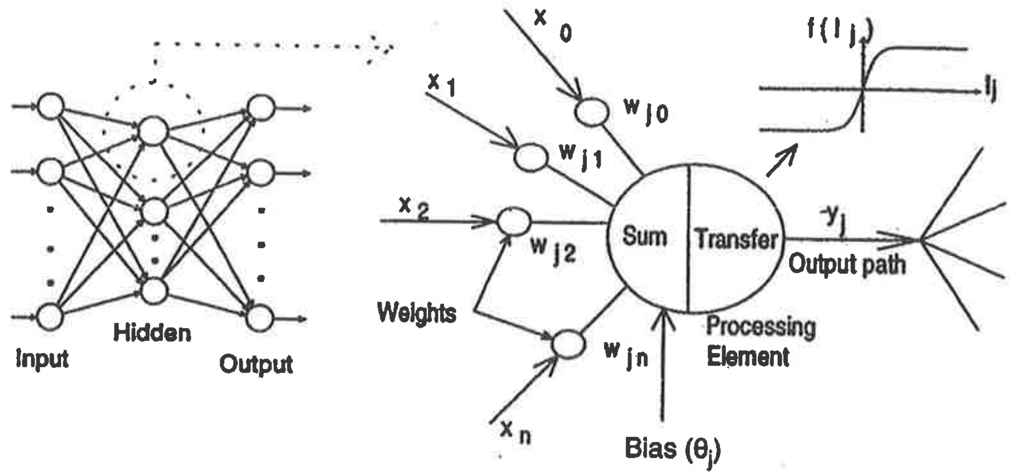


Figure 2.2: Typical structure and operation of ANNs (Maier and Dandy 1998)

Each processing element in a specific layer is fully or partially connected to many other processing elements via weighted connections. The weight in each connection represents the synaptic strength in NNNs. The scalar weights determine the strength of the connections between interconnected neurons. A zero weight refers to no connection between two neurons and a negative weight refers to a prohibitive relationship. From many other processing elements, an individual processing element receives its weighted inputs, which are summed and a bias unit or threshold is added or subtracted. The bias unit is used to scale the input to a useful range to improve the convergence properties of the neural network. The result of this combined summation is passed through a transfer function to produce the output of the processing element. For node j, this process is summarised in Equations 2.1 and 2.2 and illustrated in Figure 2.2.

$$I_j = \sum_{i=1}^{n} w_{ji} x_i + \theta_j \qquad \text{summation} \tag{2.1}$$

$$y_j = f(I_j) \qquad \text{transfer} \tag{2.2}$$

where:

I_j = the activation level of node *j*;

w_{ji} = the connection weight between nodes *i* and *j*;

x_i = the input from node *i*, $i = 0, 1, \ldots, n$;

θ_j = the threshold for node *j*;

y_j = the output of node *j*; and

$f(I_j)$ = the transfer function.

The propagation of information in ANNs starts at the input layer where the input data are presented. The inputs are weighted and received by each node in the next layer. The weighted inputs are then summed and passed through a transfer function to produce the nodal output, which is weighted and passed to processing elements in the next layer. The network adjusts its weights on presentation of a set of training data and uses a learning rule until it can find a set of weights that will produce the input-output mapping that has the smallest possible error. The above process is known as 'learning' or 'training'.

Transfer functions

Transfer functions can take a variety of forms. The logistic sigmoid and hyperbolic tangent transfer functions are the most common functions in neural networks (Fausett 1994). The logistic sigmoid function is usually used when the desired range of output values is between 0 and 1, whereas the hyperbolic tangent function is often used when the desired range of output values is between –1 and 1. The logistic sigmoid and hyperbolic tangent transfer functions are shown in Figures 2.3 and 2.4 and Equations 2.3 and 2.4, respectively. Usually, the same transfer function is used for all processing elements in a particular layer. The effect of using either the logistic sigmoid or hyperbolic tangent transfer functions in the hidden and output layers on the performance of ANN models will be investigated in Chapter 5.

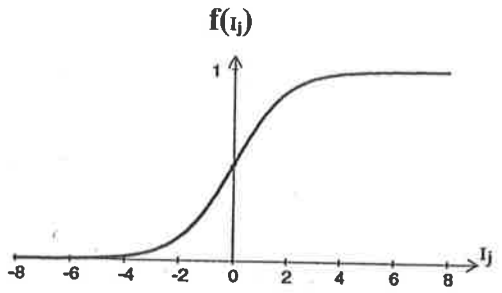


Figure 2.3: The logistic sigmoid function (Maier 1995)

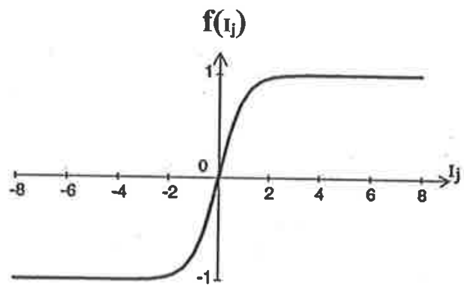


Figure 2.4: The hyperbolic tangent function (Maier 1995)

$$f(I_j) = \frac{1}{1+e^{-(I_j)}} \tag{2.3}$$

$$f(I_j) = \frac{e^{(I_j)} - e^{-(I_j)}}{e^{(I_j)} + e^{-(I_j)}} \tag{2.4}$$

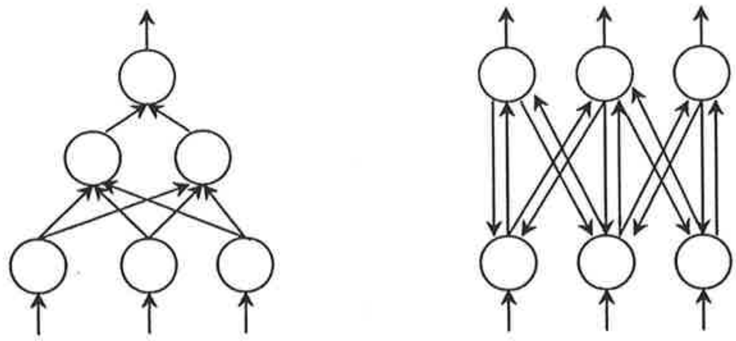
Learning (training)

Learning or training is the process of adjusting the weights in accordance with a learning rule and on the presentation of the training data. Learning in ANNs is usually divided into supervised and unsupervised learning (Masters 1993). In supervised learning, the network is presented with a historical set of model inputs and the corresponding (desired) outputs. The actual output of the network is compared with the desired output and an error is calculated. This error is used to adjust the connection weights between the model inputs and outputs to reduce the error between the historical outputs and those predicted by the ANN. The number of training samples presented between weight updates is called an epoch. The network may choose to be updated after: each training record is presented; the entire set of training data is presented or a certain number of training samples is presented.

In unsupervised learning, the network is only presented with the input stimuli and there are no desired outputs. The network itself adjusts the connection weights according to the input values. The idea of training in unsupervised networks is to cluster the input records into classes of similar features. Unsupervised learning is similar to the way learning takes place in the NNNs of the human brain.

2.4 Classification of Artificial Neural Networks

ANNs can be categorised on the basis of two major criteria: (i) the learning rule used and (ii) the connections between processing elements. Based on learning rules, ANNs, as mentioned above, can be divided into supervised and unsupervised networks. Two examples of supervised networks are multi-layer perceptrons and neurofuzzy networks. An example of an unsupervised network is the self-organising map. Based on connections between processing elements, ANNs can be divided into feed-forward and feedback networks. In feed-forward networks, the connections between processing elements are in the forward direction only (Figure 2.5a). In feedback networks, connections between processing elements are in both the forward and backward directions (Figure 2.5b).



(a) Feed-forward network **(b) Feedback network**

Figure 2.5: Connections between PEs for different neural network types

2.4.1 Multi-layer Perceptrons

Multi-layer perceptrons (MLPs) belong to the class of supervised feed-forward networks in which the processing elements are arranged in a multi-layered structure. The topology and algorithm details of MLPs are discussed in many publications (e.g. Hertz et al. 1991; Fausett 1994; Picton 1994; Ripley 1996). As mentioned previously, the structure of MLPs consists of an input layer, one or more hidden layers and an output layer. The input from each processing element (PE) in the previous layer is multiplied by a connection weight. These connection weights are adjustable and may be likened to the coefficients in statistical models. At each PE, the weighted input signals are summed and a bias or threshold value is added or subtracted. This combined input is then passed through a non-linear transfer function (e.g. logistic sigmoid or hyperbolic tangent transfer functions) to produce the output of the PE. The output of one PE provides the input to the PEs in the next layer. This process was summarised previously in Equations 2.1 and 2.2 and illustrated in Figure 2.2.

The global error between the output predicted by the network and the actual desired output is calculated using an error function. The mean squared error (MSE) function is usually preferable for the following reasons (Masters 1993): (a) the subsequent derivatives of this function are simple; (b) it gives more attention to large errors and (c) it lies close to the heart of the normal distribution in which, if the errors can be assumed to be normally distributed, minimising the MSE is optimal. Other measures can also be used and these are discussed in §2.5.7.

As mentioned earlier, the objective of the learning process is to minimise the errors between the predicted and actual outputs. This minimisation process can be achieved by the error function with respect to all variables in the neural network (e.g. connection weights, network architecture, learning rate and threshold). For simplicity and since the connection weights are the most influential variable, Rumelhart et al. (1986) proposed the back-propagation algorithm in which the error function is minimised with respect to the connection weights only. This error function is used in a backward manner to adjust the weights. The weights between the hidden layer and the output layer are adjusted first, followed by the weights between the hidden layer and the input layer. This process is repeated, which propagates the error term needed for weight adjustment until the network can obtain a set of weights, which have the input/output mapping that has the minimum error. Once the desired learning is achieved, the weights are fixed and the neural network can be deployed and used in practice.

The back-propagation training algorithm uses a gradient descent technique to adjust the weights. This process involves changing the weights from their initial random state by an amount proportional to the partial derivative of the error function, E, with respect to the given weight. For example, the error function, for node j, is calculated using the following equation:

$$E = \frac{1}{2}\sum (y_j - d_j)^2 \tag{2.5}$$

where:

E = the global error function;

y_j = the predicted output by the network; and

d_j = the desired (historical or measured) actual output.

The global error function, E, is minimised by modifying the weights using the gradient descent rule as follows:

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} \tag{2.6}$$

where:

Δw_{ji} = weight increment from node i to node j; and

η = learning rate, by which the size of the step taken along the error surface is determined.

Equation (2.6) can be further defined by the *delta rule* as follows:

$$\Delta w_{ji} = \eta \delta_j x_i \tag{2.7}$$

where:

x_i = input from node *i*, $i = 0, 1, \ldots, n$;

δ_j = error value between the predicted and desired output for node *j*.

If node *j* is in the output layer, δ_j can be calculated by applying the delta rule, as follows:

$$\delta_j = (y_j - d_j) f'(I_j) \tag{2.8}$$

where:

$f'(I_j)$ = the derivative of the activation function *f* with respect to the weighted sum of inputs of node *j*.

If node *j* is in the hidden layer, the generalised delta rule, proposed by Rumelhart et al. (1986), can be used as illustrated in Equation 2.9 and Figure 2.6.

$$\delta_j = \left[\sum_{1}^{m} \delta_m w_{mj} \right] f'(I_j) \tag{2.9}$$

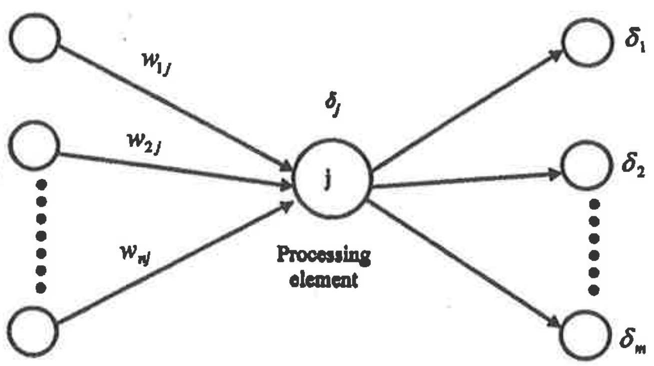


Figure 2.6: Node *j* in a hidden layer

The weights are then updated by adding the delta weight, Δw_{ji}, to the corresponding previous weight as follows:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} \tag{2.10}$$

where:

$w_{ji}(n)$ = the value of a weight from node i to node j at step n (before adjustment); and

$w_{ji}(n+1)$ = the value of the weight at step $(n+1)$ (after adjustment).

The back-propagation algorithm is sensitive to the initial conditions, i.e. the initial values of the weights, as a result of its gradient descent nature. For example, training may start with a set of initial weights that are positioned in a flat region of the error surface from which convergence becomes very slow (Hassoun 1995). Moreover, training may start from an unfavourable position in weight space from which the network may get stuck in a local minimum and cannot escape (Maier and Dandy 1998). The effect of using different initial random starting positions in weight space on the performance of ANN models will be investigated in Chapter 5.

There are two training modes for weights to be updated, on-line and batch modes. In on-line mode, the weights are updated after each training case is presented. In batch mode, the weights are not updated after each training case, rather, the weight change (Δw_{ji}) that is computed for each training case is accumulated to a certain epoch or until all training cases are presented. The average weight changes are then computed and used for weight updating. It has been suggested that on-line mode is better than batch mode as the sequence of training cases presented to the network can be easily randomised to avoid local minima (Zhang 1997). Consequently, on-line mode will be adopted for all ANN models that will be developed later in this research.

The choice of the learning rate is critical and the optimum learning rate is usually determined by trial-and-error. If the learning rate is selected to be small, convergence will be achieved, however, it will be very slow. In addition, convergence will be subject to the local minimum in the error surface that is closest to the random starting position. On the other hand, if the learning rate is selected to be large, convergence will never occur. Rumelhart et al. (1986) described a process to solve the above problem without leading to oscillation. This process is simply to add a momentum term (μ) to the weight adjustment that is proportional to the amount of the previous weight change. Once an adjustment is carried out, it is saved and used to modify all subsequent weight adjustments. This means that the weight change of the current step should carry some momentum of the weight change from the previous step. The modified adjustment equations are as follows:

$$\Delta w_{ji}(n+1) = -\eta \frac{\partial E}{\partial w_{ji}} + \mu \Delta w_{ji}(n) \tag{2.11}$$

and

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n+1) \tag{2.12}$$

A momentum value of 0.9 is customarily set, for both on-line and batch training modes, (Sarle 1994a). Ripley (1993) argued that it is often better to use momentum values of 0.99 or 0.999 for on-line training mode and a smaller value of 0.5 for batch training mode. However, Sarle (1994a) argued that the best momentum can be determined by trial-and-error. The effect of using several values of learning rates and momentum terms will be investigated in Chapter 5.

There are several other algorithms for training MLPs that are described by Hertz et al. (1991). Most of these algorithms are based on the assumption that the learning rate is constant from one epoch to the next and from one weight to another. However, some researchers (e.g. Chan and Fallside 1987; Jacobs 1988) challenged the above assumption by proposing learning rules that use varying learning rates and provided guidelines for learning rate update. This can decrease the number of cycles required for training, however, it has been argued that the automatic methods of updating learning rates have the risk of being trapped in local minima (Mukherjee and Deshpande 1997).

Despite the effectiveness of MLPs that are trained with the back-propagation algorithm for solving many engineering problems, they suffer from a number of shortcomings. MLPs trained with the back-propagation algorithm may be slow to converge (Wasserman 1989; Vitela and Reifman 1997). This is attributed to the fact that these networks rely on non-linear transfer functions for learning. If node activation is large, nodal outputs may tend to get stuck in the flat spots at the extreme values of the transfer functions as shown previously in Figures 2.3 and 2.4. The changes used to update the weights are a function of the derivative of the transfer functions. At the extreme values of the transfer functions the derivative is near zero. Consequently, very small weight changes can occur, resulting in a slow down in convergence. A number of ways are proposed in the literature to solve this problem. For example, Fahlman (1988) proposed adding a small, constant value to the derivative of the transfer function to prevent it from becoming zero. Fahlman (1988) achieved a dramatic improvement in training time by adding 0.1 to the derivative of the sigmoid transfer function. Another way to solve the above problem is the adjustment of the transfer function so that it never drops below a predefined level (Rojas 1996).

Another limitation of MLPs trained with the back-propagation algorithm is that when the network tries to find the global minimum of the error surface, it can get trapped in a local minimum. However, for many applications, local minima are not a significant problem, as they occur relatively infrequently (Weiss and Kulikowski 1991). Again, there are several ways proposed in the literature to escape local minima, including increasing the learning rate, adding a momentum term, adding a small amount of random noise to the input patterns to shake the network from the line of steepest descent, adding more hidden nodes and relocating the network along the error surface by randomising the initial weights and retraining (Sietsma and Dow 1988; Vitela and Reifman 1997; Maier and Dandy 2000). The effect of some of these approaches on the performance of ANN models will be investigated in Chapter 5.

Finally, feed-forward neural networks that are trained with the back-propagation algorithm are often criticised for being *black boxes*. The knowledge acquired by these networks during training is stored in their connection weights and bias values in a complex manner that is often difficult to interpret (Touretzky and Pomerleau 1989; Hegazy et al. 1994; Brown and Harris 1995; Shaopei and Boru 1998). Consequently, the rules governing the relationships between the input/output variables are difficult to quantify, especially for large networks that have a large number of PEs. As will be seen in the following section, one way to overcome this problem is to use neurofuzzy networks.

2.4.2 Neurofuzzy Networks

Neurofuzzy networks are modelling techniques that combine the explicit linguistic knowledge representation of fuzzy systems with the learning power of MLPs (Altrock 1995; Brown and Harris 1995). Neurofuzzy networks can be trained by processing data samples to perform input/output mappings, similar to the way MLPs do, with the additional benefit of being able to provide a set of production rules that describe the model input/output relationships and thus, they are more transparent (Brown and Harris 1994; Sayed and Razavi 2000). Neurofuzzy networks are new tools in the field of geotechnical engineering and, as will be seen in Chapter 6, they can be used to provide a better understanding of the relationships between ANN model inputs and outputs.

Neurofuzzy networks use the fuzzy logic system to store the knowledge acquired between a set of input variables (x_1, x_2, ..., x_n) and the corresponding output variable (y) in a set of linguistic fuzzy rules that can be easily interpreted, such as:

IF (x_1 *is high* AND x_2 *is low*) THEN (y *is high*), $c = 0.9$

where ($c = 0.9$) is the rule confidence which indicates the degree to which the above rule has contributed to the output. The concept of fuzzy logic was introduced by Zadeh (1965). As part of any fuzzy logic system, two main components (i.e. fuzzy sets and fuzzy rules) need to be determined. In order to determine the fuzzy sets, linguistic terms (e.g. small, medium and large) can be interpreted mathematically in the form of membership functions, and model variables are *fuzzified* to be partial members of these membership functions in the interval grade (0,1). This means that, for a fuzzy set A, an input variable x is fuzzified to be a partial member of the fuzzy set A by transforming it into a degree of membership of function $u_A(x)$ of interval (0,1). There are many forms of membership functions including B-spline and Gaussian functions (Brown and Harris 1994). Figure 2.7 is an example of B-spline basis functions of different order. For each variable, the fuzzy sets overlap and cover the necessary range of variation for that variable in a process called *fuzzification*. It should be noted that the model output of a fuzzy set is also fuzzy and, in order to obtain a real-valued output, defuzzification is needed. The *mean of maxima* and *centre of gravity* are the most popular defuzzification algorithms (Brown and Harris 1994).

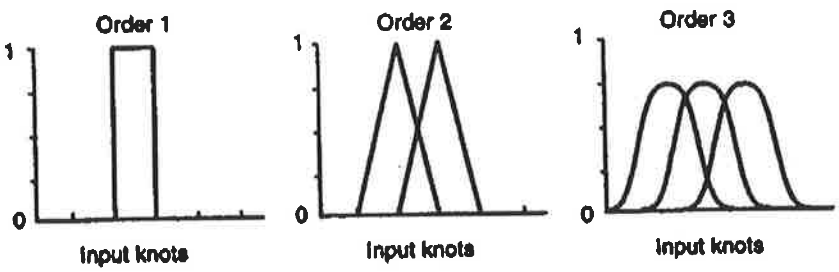


Figure 2.7: B-Spline fuzzy membership functions of different order

A typical structure of a neurofuzzy network contains three layers: an input layer; a single hidden layer and an output layer (Brown and Harris 1994). The input layer normalises the input space in a p-dimensional lattice (Figure 2.8). Each cell of the lattice represents similar regions of the input space. The hidden layer consists of basis functions (e.g. B-spline and Gaussian functions) which are defined on the lattice formed by normalising the input space. The size, shape and overlap of the basis functions determine the structure and complexity of the network. The output layer sums the weighted outputs from the basis functions to produce the network output using Equation (2.13).

$$y = \sum_{i=1}^{p} a_i w_i \qquad (2.13)$$

where:

y = model output;

a_i = output from the pth basis function; and

w_i = connection weight associated with a_i.

This output is compared with the actual measured output and a correction error (the mean squared error, MSE, is usually used) is calculated. Using this error and implementing a learning rule, the neurofuzzy network adjusts its weights and determines its fuzzy parameters (i.e. fuzzy sets and fuzzy rules). The Least Mean

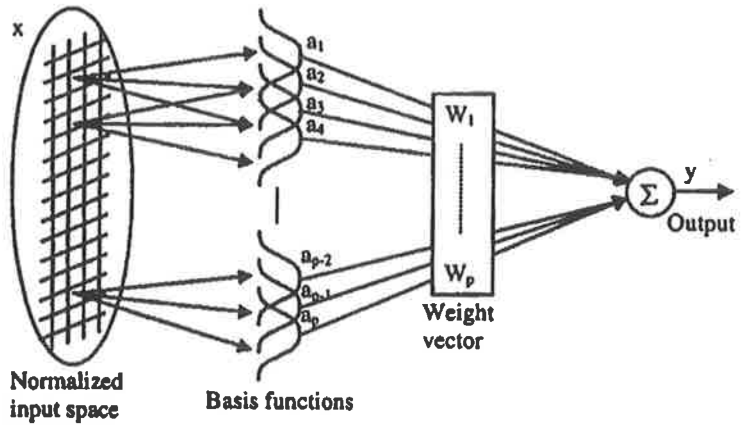


Figure 2.8: Typical structure of a neurofuzzy network (Brown and Harris 1995)

Squared (LMS) and the Normalised Least Mean Squared (NLMS) learning algorithms are generally used to update the weights (Brown and Harris 1994). At time t and as part of these algorithms, Equations 2.14 and 2.15, respectively, are used to adjust the weights for the LMS and NLMS algorithms (Brown and Harris 1995):

$$w_i(t) = w_i(t-1) + \eta(\hat{y}(t) - y(t))a_i(t) \tag{2.14}$$

$$w_i(t) = w_i(t-1) + \eta\left(\frac{\hat{y}(t) - y(t)}{\|a_i(t)\|_2^2}\right)a_i(t) \tag{2.15}$$

where:

η = learning rate; and

$\hat{y}$ = desired output.

It should be noted that, when the output error is zero, the weights are not updated, whereas if it is not zero, the weights are adjusted so as to reduce the output error. If the basis functions have nonzero outputs in only a small part of the input space, then only the numbers of weights that are contributing to the network output are updated during training. Consequently, similar network inputs result in similar sets of nonzero basis functions and therefore, the knowledge is stored locally in the network without interfering with the knowledge that is stored in other regions of the network.

One major disadvantage of B-spline networks is that the number of basis functions (i.e. fuzzy sets or membership functions) is exponentially dependent on the dimension of the input space (Brown and Harris 1995). Consequently, the number of rules is also exponentially dependent on the dimension of the input space, resulting in impractical model representation. This problem has been termed the *curse of dimensionality* (Brown and Harris 1995). To illustrate this problem, consider a fuzzy logic system that has five input variables and that each input variable is presented over five-valued membership functions. This fuzzy system will contain as many as ($5^5 = 3125$) rules. One useful approach for overcoming such problem is to use the analysis of variance (ANOVA) representation (Brown and Harris 1995). ANOVA decomposes an *n*-dimensional function to a linear combination of a number of separate functions, as follows (Brown and Harris 1995):

$$f(x) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f_{i,j}(x_i, x_j) + \dots + f_{1,2,\dots,n}(x) \tag{2.16}$$

where f_0 represents a constant (the function bias); and the other terms represent the univariate, bivariate and high-order subfunctions. In many situations, the majority of high-order terms are zero or negligible, resulting in a limited number of subfunctions (often called subnetworks) of much lower dimensions that approximate the network input/output mapping. It should be noted that each subnetwork in the ANOVA description represents a neurofuzzy system of its own and the overall model output is produced by summing outputs of all subnetworks. An example of ANOVA decomposition for the problem of five input variables and five membership functions for each of these is shown in Figure 2.9. The 5D function is decomposed into one 1D and two 2D subnetworks, resulting in 5, 25 and 25 fuzzy rules for the first, second and third subnetwork, respectively. Consequently, the network with ANOVA decomposition will produce an overall number of rules equal to 55 instead of 3125 for the non-decomposed network.

The adaptive spline modelling of observation data (ASMOD) proposed by Kavli (1993) is an automatic algorithm for obtaining the optimal structure of B-spline neurofuzzy networks. ASMOD has been found to perform well in a wide variety of modelling

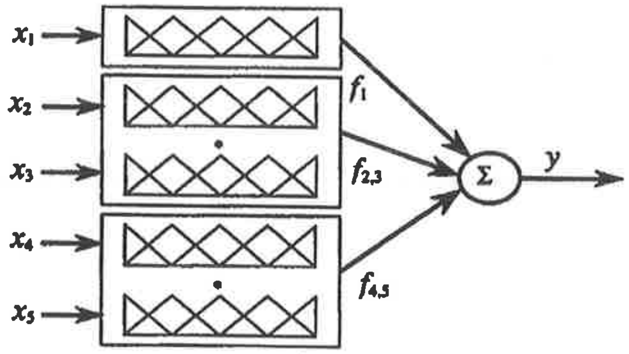


Figure 2.9: ANOVA decomposition of a neurofuzzy rule base (Brown and Harris 1995)

problems (Brown and Harris 1994). The algorithm starts with a simple model (e.g. only one variable with two membership functions) and iteratively refines the model structure during training so as to gradually increase model capability until some stopping criterion is met. Possible refinements include adding or deleting input variables, increasing the number and dimension of an individual subnetwork by linking it to an existing input, forming multi-variate subnetworks using ANOVA and changing the number and spacing of the basis functions (i.e. the optimum partitioning of the input space). Changing the order of B-spline functions for an individual input variable is also a possible refinement; however, the order of B-spline functions has to be determined in advance. It should be noted that higher order B-spline basis functions result in smoother model outputs; however, it is likely to lead to data overfitting (Brown and Harris 1994). Consequently, lower order basis functions are more desirable if they are able to model the desired relationship with a satisfied accuracy (Maier et al. 2001). For every refinement, the impact of network pruning is evaluated and the network that has the simplest structure with the best performance is chosen. As part of ASMOD, stopping criteria have to strike a balance between model performance and model size, training data and model error. Examples of such measures are given by Brown and Harris (1994), which include:

Bayesian Information Criterion (BIC): $K = n\ln(MSE) + p\ln(n)$ (2.17)

Akaike's Information Criterion (AIC): $K(\phi) = n\ln(MSE) + p\phi$, $\phi > 0$ (2.18)

$$\text{Final Prediction Error (FPE): } K = n\ln(MSE) + n\ln\left(\frac{n+p}{n-p}\right) \tag{2.19}$$

where K is the performance measure, p is the size of current model, MSE is the mean square error and n is the number of data pairs used to train the network. The effect of using the aforementioned stopping criteria on the performance of neurofuzzy networks will be investigated in Chapter 6.

2.4.3 Self-Organising Maps

Self-organising maps (SOMs) belong to the genre of unsupervised neural networks and were proposed and developed by Kohonen (1982). Unsupervised neural networks are usually used for data clustering to optimise and identify similarities associated with raw data. SOMs will be used in Chapter 5 to cluster the data variables in order to divide the data into their training, testing and validation subsets. The typical structure of SOMs consists of two layers: an input layer and a Kohonen layer (Figure 2.10).

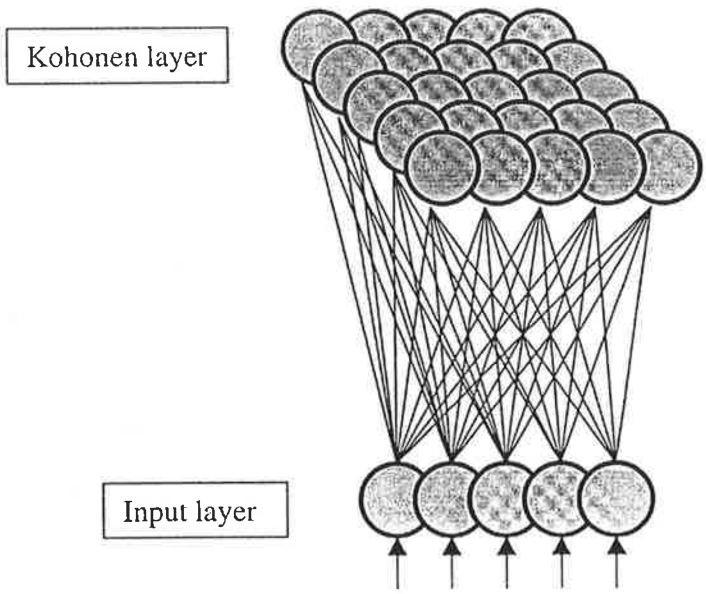


Figure 2.10: Typical structure of self-organising map

The Kohonen layer has a number of competitive processing elements (PEs) or nodes arranged in a one- or two-dimensional array. The input from each node in the input layer (x_i for $i = 1, 2, \ldots, n$) is fully connected to the Kohonen layer through connection weights (w_{ji} for $j = 1, 2, \ldots, m$). At the beginning of the self-organising process, these weights are randomly initialised. At each node in the Kohonen layer, the input (x_i) is presented without providing the desired output, and a matching value is calculated. This value is typically the Euclidean distance. For node j in the Kohonen layer, the Euclidean distance (D_j) between the weights and the corresponding input values, is given by Equation 2.20.

$$D_j = \sum_{i=1}^{n} (w_{ji} - x_i)^2 \, , j = 1, 2, \ldots, m \qquad (2.20)$$

The node that has the minimum Euclidean value is declared the winner. That is, the winner is the node whose weights are most similar to the input values. The weights of the winning node and its neighbouring nodes, in terms of topology, are then updated to match the input values more closely. The incremental weight update for node j is as follows:

$$\Delta w_{ji} = \eta (x_i - w_{ji}) \qquad (2.21)$$

where:

η = learning rate.

At step n of the training, node j can be updated as in Equation 2.10.

The process is repeated by successively presenting new input data records to the model and adjusting the connection weights until they remain unchanged. The result is a topological map in which similar data records are clustered together. A full description of the operation of self-organising maps is given by Kohonen (1997).

2.5 Development of Artificial Neural Network Models

In order to improve performance, ANN models need to be developed in a systematic manner. Such an approach needs to address major factors such as the determination of adequate model inputs, data division and pre-processing, the choice of a suitable network architecture, careful selection of some internal parameters that control the optimisation method, the stopping criteria and model validation (Maier and Dandy 2000). These factors are explained and discussed below.

2.5.1 Determination of Model Inputs

An important step in developing ANN models is to select the model input variables that have the most significant impact on model performance. A good subset of input variables can substantially improve model performance. Presenting as large a number of input variables as possible to ANN models usually increases network size, resulting in a decrease in processing speed and a reduction in the efficiency of the network (Lachtermacher and Fuller 1994). A number of techniques have been suggested in the literature to assist with the selection of input variables. An approach that is usually utilised in the field of geotechnical engineering is that a fixed number of input variables can be used in advance and assumed to be the most effective input variables in relation to the model output variables. This approach will be adopted for the ANN models that are developed in this research (see Chapter 5). Another approach used by some researchers (e.g. Goh 1994b; Najjar et al. 1996; Ural and Saka 1998) is to train many neural networks with different combinations of input variables and to select the network that has the best performance. A step-wise technique described by Maier and Dandy (2000) can also be used in which separate networks are trained, each using only one of the available variables as model inputs. The network that performs the best is then retained, combining the variable that resulted in the best performance with each of the remaining variables. This process is repeated for an increasing number of input variables, until the addition of any extra variable results in no improvement in model performance. Another useful approach is to employ a genetic algorithm to search for the best sets of input variables (NeuralWare 1997). For each possible set of input variables chosen by the genetic algorithm, a neural network is trained and used to rank different subsets of possible inputs. A set of input variables derives its fitness from the model error obtained based on those variables. The adaptive spline modelling of observation data (ASMOD) algorithm proposed by Kavli (1993) is also a useful technique that can be used for developing parsimonious neurofuzzy networks by automatically selecting a combinations of model input variables that have the most significant impact on the outputs. The ASMOD algorithm will be adopted in Chapter 6 for the neurofuzzy models that are developed in this research.

2.5.2 Division of Data

ANNs are similar to conventional statistical models in the sense that model parameters (e.g. connection weights) are adjusted in the model calibration phase (training) so as to minimise the error between model outputs and the corresponding measured values for a particular data set (the training set). ANNs perform best when they do not extrapolate beyond the range of the data used for calibration (Flood and Kartam 1994; Minns and Hall 1996; Tokar and Johnson 1999). Therefore, the purpose of ANNs is to non-linearly interpolate (generalise) in high-dimensional space between the data used for calibration. Unlike conventional statistical models, ANN models generally have a large number of model parameters (connection weights) and can therefore overfit the training data, especially if the training data are noisy. In other words, if the number of degrees of freedom of the model is large compared with the number of data points used for calibration, the model might no longer fit the general trend, as desired, but might learn the idiosyncrasies of the particular data points used for calibration leading to '*memorisation*', rather than '*generalisation*'. Consequently, a separate validation set is needed to ensure that the model can generalise within the range of the data used for calibration. It is common practice to divide the available data into two subsets; a training set, to construct the neural network model, and an independent validation set to estimate the model performance in a deployed environment (Twomey and Smith 1997; Maier and Dandy 2000). Usually, two-thirds of the data are suggested for model training (i.e. training and testing sets) and one-third for validation (Hammerstrom 1993). A modification of the above data division method is cross-validation (Stone 1974) in which the data are be divided into three sets: training, testing and validation. The training set is used to adjust the connection weights, whereas the testing set is used to check the performance of the model at various stages of training and to determine when to stop training to avoid over-fitting. The validation set is used to estimate the performance of the trained network in the deployed environment. There are no guidelines in the literature for the optimal proportion of the data to use for training, testing and validation sets. In an attempt to determine an optimal proportion of the data, the relationship between the proportion of the data used in each subset and ANN model performance will be investigated in Chapter 5.

In many situations, the available data are small enough to be solely devoted to model training and collecting any more data for validation is difficult. In this situation, the *leave-k-out* method can be used (Masters 1993) which involves holding back a small fraction of the data for validation and the rest of the data for training. After training, the performance of the trained network has to be estimated with the aid of the validation set. A different small subset of data is held back and the network is trained and tested again. This process is repeated many times with different subsets until an optimal model can be obtained from the use of all of the available data.

In the majority of ANN applications in geotechnical engineering, the data are divided into their subsets on an arbitrary basis. However, recent studies have found that the way the data are divided can have a significant impact on the results obtained (e.g. Tokar and Johnson 1999). As ANNs have difficulty extrapolating beyond the range of the data used for calibration, in order to develop the best ANN model, given the available data, all of the patterns that are contained in the data need to be included in the calibration set. For example, if the available data contain extreme data points that were excluded from the calibration data set, the model cannot be expected to perform well, as the validation data will test the model's extrapolation ability, and not its interpolation ability. If all of the patterns that are contained in the available data are contained in the calibration set, the toughest evaluation of the generalisation ability of the model is if all the patterns (and not just a subset) are contained in the validation data. In addition, if cross-validation is used as the stopping criterion (see §2.5.6), the results obtained using the testing set have to be representative of those obtained using the training set, as the testing set is used to decide when to stop training or for example which model architecture or learning rate is optimal. Consequently, the statistical properties (e.g. mean and standard deviation) of the various data subsets (e.g. training, testing and validation) need to be similar to ensure that each subset represents the same statistical population (Masters 1993). If this is not the case, it may be difficult to judge the validity of ANN models (Maier and Dandy 2000).

This fact has been recognised for some time (Masters 1993; ASCE 2000; Maier and Dandy 2000), and several studies have used ad-hoc methods to ensure that the data used for calibration and validation have the same statistical properties (Braddock et al. 1998; Campolo et al. 1999; Tokar and Johnson 1999; Ray and Klindworth 2000). Masters (1993) strongly confirms the above strategy of data division as he says "*if our training set is not representative of the data on which the network will be tested, we will be wasting our time*". However, it was not until recently that systematic approaches for data division have been proposed in the literature. Bowden et al. (2002) used a genetic algorithm to minimise the difference between the means and standard deviations of the data in the training, testing, and validation sets. While this approach ensures that the statistical properties of the various data subsets are similar, there is still a need to choose which proportion of the data to use for training, testing, and validation. Kocjancic and Zupan (2000) and Bowden et al. (2002) used a self-organising map (SOM) to cluster high-dimensional input and output data in two-dimensional space and divided the available data so that values from each cluster are represented in the various data subsets. This ensures that data in the different subsets are representative of each other and has the additional advantage that there is no need to decide what percentage of the data to use for training, testing and validation. The major shortcoming of this approach is that there are no guidelines for determining the optimum size and shape of the SOM (Cai et al. 1994; Giraudel and Lek 2001). This has the potential to have a significant impact on the results obtained, as the underlying assumption of the approach is that the data points in one cluster provide the same information in high-dimensional space. However, if the SOM is too small, there may be significant intra-cluster variation. Conversely, if the map is too large, too many clusters may contain single data points, making it difficult to choose representative subsets.

In this research, a new data division approach is introduced and compared with existing approaches. The new approach utilises a fuzzy clustering technique, which overcomes the limitations of existing methods. Shi (2002) has recently used fuzzy clustering for the evaluation and validation of neural networks. However, thus far, fuzzy clustering has not yet been used as a data division approach for ANNs.

2.5.3 Data Pre-processing

Once the available data have been divided into their subsets (i.e. training, testing and validation), it is important to pre-process the data in a suitable form before they are applied to the ANN. Data pre-processing is necessary to ensure all variables receive equal attention during the training process. Moreover, pre-processing usually speeds up the learning process. Pre-processing can be in the form of data scaling, normalisation and transformation (Masters 1993). Scaling the output data is essential, as they have to be commensurate with the limits of the transfer functions used in the output layer (e.g. between –1.0 to 1.0 for the tanh transfer function and 0.0 to 1.0 for the sigmoid transfer function). Scaling the input data is not necessary but it is almost always recommended (Masters 1993). In some cases, the input data need to be normally distributed in order to obtain optimal results (Fortin et al. 1997). However, Burke and Ignizio (1992) stated that the probability distribution of the input data does not have to be known. Transforming the input data into some known forms (e.g. linear, log, exponential, etc.) may be helpful to improve ANN performance. Shi (2000) showed that distribution transformation of the input data to a uniform distribution improves network performance by 50%. However, empirical trials (Faraway and Chatfield 1998) showed that the model fits were the same, regardless of whether raw or transformed data were used. The distribution transformation method proposed by Shi (2000) will be examined in Chapter 5 in an attempt to improve the performance of ANN models.

2.5.4 Determination of Model Architecture

Determining the network architecture is one of the most important and difficult tasks in ANN model development (Maier and Dandy 2000). It requires the selection of the optimum number of layers and the number of nodes in each of these. There is no unified theory for determination of an optimal ANN architecture. It is generally achieved by fixing the number of layers and choosing the number of nodes in each layer. There are always two layers representing the input and output variables in any neural network. It has been shown that one hidden layer is sufficient to approximate any continuous function provided that sufficient connection weights are given (Cybenko 1989; Hornik et al. 1989). Hecht-Nielsen (1989) provided a proof that a single hidden layer of neurons, operating a sigmoidal activation function, is sufficient to model any solution surface of practical interest. To the contrary, Flood (1991) stated that there are many solution surfaces that are extremely difficult to model using a sigmoidal network using one hidden layer. In addition, some researchers (Flood and Kartam 1994; Sarle 1994b; Ripley 1996) stated that, the use of more than one hidden layer provides the flexibility needed to model complex functions in many situations. Lapedes and Farber (1988) provided more practical proof that two hidden layers are sufficient, and according to Chester (1990), the first hidden layer is used to extract the local features of the input patterns while the second hidden layer is useful to extract the global features of the training patterns. However, Masters (1993) stated that using more than one hidden layer often slows the training process dramatically and increases the chance of getting trapped in local minima.

The number of nodes in the input and output layers are restricted by the number of model inputs and outputs, respectively. There is no direct and precise way of determining the best number of nodes in each hidden layer. A trial-and-error procedure, which is generally used in geotechnical engineering to determine the number and connectivity of the hidden layer nodes, can be used. It has been shown in the literature (e.g. Maren et al. 1990; Masters 1993; Rojas 1996) that neural networks with a large number of free parameters (connection weights) are more subject to overfitting and poor generalisation. Consequently, keeping the number of hidden nodes to a minimum, provided that satisfactory performance is achieved, is always better, as it: (a) reduces the computational time needed for training; (b) helps the network to achieve better generalisation performance; (c) avoids the problem of overfitting and (d) allows the trained network to be analysed more easily. For single hidden layer networks, there are a number of rules-of-thumb to obtain the best number of hidden layer nodes. One approach is to assume the number of hidden nodes to be 75% of the number of input units (Salchenberger et al. 1992). Another approach suggests that the number of hidden nodes should be between the average and the sum of the nodes in the input and output layers (Berke and Hajela 1991). A third approach is to fix an upper bound and work back from this bound. Hecht-Nielsen (1987) and Caudill (1988) suggested that upper limit of the number of hidden nodes in a single layer network may be taken as $(2I+1)$, where I is the number of inputs. The best approach found by Nawari et al. (1999) is to start with a small number of nodes and to slightly increase the number until no significant improvement in model performance is achieved. Yu (1992) showed that the error surface of a network with one hidden layer and $(I-1)$ hidden nodes has no local minima. For networks with two hidden layers, the *geometric pyramid rule* described by Nawari et al. (1999) can be used. The notion behind this method is that the number of nodes in each layer follows a geometric progression of a pyramid shape, in which the number of nodes decreases from the input layer towards the output layer. Kudrycki (1988) found empirically that the optimum ratio of the first to second hidden layer nodes is 3:1, even for high dimensional inputs.

Another way of determining the optimal number of hidden nodes that can result in good model generalisation and avoid overfitting, is to relate the number of hidden nodes to the number of available training samples. Masters (1993) stated "*the only way to prevent the network from learning unique characteristics of the training set, to the detriment of learning universal characteristics, is to flood it with so many examples that it cannot possibly learn all of their idiosyncracies*". There are a number of rules-of-thumb that have been suggested in the literature to relate the training samples to the number of connection weights. For instance, Rogers and Dowla (1994) suggested that the number of weights should not exceed the number of training samples. Masters (1993) stated that the required minimum ratio of the number of training samples to the number of connection weights should be 2 and, the minimum ratio of the optimum training sample size to the number of connection weights should be 4. Hush and Horne (1993) suggested that this ratio should be 10. Amari et al. (1997) demonstrated that if this ratio is at least 30, overfitting does not occur.

More recently, a number of systematic approaches have been proposed to automatically obtain the optimal network architecture. The *adaptive method of architecture determination*, suggested by Ghaboussi and Sidarta (1998), is an example of the automatic methods for obtaining the optimal network architecture that suggests starting with an arbitrary, but small, number of nodes in the hidden layers. During training, and as the network approaches its capacity, new nodes are added to the hidden layers, and new connection weights are generated. Training is continued immediately after the new hidden nodes are added to allow the new connection weights to acquire the portion of the knowledge base which was not stored in the old connection weights. For this process to be achieved, some training is carried out with the new modified connection weights only, while the old connection weights are frozen. Additional cycles of training are then carried out where all the connection weights are allowed to change. The above steps are repeated and new hidden nodes are added as needed to the end of the training process, in which the appropriate network architecture is automatically determined. *Pruning* is another automatic approach to determine the optimal number of hidden nodes. One such technique proposed by Karnin (1990) starts training a network that is relatively large and later reduces the size of the network by removing the unnecessary hidden nodes. *Genetic algorithms* provide evolutionary alternatives to obtain an optimal neural network architecture that have been used successfully in many situations (Miller et al. 1989). The *adaptive spline modelling of observation data* (ASMOD) (Kavli 1993) algorithm, is an automatic method for obtaining the optimal architecture of B-spline neurofuzzy networks, as shown in §2.4.2.

Cascade-Correlation (Fahlman and Lebiere 1990) is another automatic method to obtain the optimal architecture of ANNs. Cascade-Correlation is a constructive method that can be characterised by the following steps (Fahlman and Lebiere 1990). The neural network is initially trained using Fahlman's *quickprop* (Fahlman 1988) algorithm without hidden nodes and with direct connection between the input layer and the output layer. Hidden nodes are added randomly one or a few at a time. New hidden nodes receive connections from all previously established hidden nodes as well as from the original inputs. At the time new hidden nodes are added to the network, their connections with the inputs are frozen and only their output connections are trained using the *quickprop* algorithm. This process is stopped when the model performance shows no further improvement. Consequently, the architecture of ANN networks using Cascade-Correlation is that the input nodes are connected to the output nodes and the hidden nodes are connected to the input and output nodes as well as other previously established hidden nodes, as shown in Figure 2.11.

The constructive nature of the Cascade-Correlation method means that the way in which the hidden nodes are connected results in the addition of a new single-node layer to the

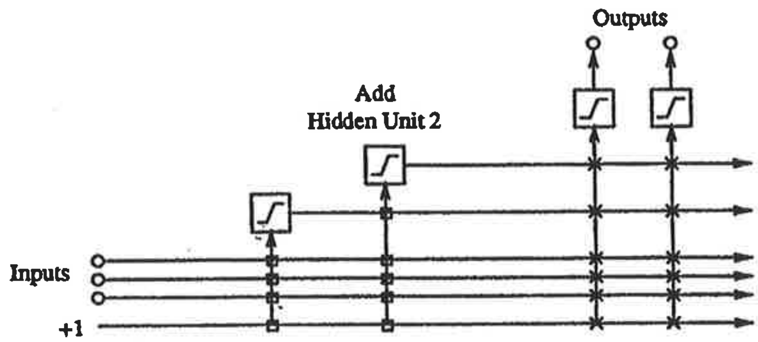


Figure 2.11: ANN architecture with the Cascade-Correlation (Fahlman and Lebiere 1990)

network each time a new node is added. This is designed to result in the smallest network that can adequately map the design input-output relationship, which has a number of advantages, including improved generalisation ability (Castellano et al. 1997) and higher processing speed (Bebis and Georgiopoulos 1994).

It should be noted that Masters (1993) has argued that the automatic approaches for obtaining optimal network architectures can be easily abused, as they do not directly address the problem of overfitting. In an attempt to exploit the benefits of both the automatic and manual approaches, the Cascade-Correlation and ad-hoc trial-and-error methods will be examined for the development of ANN models in Chapter 5. On the other hand, the ASMOD algorithm will be used for the development of the neurofuzzy networks in Chapter 6.

2.5.5 Model Optimisation (Training)

The process of optimising the connection weights is known as 'training' or 'learning'. This is equivalent to the parameter estimation phase in conventional statistical models. The aim is to find a global solution to what is typically a highly non-linear optimisation problem (White 1989). The method most commonly used for finding the optimum weight combination of feed-forward neural networks is the back-propagation algorithm (Rumelhart et al. 1986) which is based on first-order gradient descent. The use of global optimisation methods, such as simulated annealing and genetic algorithms, have also been proposed (Hassoun 1995). The advantage of these methods is that they have the ability to escape local minima in the error surface and, thus, produce optimal or near optimal solutions. However, they also have a slow convergence rate. Ultimately, the model performance criteria, which are problem specific, will dictate which training algorithm is most appropriate. If training speed is not a major concern, there is no reason why the back-propagation algorithm cannot be used successfully (Breiman 1994). Consequently, the back-propagation algorithm will be used for optimising the connection weights of the MLP models developed in Chapter 5. On the other hand, as mentioned in §2.4.2, the weights of B-spline neurofuzzy networks are generally updated using the *Least Mean Squared* or *Normalised Least Mean Squared* learning rules (Brown and Harris 1994), which will be used for the development of the neurofuzzy models in Chapter 6.

2.5.6 Stopping Criteria

Stopping criteria are used to decide when to stop the training process. They determine whether the model has been optimally or sub-optimally trained. Many approaches can be used to determine when to stop training. Training can be stopped: after the presentation of a fixed number of training records; when the training error reaches a sufficiently small value; or when no or slight changes in the training error occur. However, the above examples of stopping criteria may lead to the model stopping prematurely or over-training. As mentioned previously, the *cross-validation* technique (Stone 1974) is an approach that can be used to overcome such problems. It is considered to be the most valuable tool to ensure over-fitting does not occur (Smith 1993). Amari et al. (1997) suggested that there are clear benefits in using cross-validation when limited data are available, as is the case for many real-life case studies. The benefits of cross-validation are discussed further in Hassoun (1995). As mentioned in §2.5.2, the cross-validation technique requires the data be divided into three sets; training, testing and validation. The training set is used to adjust the connection weights. The testing set measures the ability of the model to generalise, and the performance of the model using this set is checked at many stages of the training process and training is stopped when the error of the testing set starts to increase. The testing set is also used to determine the optimum number of hidden layer nodes and the optimum values of the internal parameters (learning rate, momentum term and initial weights). The validation set is used to assess the model performance once training has been accomplished. Model validation is discussed in more detail in the following section. Cross-validation will be used for the development of all MLP models in this research, as will be seen in Chapter 5. On the other hand, as mentioned in §2.4.2, B-spline neurofuzzy networks use a number of different stopping criteria (e.g. Bayesian Information Criterion, Akaike's Information Criterion and Final Prediction Error). Unlike cross-validation, these stopping criteria require the data be divided into only two sets; a training set, to construct the model; and an independent validation set, to test the validity of the model in the deployed environment. The basic notion of these stopping criteria is that the model performance should balance the model complexity with the amount of training data and model error. The above stopping criteria will be investigated for the development of the neurofuzzy models in Chapter 6.

2.5.7 Model Validation

Once the training phase of the model has been successfully accomplished, the performance of the trained model should be validated. The purpose of the model validation phase is to ensure that the model has the ability to generalise within the limits set by the training data in a robust fashion, rather than simply having memorised the input-output relationships that are contained in the training data. The approach that is generally adopted in the literature to achieve this is to test the performance of trained ANNs on an independent validation set, which has not been used as part of the model building process. If such performance is adequate, the model is deemed to be able to generalise and is considered to be robust. In this thesis, an additional approach to test the generalisation ability and robustness of ANN models will be proposed. The proposed approach is suggested to complement the approach that is usually used in the literature and will be presented in Chapter 5.

The coefficient of correlation, r, the root mean squared error, RMSE, and the mean absolute error, MAE, are the main criteria that are often used to evaluate the prediction performance of ANN models. The coefficient of correlation is a measure that is used to determine the relative correlation and the goodness-of-fit between the predicted and observed data and can be calculated as follows:

$$r = \frac{C_{y_j d_j}}{\sigma_{y_j} \sigma_{d_j}} \tag{2.22}$$

and

$$C_{y_j d_j} = \frac{1}{n-1}\sum_{j=1}^{n}(y_j - \bar{y})(d_j - \bar{d}) = \frac{1}{n-1}\left(\sum_{j=1}^{n} y_j d_j - \frac{\sum_{j=1}^{n} y_j \sum_{j=1}^{n} d_j}{n}\right) \tag{2.23}$$

$$\sigma_{y_j} = \sqrt{\frac{\sum_{j=1}^{n}(y_j - \bar{y})^2}{n-1}} \tag{2.24}$$

$$\sigma_{d_j} = \sqrt{\frac{\sum_{j=1}^{n}(d_j - \bar{d})^2}{n-1}} \tag{2.25}$$

$$\bar{y} = \frac{\sum_{j=1}^{n} y_j}{n} \tag{2.26}$$

$$\bar{d} = \frac{\sum_{j=1}^{n} d_j}{n} \tag{2.27}$$

where:

y_j = model (predicted) output, $y_j = y_1, y_2, y_3, \ldots, y_n$;

d_j = desired (observed) output, $d_j = d_1, d_2, d_3, \ldots, d_n$;

$C_{y_j d_j}$ = covariance between the model output (y_j) and the desired output (d_j);

σ_{y_j} = standard deviation of the model output y_j;

σ_{d_j} = standard deviation of the desired output d_j;

$\bar{y}$ = mean of the model output y_j;

$\bar{d}$ = mean of the desired output d_j; and

n = number of data.

Smith (1986) suggested the following guide for values of $|r|$ between 0.0 and 1.0:

- $|r| \geq 0.8$ strong correlation exists between two sets of variables;
- $0.2 < |r| < 0.8$ correlation exists between the two sets of variables; and
- $|r| \leq 0.2$ weak correlation exists between the two sets of variables.

The RMSE is the most popular measure of error and has the advantage that large errors receive much greater attention than small errors (Hecht-Nielsen 1990). RMSE is calculated as follows:

$$\text{RMSE} = \left\{ \frac{1}{n} \sum_{j=1}^{n} (y_j - d_j)^2 \right\}^{\frac{1}{2}} \tag{2.28}$$

In contrast with RMSE, MAE eliminates the emphasis given to large errors. Both RMSE and MAE are desirable when the evaluated output data are smooth or continuous (Twomey and Smith 1997) and is calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^{n} \left| y_j - d_j \right| \tag{2.29}$$

2.6 Summary

It has been demonstrated that ANNs are a form of artificial intelligence, which, by means of their architecture, attempt to simulate the biological structure of the human brain and nervous system. It is evident from this chapter that the ANN modelling philosophy for prediction and forecasting is similar to that used in more conventional statistical models. In both cases, the purpose of the model is to capture the relationship between a set of model inputs and the corresponding outputs. To achieve this, it has been shown that ANNs rely on the data alone to determine the structure and parameters of the model. It has also been shown that the development of ANN models must address several factors which include the determination of adequate model inputs, data division and pre-processing, the choice of a suitable network architecture, careful selection of some internal parameters, the stopping criteria and model validation. The relative success of ANNs in the field of geotechnical engineering will be examined in the following chapter.

Chapter 3

Artificial Neural Network Applications in Geotechnical Engineering

3.1 Introduction

The engineering properties of soil and rock exhibit uncertain behaviour from one location to another due to the complex and varied physical processes associated with the formation of these materials (Jaksa 1995). This is in contrast to most other civil engineering materials, such as steel, concrete and timber, which exhibit far greater homogeneity and isotropy. In order to cope with the complexity of geotechnical behaviour and the spatial variability of these materials, traditional forms of engineering design models are justifiably simplified. An alternative approach, which has shown some promise in the field of geotechnical engineering, is artificial neural networks (ANNs).

Over the last few years, the use of ANNs has increased in many areas of engineering. In particular, ANNs have been applied to many geotechnical engineering problems and have demonstrated some degree of success. A review of the literature reveals that ANNs have been used successfully in pile capacity prediction, predicting the settlement of structures, modelling soil properties and behaviour, determination of liquefaction potential, site characterisation, modelling earth retaining structures, evaluating stability of slopes and the design of tunnels and underground openings. In the majority of these applications, multi-layer perceptrons (MLPs) trained with the back-propagation algorithm were used. The aim of this chapter is to provide an overview of most ANN applications that have appeared to-date in geotechnical engineering to reveal the relative success of ANNs in predicting various geotechnical engineering properties and behaviour. It is not intended to cover every single application or scientific paper that can be found in the literature. Rather, the intention is to provide a general overview of some of the more relevant ANN applications in geotechnical engineering problems. Some works are selected to be described in some detail while, others are acknowledged for reference purposes.

3.2 Pile Capacity

The prediction of the load capacity, particularly that based on pile driving data, has been examined by several ANN researchers. Goh (1994a; 1995b) presented a neural network to predict the friction capacity of piles in clays. The neural network was trained with field data of actual case records. The model inputs were considered to be the pile length, pile diameter, mean effective stress and the undrained shear strength. The skin friction resistance was the only model output. The results obtained by utilising the neural network were compared with the results obtained by the method of Semple and Rigden (1986) and the β method (Burland 1973). The methods were compared using regression analysis as well as the error rate (Yeh et al. 1993) as shown in Table 3.1. It is evident from Table 3.1 that the ANN model outperforms the conventional methods. The study also pointed out that the main criticism of the ANN methodology is its inability to trace and explain the logic it uses to arrive at the prediction.

Table 3.1: Summary of correlation coefficients and error rate for friction pile capacity (Goh 1995b)

Method	Coefficient of correlation, *r*		Error rate (kPa)	
	Training	Testing	Training	Testing
ANN	0.985	0.956	1.016	1.194
Semple and Rigden (1986)	0.976	0.885	1.318	1.894
β method	0.731	0.704	4.824	3.096

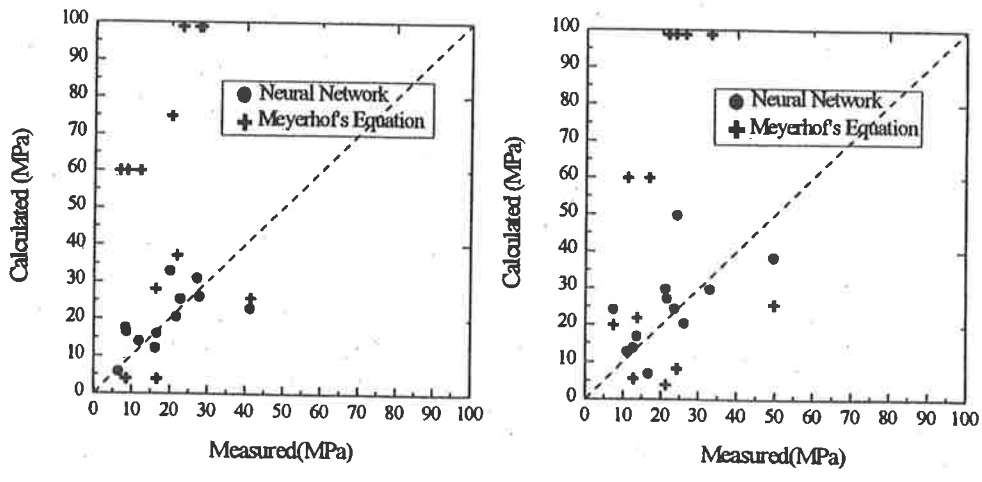
Goh (1995a; 1996b), soon after, developed another neural network to estimate the ultimate load capacity of driven piles in cohesionless soils. In this study, the data used were derived from the results of load tests on timber, precast concrete and steel piles driven into sandy soils. The inputs to the ANN model that were found to be more significant were the hammer weight, hammer drop, pile length, pile weight, pile cross sectional area, pile set, pile modulus of elasticity and the hammer type. The model output was the pile load capacity. When the model was examined using a testing set, it was observed that the neural network successfully modelled the pile load capacity. By examining the connection weights, it was observed that the more important input factors are the pile set, the hammer weight and the hammer type. The study compared the results of the ANNs with the following common formulae: Engineering News formula (Wellington 1892), Hiley formula (Hiley 1922) and Janbu formula (Janbu 1953). Regression analysis was carried out to obtain the coefficients of correlation, *r*, of the predicted versus measured results for the ANNs and the traditional methods. Table 3.2 summarises the regression analysis results, which indicate that the neural network predictions of the load capacity of driven piles were found to be significantly better than those obtained using the other methods.

Table 3.2: Summary of regression analysis results of pile capacity prediction (Goh 1995a)

Method	Coefficient of correlation, *r*	
	Training data	Testing data
ANN	0.96	0.97
Engineering News	0.69	0.61
Hiley	0.48	0.76
Janbu	0.82	0.89

Chan et al. (1995) developed a neural network as an alternative to the pile driving formulae. The network was trained with the same input parameters listed in the simplified Hiley formula (Broms and Lim 1988), including the elastic compression of the pile and soil, the pile set and the driving energy delivered to the pile. The model output considered was, again, the pile capacity. The desired output value of the pile capacity that was used in the training process was estimated by using the computer algorithm CAPWAP (Rausche et al. 1972) or the CASE method (Goble et al. 1975). The root mean squared percentage error, EN, of the neural network ($EN = \sqrt{\sum_{j=1}^{n}\left(\frac{y_j - d_j}{d_j}\right)^2 \Big/ n}$ where y_j and d_j are the network predicted output and the desired output, respectively) was 13.5% for the training set, and 12.0% for the testing set, compared with the results of 15.7% in the training and testing sets for the simplified Hiley formula.

Lee and Lee (1996) utilised ANNs to predict the ultimate bearing capacity of piles. The problem was simulated using data obtained from model pile load tests using a calibration chamber and results of in-situ pile load tests. For the simulation using the model pile load test data, the model inputs were the penetration depth ratio (i.e. penetration depth of pile/pile diameter), the mean normal stress of the calibration chamber and the number of blows. The ultimate bearing capacity was the model output. The prediction of the ANN model showed a maximum error not greater than 20% and an average summed square error of less than 15%. For the simulation using the in-situ pile load test data, five input variables were used representing the penetration depth ratio, the average standard penetration number along the pile shaft, the average standard penetration number near the pile tip, pile set and hammer energy. The data were arbitrarily partitioned into two parts, odd and even numbered sets and two neural network models were developed. The results of these models were compared with Meyerhof's equation (Meyerhof 1976), based on the average standard penetration value. Figure 3.1 shows the plots of the testing set results of the estimated versus measured pile bearing capacity obtained from the neural network models and Meyerhof's equation. The plots in Figure 3.1 show that the predicted values from the neural networks matched the measured values much better than those obtained from Meyerhof's equation.



(a) odd numbered data trained **(b) even numbered data trained**

Figure 3.1: Testing results of predicted vs measured pile bearing capacity from in-situ pile load test (Lee and Lee 1996)

Abu-Kiefa (1998) introduced three ANN models (referred to in the paper as GRNNM1, GRNNM2 and GRNNM3) to predict the capacity of driven piles in cohesionless soils. The first model was developed to estimate the total pile capacity. The second model was employed to estimate the tip pile capacity, whereas the final model was used to estimate the shaft pile capacity. In the first model, five variables were selected to be the model inputs. These inputs were the angle of shear resistance of the soil around the shaft, the angle of shear resistance at the tip of the pile, the effective overburden pressure at the tip of the pile, pile length and the equivalent cross-sectional area of the pile. The model, again, had one output representing the total pile capacity. In the model used to evaluate the pile tip capacity, the above variables were also used. The number of input variables used to predict the pile shaft capacity was four, representing the average standard penetration number around the shaft, the angle of shear resistance around the shaft, pile length and pile diameter. The results of the networks obtained in this study were compared with four other empirical techniques. These techniques were those proposed by Meyerhof (1976), Coyle and Castello (1981), the American Petroleum Institute (1984) and Randolph (1985). The results of the total pile capacity prediction demonstrated high coefficients of determination (0.95) for all data records obtained from the neural network model, while those for the other methods ranged between 0.52 and 0.63. Figures 3.2 to 3.4 show the measured versus predicted values of all data records for the pile capacity, tip pile capacity and shaft pile capacity, respectively. It can be seen from these figures that the predictions of the ANNs produce less scatter than the predictions of all other methods, and thus provide the best prediction of pile load capacity, tip pile capacity and shaft pile capacity.

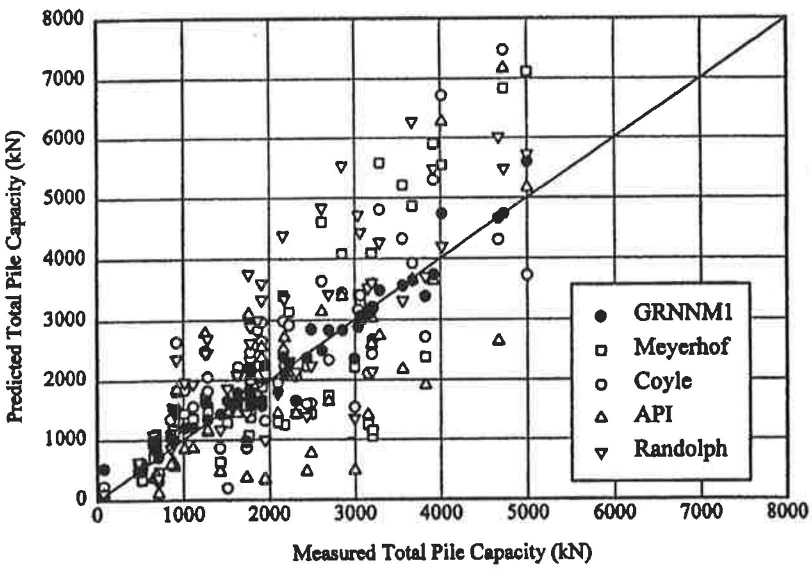


Figure 3.2: Comparison of predicted and measured total pile capacity (Abu-Kiefa 1998)

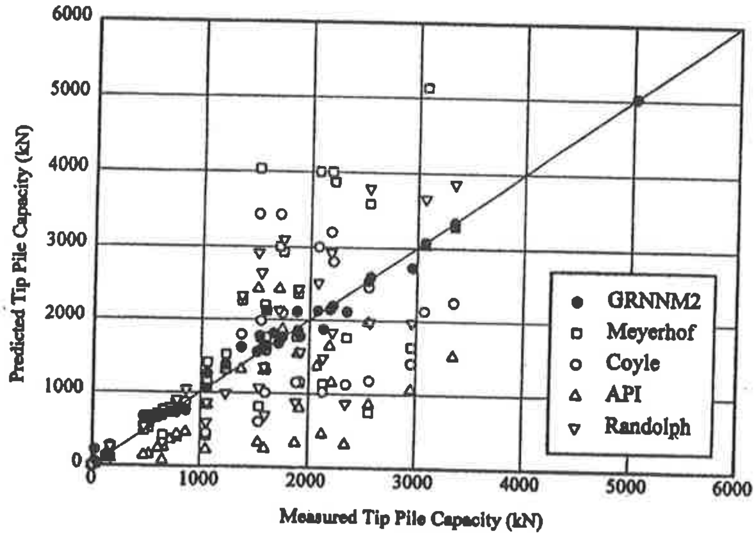


Figure 3.3: Comparison of predicted and measured tip pile capacity (Abu-Kiefa 1998)

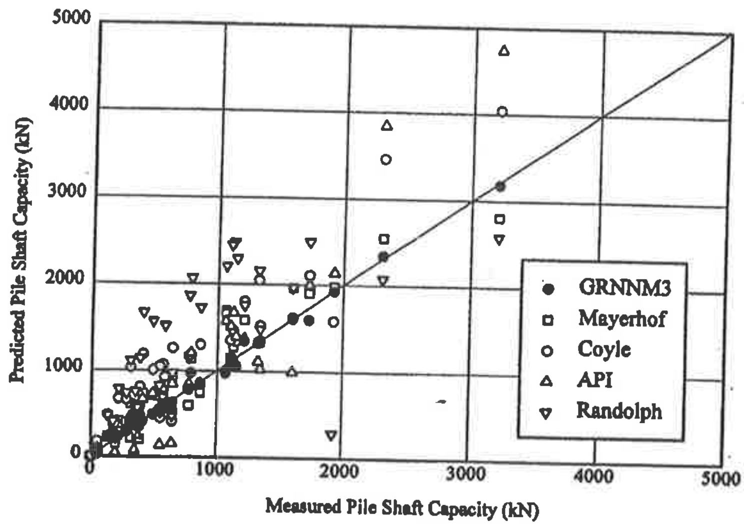


Figure 3.4: Comparison of predicted and measured shaft pile capacity (Abu-Kiefa 1998)

Teh et al. (1997) proposed a neural network for estimating the static pile capacity determined from dynamic stress-wave data for precast reinforced concrete piles with a square section. The networks were trained to associate the input stress-wave data with capacities derived from the CAPWAP technique (Rausche et al. 1972). The study was concerned with predicting the 'CAPWAP predicted capacity' rather than the true bearing capacity of the pile. The neural network learned the training data set almost perfectly for predicting the static total pile capacity with a root mean square error of less than 0.0003. The trained neural network was assessed for its ability to generalise by means of a testing data set. Good prediction was obtained for seven out of ten piles as shown in Figure 3.5.

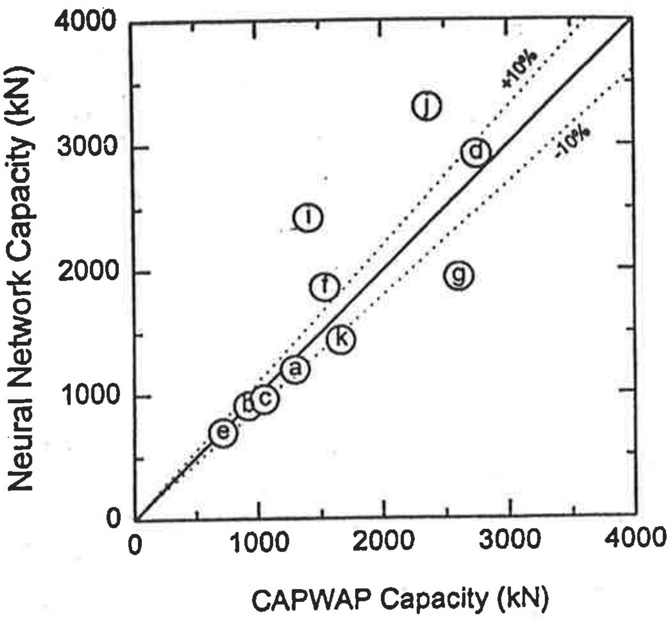


Figure 3.5: Static capacity predicted by CAPWAP and neural network for testing set (Teh et al. 1997)

Another application of ANNs includes the prediction of axial and lateral load capacity of steel H-piles, steel piles and prestressed and reinforced concrete piles by Nawari et al. (1999). In this application, ANNs were found to be an accurate technique for the design of pile foundations.

3.3 Settlement of Foundations

As mentioned previously, the design of foundations is generally controlled by the criteria of bearing capacity and settlement, the latter often governing. The estimation of the settlement of foundations is very complex, uncertain and not yet entirely understood. This fact has encouraged some researchers to apply the ANN technique to settlement prediction. Goh (1994a) developed a neural network for the prediction of settlement of a vertically loaded pile foundation in a homogeneous soil stratum. The input variables for the neural network consisted of the 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 variable was the pile settlement. The desired output that was used for the ANN model training was obtained by means of finite element and integral equation analyses developed by Randolph and Wroth (1978). A comparison of the theoretical and predicted settlements for the training and testing sets is given in Figure 3.6. The results show that the neural network was able to successfully model the settlement of pile foundations.

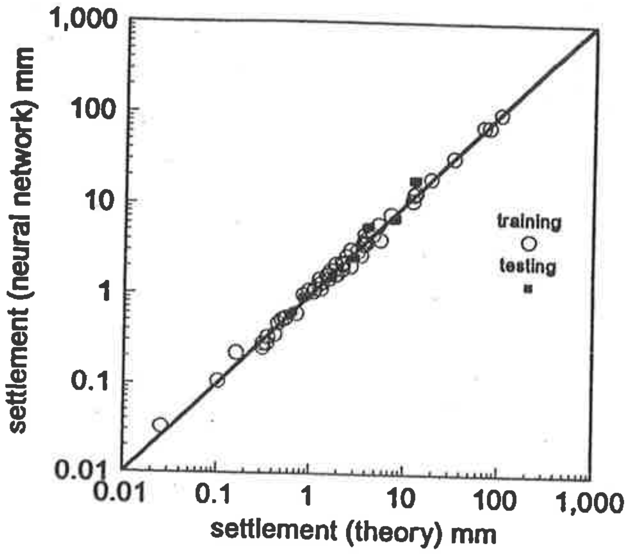
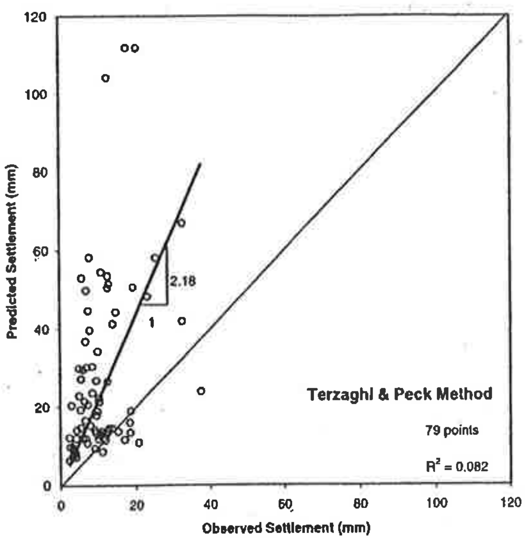


Figure 3.6: Comparison of theoretical settlements and neural network predictions (Goh 1994a)

Sivakugan et al. (1998) carried out a preliminary study on a small set of data to explore the possibility of using neural networks to predict the settlement of shallow foundations on sands. A neural network was trained with five inputs representing the net applied pressure, average blow count from the standard penetration test, width of the foundation, shape of the foundation and depth of the foundation. The output was the settlement of the foundation. With the aid of Cascade-Correlation, a network with one hidden layer and 11 hidden nodes was found optimal. The results obtained by the neural network were compared with methods proposed by Terzaghi and Peck (1967) and Schmertmann (1970). Based on the results obtained, it was shown that the traditional methods of Terzaghi and Peck and Schmertmann overestimate the settlements by about 2.2 times and 3.4 times, respectively, as shown in Figure 3.7. In contrast, the predictions using the ANN model were good (Figure 3.8). Using the same neural network features, Arnold (1999) extended the work done by Sivakugan et al. (1998) with a database containing a larger number of data cases. His work, although relatively superficial, found that the best network consisted of 18 hidden layer nodes with correlation coefficients equal to 0.954, 0.955 and 0.944 for the training, testing and validation sets, respectively. It should be noted that 18 hidden layer nodes are considered to be large for a network with 5 input variables, which may affect the generalisation ability of the model, as discussed in §2.5.4.



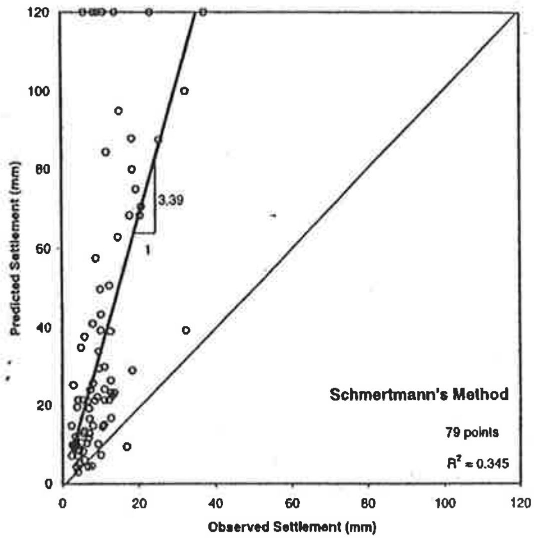


Figure 3.7: Settlements predicted using traditional methods (Sivakugan et al. 1998)

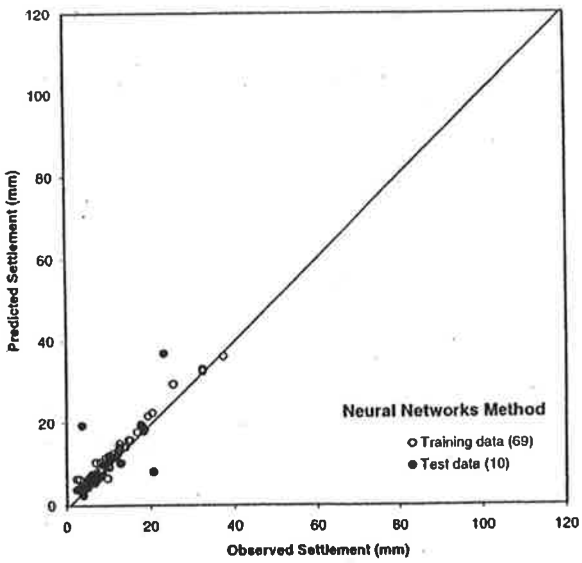


Figure 3.8: Settlement prediction using artificial neural network (Sivakugan et al. 1998)

3.4 Soil Properties and Behaviour

Soil properties and behaviour is an area that has attracted many researchers to modelling using ANNs. Developing engineering correlations between various soil parameters is an issue discussed by Goh (1995a; 1995c). Goh used neural networks to model the correlation between the relative density and the cone resistance from the cone penetration test (CPT), for both normally consolidated and overconsolidated sands. Laboratory data, based on calibration chamber tests, were used to successfully train and test the neural network model. The neural network model used the relative density and the mean effective stress of soils as inputs and the CPT cone resistance as a single output. The ANN model was found to give high coefficients of correlation of 0.97 and 0.91 for the training and testing data, respectively, which indicated that the neural network was successful in modelling the non-linear relationship between the CPT cone resistance and the input parameters. Many other studies have successfully used ANNs for modelling soil properties and behaviour, which, for brevity, are acknowledged for reference purposes in the following paragraphs.

Ellis et al. (1995) developed an ANN model for sands based on grain size distribution and stress history. Sidarta and Ghaboussi (1998) employed an ANN model within a finite element analysis to extract the geomaterial constitutive behaviour from non-uniform material tests. Penumadu and Jean-Lou (1997) used neural networks for representing the behaviour of sand and clay soils. Ghaboussi and Sidarta (1998) used neural networks to model both the drained and undrained behaviour of sandy soil subjected to triaxial compression-type testing. Penumadu and Zhao (1999) also used ANNs to model the stress-strain and volume change behaviour of sand and gravel under drained triaxial compression test conditions. Zhu et al. (1998a; 1998b) used neural networks for modelling the shearing behaviour of a fine-grained residual soil, dune sand and Hawaiian volcanic soil. Cal (1995) used a neural network model to generate a quantitative soil classification from three main factors (plastic index, liquid limit and clay content). Najjar et al. (1996) showed that neural network-based models can be used to accurately assess soil swelling, and that neural network models can provide significant improvements in prediction accuracy over statistical models. Romero and Pamukcu (1996) showed that neural networks are able to effectively characterise and estimate the shear modulus of granular materials. Agrawal et al. (1994), Gribb and

Gribb (1994) and Najjar and Basheer (1996b) all used neural network approaches for estimating the permeability of clay liners. Basheer and Najjar (1995) and Najjar et al. (1996) presented neural network approaches for soil compaction.

ANNs are also used successfully in other applications including: modelling the mechanical behaviour of medium-to-fine sand (Ellis et al. 1992), modelling rate-dependent behaviour of clay soils (Penumadu et al. 1994), simulating the uniaxial stress-strain constitutive behaviour of fine-grained soils under both monotonic and cyclic loading (Basheer 1998; Basheer and Najjar 1998), characterising the undrained stress-strain response of Nevada sand subjected to both triaxial compression and extension stress paths (Najjar and Ali 1999; Najjar et al. 1999), predicting the axial and volumetric stress-strain behaviour of sand during loading, unloading and reloading (Zhu and Zaman 1997) and predicting the anisotropic stiffness of granular materials from standard repeated load triaxial tests (Tutumluer and Seyhan 1998).

3.5 Liquefaction

Liquefaction is a phenomenon which occurs mainly in loose and saturated sands as a result of earthquakes. It causes the soil to lose its shear strength due to an increase in pore water pressure, often resulting in large amounts of damage to most civil engineering structures. Determination of liquefaction potential due to earthquakes is a complex geotechnical engineering problem. Goh (1994b) used neural networks to model the complex relationship between seismic and soil parameters in order to investigate liquefaction potential. The neural network used in this work was trained using case records from 13 earthquakes that occurred in Japan, United States and Pan-America during the period 1891–1980. The study used eight input variables and a single output variable. The input variables were the SPT N-value, fines content, mean grain size, total stress, effective stress, equivalent dynamic shear stress, earthquake magnitude and the maximum horizontal acceleration at the ground surface. The output was assigned a binary value of 1 for sites with extensive or moderate liquefaction, and a value of 0 for sites with marginal or no liquefaction. The results obtained by the neural network model were compared with those obtained using the method of Seed et al. (1985). The study showed that the neural network gave correct predictions in 95% of cases, whereas the method of Seed et al. (1985) gave a success rate of 84%. Goh (1996a) also used neural networks to assess liquefaction potential from CPT resistance data. The data records were taken for sites of sand and silty sand deposits in Japan, China, United States and Romania, representing five earthquakes that occurred during the period 1964–1983. A similar neural network modelling strategy, as used by Goh (1994b), was used for this study and the results were compared with the method of Shibata and Teparaksa (1988). The neural network showed a 94% success rate, which is equivalent to the same number of error predictions as the conventional method by Shibata and Teparaksa (1988).

Two other papers (Najjar and Ali 1998; Ural and Saka 1998) also used CPT data to evaluate soil liquefaction potential and resistance. Najjar and Ali (1998) used neural networks to characterise the soil liquefaction resistance utilising field data sets representing various earthquake sites from around the world. The ANN model that was developed in this work was generated to produce a liquefaction potential assessment chart that could be used by geotechnical engineers in liquefaction assessment. Ural and Saka (1998) also used neural networks to analyse liquefaction. Comparisons between the ANN approach and a simplified liquefaction procedure indicated a similar rate of success for the neural network approach as for the conventional approach.

Other applications of ANNs for liquefaction prediction include the prediction of liquefaction resistance and potential (Juang and Chen 1999), investigation of the accuracy of liquefaction prediction of ANNs compared with fuzzy logic and statistical approaches (Ali and Najjar 1998) and assessment of liquefaction potential using standard penetration test results (Agrawal et al. 1997).

3.6 Site Characterisation

Site characterisation is an area concerned with the analysis and interpretation of geotechnical site investigation data. Zhou and Wu (1994) used a neural network model to characterise the spatial distribution of rockhead elevations. The data used to train the model were taken from seismic refraction surveys on more than 11 km of transverse lines. The network used the spatial position (x- and y-coordinate) and the surface elevation as inputs, and was used to estimate the rockhead elevation at that location as the output. The trained network was tested to estimate the rockhead elevations for all locations within the area of investigation by producing a contour map. Results from the neural network model compared well with similar contour maps generated using kriging (Journel and Huijbregts 1978), with the additional benefit that neural networks do not make assumptions or simplify spatial variations.

A similar application relevant to ground water characterisation was described by Basheer et al. (1996). Basheer et al. (1996) indicated that neural networks can be used to map and logically predict the variation of soil permeability in order to identify landfill boundaries and to construct a waste landfill. Rizzo et al. (1996) presented a new site characterisation method called SCANN (Site Characterisation using Artificial Neural Networks) that is based on the use of neural networks to map discrete spatially-distributed fields. Other applications were presented by Najjar and Basheer (1996a) and Rizzo and Dougherty (1994).

3.7 Earth Retaining Structures

Goh et al. (1995) developed a neural network model to provide initial estimates of maximum wall deflections for braced excavations in soft clay. The neural network was used to synthesise data derived from finite element studies on braced excavations in clay. The input parameters used in the model were the excavation width, soil thickness/excavation width ratio, wall stiffness, height of excavation, soil undrained shear strength, undrained soil modulus/shear strength ratio and the unit weight of the soil. The maximum wall deflection was the only output. Using regression analysis, the scatter of the predicted neural network deflections relative to the deflections obtained using the finite element method were assessed. The results produced high coefficients of correlation for the training and testing data of 0.984 and 0.967, respectively. Some additional testing data from actual case records were also used to confirm the performance of the trained neural network model. The agreement of the neural network predicted and measured wall deflections was encouraging, as shown in Table 3.3. The study intended to use the neural network model as a time-saving and user-friendly alternative to the finite element method.

Table 3.3: Comparison of neural network and field measurements (Goh 1995)

Case history	Measured wall deflection (mm)	Predicted wall deflection (mm)
Laveder (Singapore)	32	31
Laveder (Singapore)	36	28
Telecom (Singapore)	56–84	65
Vaterland 3	76	76
(NGI 1962)	114–140	107
San Francisco	20–60	59
(Mana 1977)	72–150	122

3.8 Slope Stability

Ni et al. (1996) proposed a methodology of combining fuzzy sets theory with artificial neural networks for evaluating the stability of slopes. In this approach, the input parameters were gradient, horizontal profile, vertical profile, location, height, geological origin, soil texture, depth of weathering, direction of slopes, vegetation, land use, maximum daily precipitation and maximum hour precipitation. The output was the slope failure potential. A number of hypothetical natural slopes were evaluated by both neural networks and an analytical model, and the results of the neural network approach were in good agreement when compared with those obtained by the analytical model.

3.9 Tunnels and Underground Openings

Shi et al. (1998) presented a study of neural networks for predicting settlements of tunnels. A general neural network model was trained and tested using data from the 6.5 km Brasilia Tunnel, Brazil. The study identified many factors to be used as the model inputs and three settlement parameters as the model outputs. The input parameters were the length of excavation from drive start, the depth of soil cover above tunnel crown, the area of tunnel section, the delay for closing the invert, water level depth, the rate of advance of excavation, construction method, the SPT mean blow count at the tunnel crown level, the SPT mean blow count at the tunnel spring-line level and the SPT mean blow count at tunnel inverted arch level. The three output parameters were the settlement at the face passage, settlement at the invert closing and the final settlement after stabilisation. The results showed that the neural network model could not achieve an appropriate level of accuracy. To improve the prediction accuracy, the study proposed a modular neural network model based on the concept of integrating multiple neural network modules in one system, with each module being constrained to consider one specific situation of a complicated real world problem. The modular concept showed an improvement in terms of model convergence and prediction. The capability to improve the models developed in this work was later extended by Shi (2000) by applying input data transformation. This extended study indicated that distribution transformation of the input variables reduced the prediction error by more than 13%.

Lee and Sterling (1992) developed a neural network for identification of probable failure modes for underground openings from prior case history information. The study used the knowledge obtained by the neural network to produce an assistance system for the design of tunnels. Sterling and Lee (1992) used the neural network as part of a knowledge-based expert system for assisting with tunnel design. Moon et al. (1995) also used ANNs, integrated with an expert system, for the preliminary design of tunnels.

3.10 Summary

It is evident from the review presented in this chapter that ANNs have been applied successfully to many geotechnical engineering areas. This includes the prediction of pile capacity, predicting the settlement of foundations, modelling soil properties and behaviour, determination of liquefaction potential, site characterisation, modelling earth retaining structures, evaluating slope stability and the design of tunnels and underground openings. Perhaps the most successful and well-established applications are the capacity prediction of driven piles, liquefaction and the prediction of soil properties and behaviour. There are several areas in which the feasibility of ANNs has yet to be tested, such as bearing capacity prediction of shallow foundations, capacity of bored piles, design of sheet pile walls and dewatering, among others. The feasibility of ANNs for some other applications such as settlement of shallow foundations has been tested in a preliminary fashion and has shown some degree of success. However, thus far, a comprehensive study has yet to be achieved. The ANN models that have been developed in the literature for settlement prediction of shallow foundations have been built on either a limited number of data cases (e.g. Sivakugan et al. 1998) or have suffered from the lack of a comprehensive procedure for testing their robustness and generalisation ability (e.g. Arnold 1999). Consequently, these models need to be treated with caution until the development of more well-established models, which will be one of the main focuses of the present research.

Based on the results of the studies reviewed in this chapter, it is also evident that ANNs perform better than, or as well as, the conventional methods used as a basis for comparison in many situations, whereas they fail to perform well in a few. This is implies that ANNs are a powerful and practical tool for solving many problems in the field of geotechnical engineering. The predictions of ANN models developed in the applications reviewed in this chapter were based on an assumption that the data used for ANN model development are ideal (i.e. have no parameter uncertainty) and that the model is a perfect predictor (i.e. has no model uncertainty). It may be possible to achieve better predictions if these uncertainties are considered. In Chapter 7, the issue of model and parameter uncertainties will be considered and discussed in detail for settlement prediction of shallow foundations on cohesionless soils. In the following chapter, the problem of settlement prediction of shallow foundations on cohesionless soils will be discussed.

Chapter 4

Settlement of Shallow Foundations on Cohesionless Soils

4.1 Introduction

This chapter provides a background to the settlement of foundations and focuses on the settlement of shallow foundations on cohesionless soils. The major causes of settlement of shallow foundations are presented and the factors affecting the settlement prediction of shallow foundations are reviewed and discussed. In addition, the available methods for settlement prediction of shallow foundations on cohesionless soils are categorised and the more successful are described in more detail, with some discussion of their relative advantages and disadvantages.

4.2 Causes of Settlement of Shallow Foundations

Settlement of shallow foundations may arise from a number of causes (Poulos 1975), including:

1. Static loads imposed by the weight of structures;
2. Dynamic loads produced by machinery, earthquakes, moving loads on roads or airfield pavements;
3. Changes in moisture content from seasonal fluctuation in the water table, rainfall and evaporation or the absorption of water by roots of large trees; or
4. The effects of nearby construction resulting from adjacent excavation, pile driving and dewatering.

This chapter concentrates on settlement of shallow foundations on cohesionless soils due to static loads only, as this is the focus of the research described in this thesis.

4.3 Factors Affecting Settlement of Shallow Foundations on Cohesionless Soils

A proper estimation of settlement of shallow foundations on cohesionless soils can be obtained through a thorough understanding of the factors affecting settlement. Such factors can be categorised as 'primary' and 'secondary' factors, which are discussed below.

4.3.1 Primary Factors

In a statistical analysis carried out by Burland and Burbidge (1985), for more than 200 case records of settlement of shallow foundations on sands and gravels, three factors were found to govern settlement prediction:

- Footing width, B;
- Footing net applied pressure, q; and
- Soil compressibility within the depth of influence of foundation.

This conclusion has also been recognised by most traditional methods for settlement prediction of shallow foundations on cohesionless soils.

Soil compressibility within the depth of influence of a foundation requires the assignment of a depth over which the compressibility of the soil beneath the footing significantly influences the settlement and the assignment of soil properties that can accurately reflect this compressibility. There is no unanimous agreement in the literature on the definition of the depth of influence of a foundation. For example, Terzaghi and Peck (1948), Bazaraa (1967) and D'Appolonia et al. (1968) recommended taking a depth of influence equal to the width of the footing, B. Parry (1971) and Schultze and Sherif (1973) took a depth of influence equal to $2B$ in their settlement analyses. Schmertmann et al. (1978) considered a depth of influence of $2B$ for square and circular footings and $4B$ for continuous footings. Gupta (1993) assumed this depth to be $2B$ for $L/B \leq 3$ and $4B$ for $L/B > 3$, where L is the footing length. When the average SPT blow count N decreases with depth, Burland and Burbidge (1985) took this depth to be equal to the lesser of $2B$ or the depth from the bottom of the footing to bedrock. On the other hand, when the blow count is constant or increasing with depth, Burland and Burbidge (1985) considered a depth of influence of approximately $B^{0.75}$. In this research, the guidelines proposed by Burland and Burbidge (1985) are used for the definition of the depth of influence, as most case records in the database used for the purpose of this thesis were obtained from Burland and Burbidge (1985).

Determining an appropriate value of soil compressibility is the most difficult part of settlement prediction (Sivakugan et al. 1998). Due to the difficulty in obtaining undisturbed soil samples from cohesionless soils, soil compressibility is often obtained from in-situ tests. Penetration tests such as the cone penetration test (CPT) and the standard penetration test (SPT) are commonly used for estimating the compressibility of soils. The CPT is one of the best available penetration methods for determining soil compressibility and has the advantage of giving a continuous profile of soil strength with penetration. The CPT is also rapid and inexpensive compared with other soil profiling techniques (Orchant et al. 1987). However, the CPT suffers from a number of shortcomings. It does not provide soil samples for visual inspection and, thus, additional boreholes are necessary to correlate the penetration resistance with the soil profile. Moreover, small variations in sand density and grain size often produce very large changes in CPT penetration resistance, making it difficult to interpret the general nature of the in-situ soil deposit (D'Appolonia and D'Appolonia 1970). In addition, in dense and very dense soils, it is often difficult to push the penetrometer to the required depth, which often needs substantial jacking reaction (Bazaraa 1967).

The SPT, on the other hand, is one of the most commonly used tests in practice for measuring the compressibility of cohesionless soils (D'Appolonia and D'Appolonia 1970). The SPT has been used successfully for many years with thousands of foundations and other structures (Gordon and Fletcher 1965). The SPT has the advantage that it is often conducted as part of a routine subsurface exploration program, which enables the visual inspection of soil samples. The SPT can also detect the important variations in granular soil density and its procedure is simple, relatively inexpensive and can be applied both below and above the water table with a reasonable degree of accuracy. However, the SPT has also several notable disadvantages. The SPT does not give a continuous soil profile and is notorious for its unreliability with regard to the reflection of soil compressibility. The results of the SPT are also uncertain due to several factors associated with the test equipment and procedures (see Gordon and Fletcher 1965; Wang and Lu 1982).

It has been argued that the CPT gives a better indication of the soil properties than the SPT. In fact, as time proceeds, the CPT is replacing the SPT as the industry standard. Whilst the SPT is not the most accurate in-situ method for measuring soil compressibility, it is used extensively worldwide and most available data sets in the literature include SPT measurements rather than more accurate estimates of soil properties from the CPT. The performance of ANN models depends mainly on collecting as much reliable data as possible. Consequently, the SPT is used as a measure of soil compressibility for the development of the ANN models introduced in this thesis. As a result, a brief description of the SPT test procedure is given below.

Standard Penetration Test (SPT)

The SPT is carried out using the following steps (Standards Australia 1993):

1. A vertical hole of at least 65 mm diameter is drilled to the depth at which the test is to be conducted for the first time.
2. A split spoon sampler (Figure 4.1) is inserted into the hole via steel rods.
3. A 63.5 ± 1 kg hammer, as shown in Figure 4.2, is raised a distance of 760 ± 15 mm using a self-tripping mechanism and is allowed to fall freely due to lifting winch inertia.
4. The process is repeated until the sampler penetrates the soil for a total distance of 450 mm.
5. The number of hammer blows required for each 150 mm interval is recorded.
6. The test is stopped if (i) a total of 30 blows causes less than 100 mm penetration at any stage or (ii) there is no measurable penetration or the hammer bounces for 5 consecutive blows.
7. The blow counts for the last 300 mm of penetration are summed and the number of blows of the standard penetration test (N) is computed, noting that the blow counts for the first 150 mm are not used for computing N, as this soil is assumed to be disturbed by the drilling process.

8. The sampler is extracted and the soil sample is removed, inspected and placed in an airtight container to maintain the sample moisture content, if required.
9. The hole is then drilled to the depth required for the next test and steps 2 to 8 are repeated, as required.

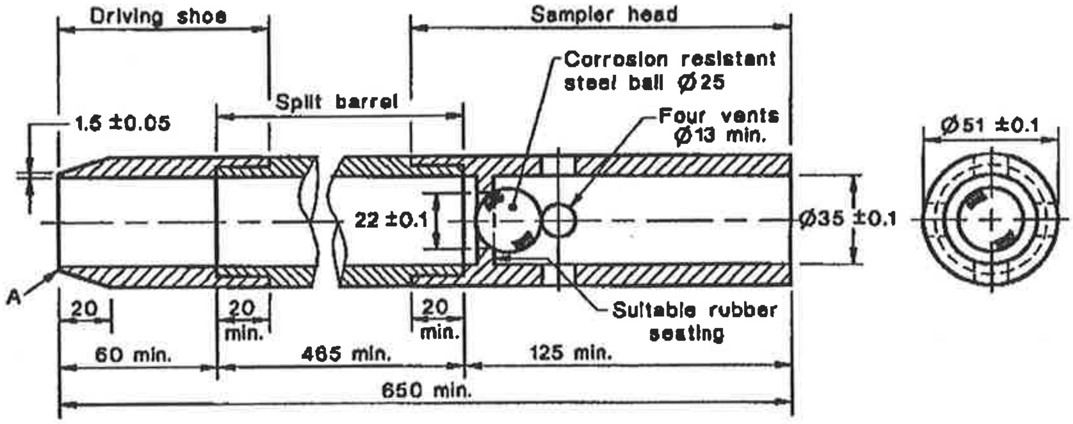


Figure 4.1: The split spoon sampler (Standards Australia 1993)

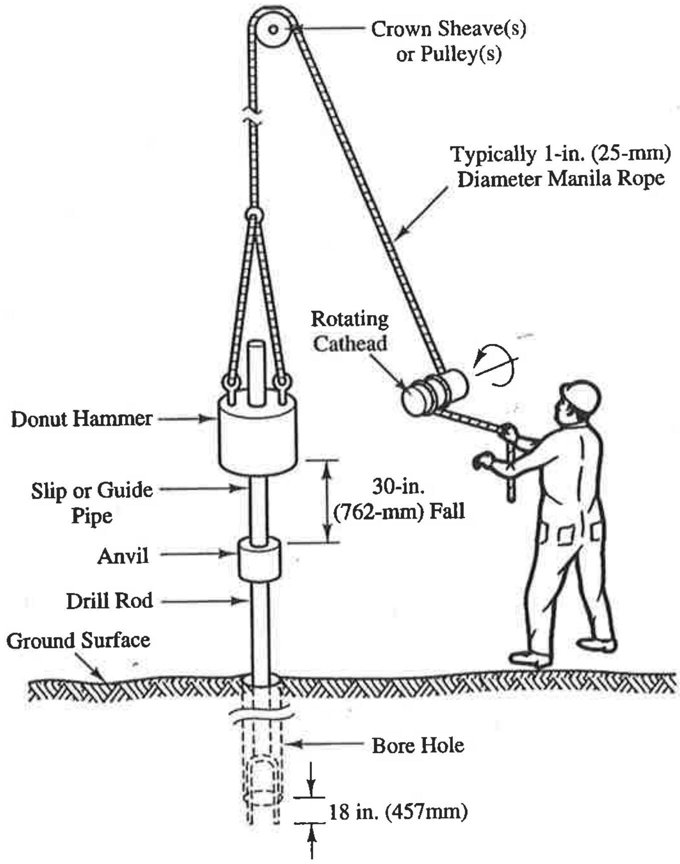


Figure 4.2: The standard penetration test (Coduto 1999)

The blow counts obtained from the SPT may be affected by submergence and overburden pressure. Terzaghi and Peck (1948) found that the SPT blow count, *N*, for dense, fine or silty sand beneath the water table is abnormally high if the measured value of N is greater than 15. This is due to the tendency of dense, fine or silty submerged sand to dilate during shear in undrained conditions. Consequently, Terzaghi and Peck (1948) recommended a correction for reducing the measured N values for dense, fine or silty submerged sand when $N > 15$ as follows:

$$N_{corrected} = 15 + 0.5(N - 15) \tag{4.1}$$

The above proposal was confirmed by a study carried out by Bazaraa (1967) for a large number of results of SPT tests within 1 m above and below the water table. In his study, Bazaraa (1967) concluded that the effect of submergence on very fine or silty sand, in general, increases the number of blows of the SPT test. However, Bazaraa (1967) established his conclusion for loose rather than dense sand. As a consequence, Bazaraa (1967) proposed an alternative correction for the SPT blow count beneath the water table in fine or silty sand as:

$$N_{corrected} = 0.6N \tag{4.2}$$

Schultze and Menzenbach (1961) and Bazaraa (1967) have shown that medium to coarse sand and gravel are not affected by submergence. Burland and Burbidge (1985) recommended no correction to N be taken for submergence. However, for very fine and silty sand below the water table, Burland and Burbidge (1985) used the submergence correction proposed by Terzaghi and Peck (1948) when $N > 15$. For soils consisting of gravel or sandy gravel, Burland and Burbidge (1985) proposed a correction for N as follows:

$$N_{corrected} = 1.25N \tag{4.3}$$

Another factor that may affect the results of the SPT blow count is the influence of the effective overburden pressure, which is a measure of the confining pressure at the level where the SPT is carried out. Many researchers (e.g. Sutherland 1963; Thorburn 1963; Alpan 1964) recommended that the SPT blow count should be corrected for overburden pressure. Mansur and Kaufman (1958), Philox (1962) and Zolkov and Wiseman (1965) confirmed the substantial influence of overburden pressure on the penetration resistance of the SPT. Gibbs and Holtz (1957) and Bazaraa (1967) developed SPT correction charts for overburden pressure. On the other hand, Peck et al. (1974) suggested no correction for overburden pressure should be applied unless footings are below 6 m or above 2 m in depth. Burbidge (1982) suggested that overburden pressure is not usually an important correction to the SPT blow counts of granular soils because the overburden pressure is normally restored after the construction of foundations and before the beginning of any settlement monitoring. Burland and Burbidge (1985) recommended no correction to N be taken for overburden pressure. Since most case records in the database used for the purpose of this thesis were obtained from Burland and Burbidge (1985), the correction factors recommended by Burland and Burbidge (1985) were applied.

4.3.2 Secondary Factors

In addition to the three primary factors discussed in the previous section, there are many other factors that contribute to a lesser degree to the settlement of shallow foundations on cohesionless soils and, thus, can be considered as secondary factors. These factors include:

- Depth of the water table beneath the foundation level;
- Time since load application;
- Footing geometry or shape (length-to-width), L/B;
- Depth of footing embedment, D_f (usually expressed as a ratio of B and known as the footing embedment ratio); and
- Thickness of the soil layer beneath the foundation.

Burland and Burbidge (1985) stated that the depth of water table beneath the foundation level has little influence on settlement. Meyerhof (1965) also concluded that the water table is reflected in the measured blow count of the SPT. Consequently, the effect of depth of water table can be ignored without any significant error to the settlement prediction.

Burland and Burbidge (1985) also found that settlement of shallow foundations on sand and gravels exhibit time-dependence. However, no distinct pattern emerged from their statistical analysis in relation to the effect of time-dependence. Fang (1991) observed that time-dependence for settlement of shallow foundations on cohesionless soils may result from the consolidation of thin layers of silt and clay within sand or gravel soils and, consequently, the impact of this factor is not recommended in the calculation of settlement of foundations on cohesionless soils.

Amar and Baguelin (1984) concluded that foundation geometry, L/B, was not shown to have any significant influence on settlement. On the other hand, Burland and Burbidge (1985) concluded from their statistical study that there is a correlation between settlement and L/B of the foundation. However, they stated that the correction factor for L/B is quite small. Burbidge (1982) found that there is no significant difference between the settlement of square, circular and continuous strip foundations having the same width, B, on the same soil. Consoli et al. (1998) also demonstrated that the shape of the loaded area does not influence initial soil compressibility.

In their study, Burland and Burbidge (1985) found that, for foundations with $D_f / B < 3.3$, there is no obvious correlation between settlement and depth of footing embedment. This conclusion confirms the results obtained previously by D'Appolonia et al. (1968) who found that, from the analysis of a number of results on a single site, only 12% reduction in settlement occurred when D_f/B increases from 0.5 to 1.0. Moreover, Christian and Carrier (1978) stated that the material above the foundation level does not seem to contribute very significantly to the settlement behaviour. Consequently, Christian and Carrier (1978) demonstrated that ignoring the depth of footing embedment in calculating settlement gives reasonably satisfactory results, especially when some other factors such as heterogeneity and non-linearity of the soil are to be taken into consideration. The relative importance of the previous factors (main and secondary) on settlement will be investigated in Chapter 5 by carrying out a sensitivity analysis on the optimal ANN model.

In the present research and for the purpose of predicting the settlement of shallow foundations on cohesionless soils using ANNs, the following aspects are considered.

The depth of the water table and time-dependence are not included in the study as proposed by Meyerhof (1965) and Fang (1991), respectively. There are also insufficient data for the thickness of the soil layer in the database used in this research and thus it is not considered in this work. The conclusion made by Burbidge (1982) that there is no significant difference between the settlement of circular or square footings having the same width, *B*, on the same soil is considered and thus *L* /*B* is assumed to equal unity for circular footings. In the following section, three of the most commonly used methods for settlement prediction of shallow foundations on cohesionless soils will be described in some detail. These methods will be the basis of assessing the relative performance of the ANN models. These include the methods proposed by Meyerhof (1965), Schultze and Sherif (1973) and Schmertmann et al. (1978).

4.4 Methods of Settlement Prediction of Shallow Foundations

A remarkable number of methods for calculating settlement of shallow foundations on cohesionless soils are available in the literature. Some methods are simple and direct empirical correlations between field settlements and field tests such as the plate load, standard penetration, cone penetration, dilatometer and pressuremeter tests. Other methods apply the theory of elasticity with soil properties obtained from empirical correlations with in-situ or laboratory tests. In accordance with Poulos (1999), the available methods can be classified into three main categories (Table 4.1).

Table 4.1: Categories for classification of settlement methods (Poulos 1999)

Category	Characteristics	Method of parameter estimation
1	Empirical; not based on soil mechanics principles	Simple in-situ or laboratory tests, with correlations
2	Based on simplified theory or charts; uses soil mechanics principles; linear or non-linear elastic, rigid or elasto-plastic soil models	Routine; relevant in-situ or laboratory tests; may require some correlations
3	Based on theory using site-specific analysis; uses soil mechanics principles, linear or non-linear elastic, rigid plastic	Careful laboratory and/or in-situ tests which follow the appropriate stress paths

It is extremely difficult and expensive to obtain relatively undisturbed samples from cohesionless soils to represent the in-situ properties of such soils. As a result, methods that use the results of field tests are much more common than those based on laboratory tests. As previously mentioned in Chapter 1, as yet, no universally accepted method exists for predicting the settlement of shallow foundations on cohesionless soils. The purpose of this section is not to summarise all the available methods that are found in the literature, but, rather describe and discuss, in some detail, some of the most commonly used, and more relevant ones. Among these, three are selected: Meyerhof (1965); Schultze and Sherif (1973) and Schmertmann et al. (1978), for the following reasons:

- They are in common use;
- They represent the chronological development of settlement prediction;
- Each of them falls into one or more of the categories of settlement classification methods given in Table 4.1; and
- The database used later in this research contains most parameters required to calculate the settlement by these methods, which is necessary for the purposes of comparison with the techniques proposed in this thesis.

The finite element method (FEM) is one of the well-established techniques that might be used for settlement prediction, as it has been used successfully for solving many problems in the field of geotechnical engineering. FEM has the advantage of dealing with complicated geometry and boundary conditions and non-linear stress-strain behaviour of soil (Poulos 1975). However, it is not necessarily the case when it comes to the settlement prediction of shallow foundations on cohesionless soils. The reason is that the key to success of FEM lies mainly in an appropriate evaluation of the stress-strain behaviour of soil (Poulos 1999), which is a difficult task in the case of cohesionless soils (Moorhouse 1972). With regard to applying FEM for settlement prediction of shallow foundation on cohesionless soils, Poulos (1999) stated that "*...the more complex finite element methods appear to require far more development before being able to be used with confidence*". In addition, this method requires a greater amount of soil data, as well as data that are costly to measure. Consequently, FEM will not be used as a basis of comparison with ANNs in this research, primarily because sufficient data are unavailable.

4.4.1 Meyerhof's Method

Meyerhof (1965) suggested the following equations for estimating the settlement, S_c, of shallow foundations on sand:

$$S_c = \frac{8q}{N} \qquad \text{for } B \leq 4 \text{ ft (1.2 m)} \tag{4.4}$$

or

$$S_c = \frac{12q}{N}\left(\frac{B}{B+1}\right)^2 \qquad \text{for } B > 4 \text{ ft (1.2 m)} \tag{4.5}$$

where:

S_c = calculated settlement (inch);

q = footing net applied pressure (ton/ft²);

B = footing width (ft); and

N = average SPT blow count to a depth of $2B$ below the foundation level.

There is no overburden pressure correction for the average blow count, N, except for dense submerged silty sand when the minimum average blow count exceeds 15. The corrected, $N_{corrected}$, is given as:

$$N_{corrected} = 15 + 0.5(N - 15) \tag{4.6}$$

Meyerhof also suggested no correction be made for the effect of the ground water table, as its influence would be implicitly incorporated into the measured SPT results.

Meyerhof's method is simple to implement and has been one of the most popular methods for calculating settlement of shallow foundations based on SPT data (Coduto 1994). However, the method tends to be conservative, as it overestimates the settlement about 75% of the time (Coduto 1994).

4.4.2 Schultze and Sherif's Method

Shultze and Sherif (1973) proposed a procedure for settlement estimation based on elastic theory. By applying a statistical analysis to settlements obtained from 48 commercial and industrial structures, Schultze and Sherif (1973) proposed the following equation:

$$S_c = \frac{qF}{N^{0.87}(1+0.4D_f/B)} \tag{4.7}$$

where:

S_c = calculated settlement (cm);

q = footing net applied pressure (kg/cm^2);

F = settlement coefficient (obtained from a chart);

N = average SPT blow count to a depth of $2B$ or thickness of the compressible layer, whichever is the lesser;

D_f = depth of footing embedment (cm); and

B = footing width (cm);

Schultze and Sherif's method relies mainly on elastic theory. The limitation concerning methods involving elastic theory is the difficulty of evaluating the in-situ stress-strain properties and that, in most cases, settlement is non-linear (Moorhouse 1972).

4.4.3 Schmertmann's Method

Schmertmann (1970) proposed a procedure for estimating the settlement of foundations on granular soils, which was later updated by Schmertmann et al. (1978). The procedure is based on the theory of elasticity, finite element analyses, observations from field measurements and laboratory model studies. The method proposed that the calculated settlement, S_c, at the surface of a profile for a granular mass, can be expressed in terms of the vertical strain, ε_z, as follows:

$$S_c = \int_{z=0}^{\infty} \varepsilon_z dz \tag{4.8}$$

The profile can be considered as consisting of a series of homogeneous sub-layers with approximately constant values of cone resistance and N, and the settlement, S_c, can be computed as:

$$S_c = C_1 C_2 q \sum_{i=1}^{n} \sum_{i=1}^{n} (\frac{I_z}{E_s})_i \Delta z_i \tag{4.9}$$

in which:

S_c = calculated settlement (m);

q = footing net applied pressure (kPa);

I_z = strain influence factor (obtained from a chart);

E_s = Young's modulus at the middle of the ith layer of thickness Δz_i (kPa);

Δz_i = thickness of the ith layer (m); and

C_1, C_2 = correction factors for embedment and creep.

A chart was developed to obtain I_z, and correlations were suggested to obtain E_s from the results of CPT and/or SPT tests. Schmertmann (1978) also suggested two correction factors, C_1, C_2, which account for the effect of strain relief due to embedment and the effect of time-dependence or creep, respectively. The chart, correction factors and other details are given by Schmertmann (1970) and Schmertmann et al. (1978).

Schmertmann's method provides a more reliable method of estimating settlement of shallow foundations on granular soils (Moorhouse 1972). This method is also popular, useful and more precise than other methods (Coduto 1994). The reliability and accuracy of these three methods will be examined later in this thesis and compared with the model based on ANNs.

4.5 Summary

It has been argued that there are three main factors affecting settlement of shallow foundations on cohesionless soils that have more than marginal effects. These are the footing width, B, footing net applied pressure, q, and soil compressibility, which can be reflected by the average blow count from the standard penetration test (SPT). It has also been shown that other factors affecting settlement, including depth of water table, time dependence, footing geometry, L/B, and depth of footing embedment, D_f, are secondary compared with the three main factors illustrated above. In the following chapter, the use of multi-layer perceptrons (MLPs) that are trained with the back-propagation algorithm will be examined for predicting settlement of shallow foundations on cohesionless soils.

Chapter 5

Settlement Prediction by Multi-layer Perceptrons

5.1 Introduction

Over the years, many methods have been developed to predict the settlement of shallow foundations on cohesionless soils, as discussed in Chapter 4. However, methods for making such predictions with the required degree of accuracy and consistency have yet to be developed. Accurate prediction of settlement is essential since settlement, rather than bearing capacity, generally controls foundation design. In this chapter, artificial neural networks (ANNs) are used in an attempt to obtain more accurate settlement prediction. A large database of actual measured settlements is used to develop and verify the ANN models. As the prediction of settlement of shallow foundations on cohesionless soils does not involve any time-related parameter components, feed-forward multi-layer perceptrons (MLPs) are used. Feed-forward MLPs that are trained with the back-propagation algorithm are the most commonly used neural network type (Maren et al. 1990), as they have a high capability of data mapping (Hecht-Nielsen 1990). MLPs trained with the back-propagation algorithm have been applied successfully to many geotechnical engineering problems (e.g. Goh 1994a, b; Najjar and Basheer 1996a, b), and are thus used in this work.

The objectives of this chapter are to:

1. Investigate the feasibility of the ANN technique for predicting the settlement of shallow foundations on cohesionless soils.
2. Introduce a method of model validation that tests the robustness of the predictive ability of ANN models.
3. Study the effect of ANN geometry and internal parameters on the performance of ANN models.
4. Investigate the relationship between the statistical properties of the data subsets used to develop ANN models and model performance.
5. Investigate the relationship between the proportion of the data in each of the subsets used to develop ANN models and model performance.

6. Investigate the use of different approaches for dividing the available data into the subsets needed to develop ANN models and introduce and evaluate a new approach for data division that is based on fuzzy clustering.
7. Examine the effect of data transformation on the performance of ANN models.
8. Provide information on the relative importance of the factors affecting the settlement of shallow foundations on cohesionless soils.
9. Compare the performance of ANNs with some of the most commonly used traditional methods.
10. Provide a practical equation and a series of design charts for settlement prediction of shallow foundations on cohesionless soils from the developed ANN model for routine use in practice.

5.2 Development of ANN Models

The data used to calibrate and validate the neural network models are obtained from the literature, and include field measurements of settlement of shallow foundations as well as the corresponding information regarding the footings and soil. The data cover a wide range of footing dimensions and cohesionless soil types and properties. The database comprises a total of 189 individual cases; 125 cases were reported by Burland and Burbidge (1985), 22 cases by Burbidge (1982), 5 cases by Bazaraa (1967) and 30 cases by Wahls (1997). Another 4 cases are given by Briaud and Gibbens (1999), one case by Picornell and Del Monte (1988) and 2 cases by Maugeri et al. (1998). Full details of the database are given in Appendix A.

The steps for developing ANN models, as outlined by Maier and Dandy (2000) and given in §2.5, are used as a guide in this work. These include the determination of model inputs and outputs, division and pre-processing of the available data, the determination of appropriate network architecture, optimisation of the connection weights (training) and model validation. Two PC-based commercial software systems are used to simulate neural network operation. The first is NeuralWorks *Predict* Release 2.1 (NeuralWare 1997), which adopts the Cascade-Correlation algorithm to automatically determine optimal network architecture. The second is *Neuframe* Version 4.0 (Neusciences 2000), in which optimal network architecture is determined by trial-and-error. ANN models that are developed using the aforementioned software systems are described in detail below.

5.2.1 ANN Models Developed Using *Predict*

- ***Model Inputs and Outputs***

As discussed in Chapter 4, it is generally accepted that five parameters have the most significant impact on the settlement of shallow foundations on cohesionless soils (Burland and Burbidge 1985), and are thus used as the ANN model inputs. These include footing width (B), footing net applied pressure (q), the average SPT blow count (N) over the depth of influence of the foundation, footing geometry (L/B) and footing embedment ratio (D_f/B). The model output is the average measured settlement (S_m) of the foundation, considered in its final state. A sensitivity analysis to investigate the relative importance of the ANN model inputs will be explored in §5.5.

- ***Data Division***

The next step in the development of ANN models is dividing the available data into their subsets. For reasons given in Chapter 2, cross-validation (Stone 1974) is used as the stopping criteria in this research. Consequently, the data are randomly divided into three sets: training, testing and validation, as is standard practice in the development of ANN models in geotechnical engineering. In total, 80% of the data are used for training and 20% are used for validation. The training data are further divided into 70% for the training set and 30% for the testing set. For reasons discussed in §2.5.2, the training, testing and validation sets are also 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 tried until three statistically consistent data sets are obtained. The statistics of the training, testing and validation sets are shown in Table 5.1. The statistical parameters considered include the mean, standard deviation, minimum, maximum and range. The relationship between the statistical properties of the data subsets, and the effect of the proportion of the data used in each subset on model performance will be investigated in §5.3. The use of different available data division methods for ANN models will also be investigated in §5.3 and a new approach that is based on fuzzy clustering will also be introduced and evaluated in this section.

Table 5.1: Input and output statistics for the ANN models

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Footing width, *B* (m)					
Training set	8.3	9.8	0.8	60.0	59.2
Testing set	9.3	10.9	0.9	55.0	54.1
Validation set	9.4	10.1	0.9	41.2	40.3
Footing net applied pressure, *q* (kPa)					
Training set	188.4	129.0	18.3	697.0	678.7
Testing set	183.2	118.7	25.0	584.0	559.0
Validation set	187.9	114.6	33.0	575.0	542.0
Average SPT blow count, *N*					
Training set	24.6	13.6	4.0	60.0	56.0
Testing set	24.6	12.9	5.0	60.0	55.0
Validation set	24.3	14.1	4.0	55.0	51.0
Footing geometry, *L/B*					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.3	1.9	1.0	9.9	8.9
Validation set	2.1	1.8	1.0	8.0	7.0
Footing embedment ratio, D_f/B					
Training set	0.52	0.57	0.0	3.4	3.4
Testing set	0.49	0.52	0.0	3.0	3.0
Validation set	0.59	0.64	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training set	20.0	27.2	0.6	121.0	120.4
Testing set	21.4	26.6	1.0	120.0	119.0
Validation set	20.4	25.2	1.3	120.0	118.7

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 examines the null hypothesis of no difference in the means of two data sets and the *F*-test examines 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). Consequently, this level of significance is used in this research. This means that there is a confidence level of 95% that the training, testing and validation sets are statistically consistent. A detailed description of these tests is given by Levine et al. (1999). The results of the *t*- and *F*-tests are given in Table 5.2. These results indicate that the training, testing and validation sets are generally representative of a single population.

Table 5.2: Null hypothesis tests for the ANN input and output variables

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.58	–1.97	1.97	Accept	0.81	0.59	1.87	Accept
Validation	–0.61	–1.97	1.97	Accept	0.94	0.61	1.77	Accept
q								
Testing	0.23	–1.97	1.97	Accept	1.18	0.59	1.87	Accept
Validation	0.02	–1.97	1.97	Accept	1.27	0.61	1.77	Accept
N								
Testing	0.00	–1.97	1.97	Accept	1.11	0.59	1.87	Accept
Validation	0.11	–1.97	1.97	Accept	0.93	0.61	1.77	Accept
L/B								
Testing	–0.64	–1.97	1.97	Accept	0.80	0.59	1.87	Accept
Validation	0.00	–1.97	1.97	Accept	0.89	0.61	1.77	Accept
D_f/B								
Testing	0.31	–1.97	1.97	Accept	1.20	0.59	1.87	Accept
Validation	–0.62	–1.97	1.97	Accept	0.79	0.61	1.77	Accept
S_m								
Testing	–0.29	–1.97	1.97	Accept	1.05	0.59	1.87	Accept
Validation	–0.08	–1.97	1.97	Accept	1.17	0.61	1.77	Accept

- ***Pre-processing of Data***

Once the available data have been divided into their subsets, the input and output variables are pre-processed by scaling them 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 (i.e. −1.0 to 1.0 for tanh transfer function and 0.0 to 1.0 for sigmoid transfer function). The simple linear mapping of the variables' practical extremes to the neural network's practical extremes is adopted for scaling, as it is the most commonly used method (Masters 1993). As part of this method, 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:

$$x_n = (x - x_{\min})/(x_{\max} - x_{\min}) \tag{5.1}$$

Transformation of the input data as a way of improving the performance of ANN models will be examined in §5.4.

- ***Model Architecture, Optimisation and Stopping Criteria***

As discussed in Chapter 2, one of the most important and difficult tasks in the development of ANN models is determining the model architecture (i.e. the number and connectivity of the hidden layer nodes). In order to obtain the optimum number of hidden layer nodes, it is important to strike a balance between having sufficient free parameters (weights) to enable representation of the function to be approximated, and not having too many so as to avoid overtraining. Overtraining is not an issue in this study, as cross-validation is used as the stopping criterion. However, physical interpretation of the model is important, and hence the smallest network that is able to map the desired relationship should be used. In NeuralWorks *Predict*, the optimal network architecture is found automatically with the aid of the *Cascade-Correlation* algorithm (see §2.5.4). The process of optimising the connection weights is applied using the default parameters of the software and are given as follows:

- Learning rate: 100 for the hidden layer and 0.01 for the output layer; and
- Transfer function: tanh for the hidden and sigmoid for the output layer.

As *Predict* uses an automatic method for finding the optimal network architecture, the final network architecture obtained might be different for different models, even though the same data sets and learning parameters have been used. As a result, studying the effect of changing the learning parameters on the performance of ANN models could be misleading and was therefore not done when *Predict* was used. However, because *Neuframe* requires a manual process for determining optimal network architecture and internal parameters, the effect of varying ANN internal parameters and geometry on the performance of ANN models was investigated (§5.2.2).

- ***Model Validation***

Using the above method, a number of networks are developed and the structure and performance results of the developed models are shown in Table 5.3. A code is used in this chapter to identify the names of the different models developed. The code consists of two parts separated by a hyphen. The first part represents an abbreviation to the current chapter. The second part begins with an abbreviation that denotes the software used followed by the model number. Hence, for example, "CHP5-PD1" implies Chapter 5, Predict Model No. 1.

It can be seen from Table 5.3 that the performance of the developed models is quantified using three different measures; the coefficient of correlation (*r*), the root mean squared error (RMSE) and the mean absolute error (MAE), as discussed in §2.5.7. Three measures are used for assessing the models, including: (i) the performance of the model on the testing set; (ii) the number of hidden nodes and (iii) the general consistency of model performance on the validation set with those obtained on the training and testing sets. Consequently, an ANN model is deemed to be optimal if the model provides satisfactory performance on the testing set coupled with a small number of hidden nodes and consistent performance on the validation set with that obtained on the training and testing sets. It can be seen from Table 5.3 that model CHP5-PD14 performs well, as it has a high coefficient of correlation and low RMSE and MAE between the measured and predicted settlement on the testing set, coupled with a smaller number of hidden nodes and consistent performance on the training, testing and validation sets. The above results indicate that model CHP5-PD14 has the capability of predicting the settlement of shallow foundations on cohesionless soils with a high degree of accuracy and can thus be used as a practical tool for predictive purposes.

Table 5.3: Structure and performance of ANN models developed using *Predict*

Model No.	No. hidden nodes	Performance measures								
		Correlation coefficient, *r*			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-PD1	23	0.97	0.96	0.82	6.2	7.4	15.5	4.2	4.3	9.8
CHP5-PD2	31	0.97	0.97	0.80	5.6	6.5	16.4	4.1	4.5	10.4
CHP5-PD3	35	0.98	0.96	0.81	4.5	7.6	16.5	3.3	5.1	9.7
CHP5-PD4	17	0.97	0.95	0.87	5.6	8.3	13.1	4.0	5.5	9.8
CHP5-PD5	47	0.98	0.97	0.83	4.2	6.7	14.6	2.7	4.8	8.8
CHP5-PD6	42	0.98	0.97	0.84	4.4	6.1	14.0	3.1	3.9	9.0
CHP5-PD7	48	0.98	0.96	0.86	4.9	6.6	13.3	3.4	4.6	8.3
CHP5-PD8	20	0.98	0.96	0.86	3.9	7.4	14.0	2.4	4.9	8.2
CHP5-PD9	50	0.97	0.95	0.84	5.7	8.3	14.9	4.2	6.0	9.9
CHP5-PD10	38	0.98	0.96	0.83	5.4	6.6	14.7	4.0	4.4	9.9
CHP5-PD11	31	0.98	0.97	0.85	5.2	6.4	13.6	3.8	4.5	8.3
CHP5-PD12	40	0.98	0.96	0.84	4.9	6.9	14.2	3.6	5.1	9.7
CHP5-PD13	39	0.98	0.97	0.87	4.7	6.3	12.9	3.4	3.9	8.3
CHP5-PD14	11	0.92	0.92	0.88	11.5	11.5	12.0	7.3	7.5	9.5
CHP5-PD15	48	0.98	0.96	0.87	4.7	7.2	12.5	3.2	4.7	8.8
CHP5-PD16	44	0.98	0.96	0.86	5.0	7.1	12.8	3.6	5.3	8.9
CHP5-PD17	23	0.97	0.96	0.86	6.2	7.0	13.7	4.2	5.2	9.9

T = training, S = testing and V= validation

In order to confirm the generalisation ability and robustness of model CHP5-PD14, an additional validation approach is proposed. The approach suggests carrying out a parametric study in which the response of the ANN model output to changes in its inputs is investigated. All input variables, except one, are fixed to their mean values used for training and a set of synthetic data, between the minimum and maximum values used for model training, are generated for the input that is not set to a fixed value. The synthetic data are generated by increasing their values in increments equal to 5% of the total range between the minimum and maximum values. The response of the model is then examined. This process is repeated using another input variable and so on until the model response is tested for all input variables. The robustness of the model can be determined by examining how well the predicted settlements are in agreement with the known underlying physical process over a range of inputs.

The above approach is applied to model CHP5-PD14 and the results are shown in Figure 5.1. It can be seen that some of the results obtained are contrary to what one would expect based on the known physical behaviour of settlement of shallow foundations on cohesionless soils. For example, in Figure 5.1(a), one would expect that the predicted settlement would increase as footing width increases in a relatively consistent and smooth fashion. In Figure 5.1 (c), one would expect that the predicted settlement would decrease as the average SPT blow count increases. The model behaviour displayed in Figure 5.1 (d) is also unexpected, as there is no obvious trend in the relationship between predicted settlement and footing geometry. One would expect that the predicted settlement would increase with an increase in footing geometry. In addition, the odd shape of the curves is difficult to justify from a physical perspective.

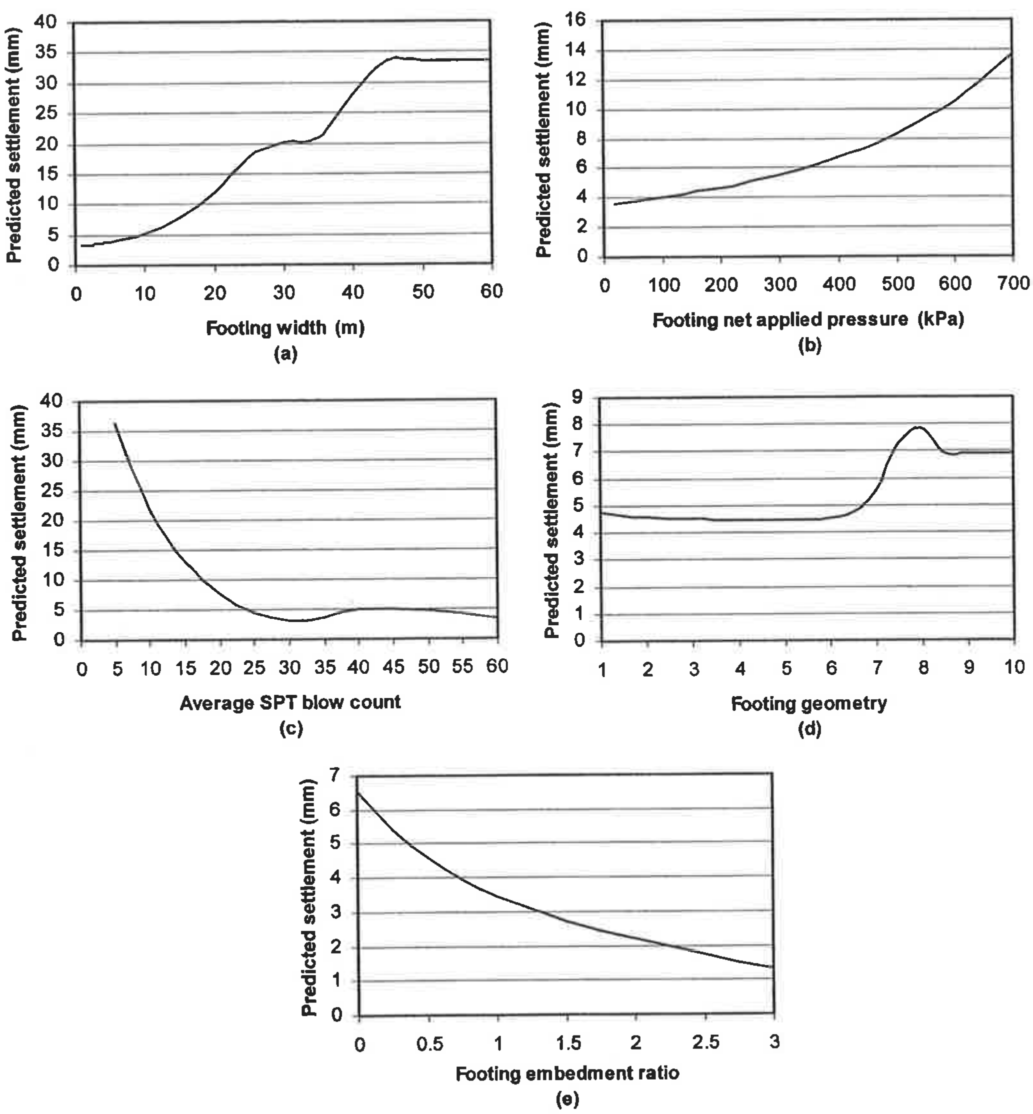


Figure 5.1: Results of parametric study for Model CHP5-PD14

The above results indicate that there may be some level of model overfitting, even though cross-validation was adopted and an independent validation set was used. Consequently, further work is needed to determine the cause of this behaviour. Possible factors affecting the generalisation ability and robustness of ANN models include the number and type of connection weights, the degree of noise in the data and the software implementation used. These are investigated and discussed below.

- ***Effect of Number and Type of Connection Weights***

As mentioned in Chapter 2, one of the difficulties in using ANN models is that the potential number of free model parameters (i.e. connection weights) is generally large and there is therefore a danger of overfitting the training data. In other words, if the number of degrees of freedom of the model is large compared with the number of data points used for training, the model might no longer fit the general trend, as desired, but might learn the idiosyncrasies of the particular data points used for training. In general, one of two methods is used to overcome this problem. The first is to restrict the ratio of the number of connection weights to the number of data points in the training set, and several rules-of-thumb have been given in Chapter 2 as a guide. The second approach to avoiding overfitting is to use a non-convergent method (Finnhoff et al. 1993) in which training is stopped early once the error in an independent test set starts to increase. This method is described in Chapter 2 and is commonly called cross-validation (Stone 1974).

As mentioned above, if the ratio of the number of connection weights to the number of data samples in the training set is too large, the model might no longer be able to fit the desired trends. This makes it difficult to interpret the physical meaning of the relationship found by ANN models. In order to investigate the impact of reducing the number of connection weights and thus reducing the ratio of the number of connection weights to the number of data samples, more networks are trained eliminating the direct connections between the input and output nodes but allowing the cascaded connections between the new hidden nodes and previously established ones (see §2.5.4). The structure and performance results of the developed models are given in Table 5.4. It can be seen that model CHP5-PD26 with 4 hidden layer nodes gives as good performance as model CHP5-PD14, which has 11 hidden layer nodes. The result of the parametric study for model CHP5-PD26 is given in Figure 5.2. As model CHP5-PD26 performs well on an independent validation set (Table 5.4), one would expect that the model is adequately trained so as to be used for predictive purposes. However, Figure 5.2 shows that, although some improvement of model robustness has been achieved compared with model CHP5-PD14 (see Figure 5.1), reducing the number of connection weights does not completely remedy the problem. There are still some unexpected deviations in the trends that relate predicted settlement to the footing width, average SPT blow count and footing geometry, as shown in Figures 5.2 (a), (c) and (d), respectively.

Table 5.4: Structure and performance of ANN models where direct connections between the input and output nodes are prohibited and cascaded connections between hidden nodes are permitted

Model No.	No. hidden Nodes	Performance measures								
		Correlation coefficient, *r*			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-PD18	14	0.98	0.96	0.76	5.3	7.6	19.0	4.0	5.4	11.0
CHP5-PD19	14	0.97	0.97	0.82	5.7	6.5	16.6	4.2	4.7	10.4
CHP5-PD20	46	0.97	0.96	0.85	5.5	6.7	14.1	4.1	4.9	9.8
CHP5-PD21	26	0.99	0.96	0.84	3.8	6.9	14.2	2.4	5.0	9.2
CHP5-PD22	45	0.99	0.96	0.83	3.6	6.9	16.0	2.5	4.8	10.8
CHP5-PD23	28	0.98	0.96	0.85	5.2	7.5	13.7	2.7	5.2	9.1
CHP5-PD24	11	0.94	0.94	0.91	10.4	10.5	11.8	7.4	7.5	9.4
CHP5-PD25	25	0.97	0.95	0.85	5.8	7.9	13.6	3.8	4.9	8.8
CHP5-PD26	4	0.90	0.93	0.87	12.1	9.9	12.7	7.7	6.1	9.4
CHP5-PD27	9	0.97	0.96	0.92	6.3	8.0	10.2	4.4	5.8	7.7
CHP5-PD28	26	0.98	0.96	0.86	5.4	6.9	13.2	3.9	5.0	8.8
CHP5-PD29	5	0.87	0.89	0.79	13.4	11.9	15.9	8.6	8.2	11.9
CHP5-PD30	5	0.91	0.92	0.85	11.2	9.9	13.2	7.0	6.0	10.2

T = training, S = testing and V= validation

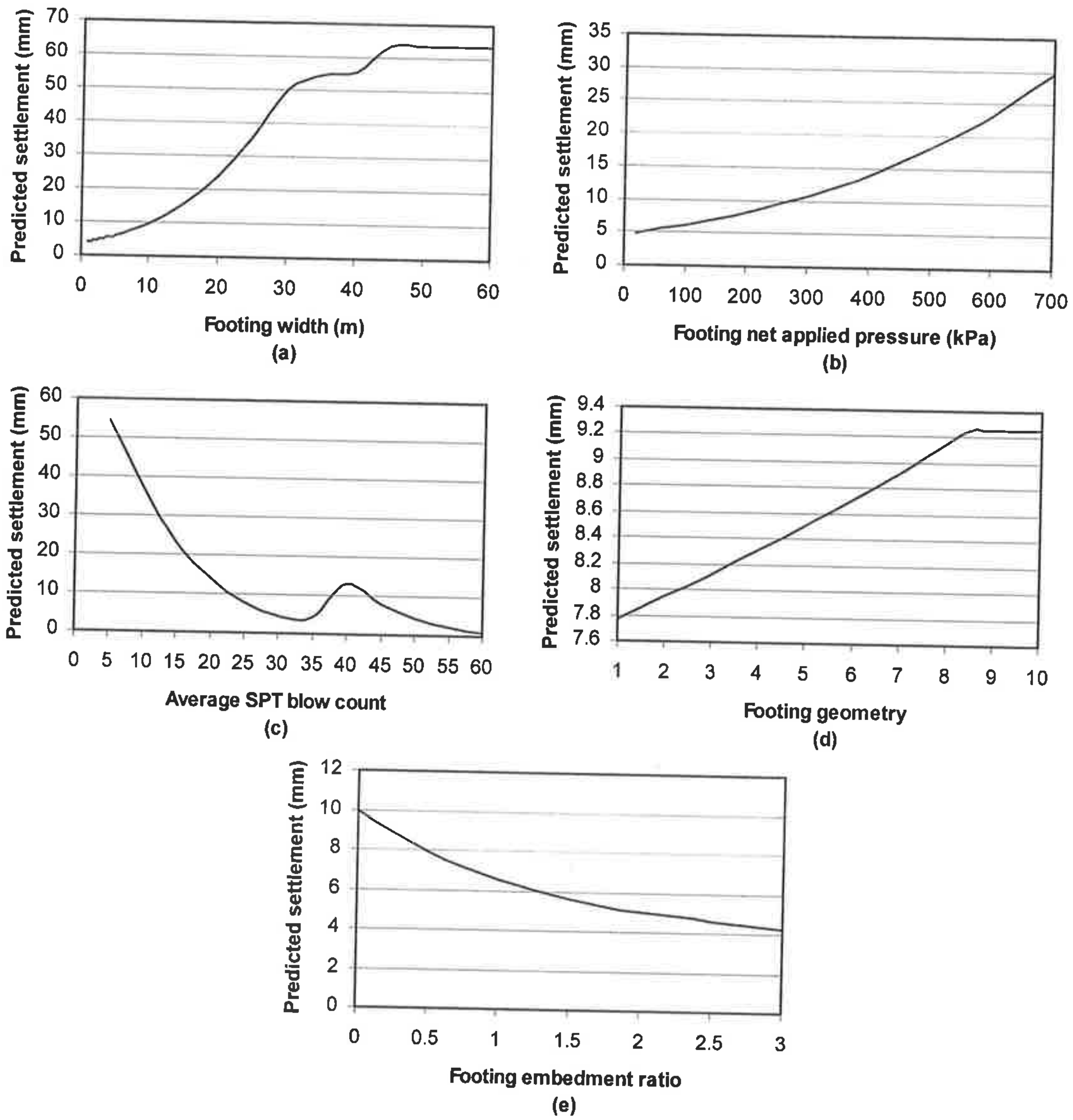


Figure 5.2: Results of parametric study for Model CHP5-PD26

In an attempt to reduce the number of connection weights even further, both the direct connections between the input and output nodes and cascaded connections between hidden nodes are removed. Consequently, only the direct connections between the input layer nodes to the newly added hidden layer nodes and between hidden nodes to the output layer nodes are allowed. By following this procedure, a number of networks are developed and the structure and performance results of the developed models are given in Table 5.5. It can be seen that model CHP5-PD43, with 4 hidden layer nodes, performs well and its performance, without both the direct connections between the input and output nodes and cascaded connections between hidden nodes, is similar to previously developed models (i.e. models CHP5-PD14 and CHP5-PD26). It should be noted that the ratio of the number of connection weights to the number of training samples for model CHP5-PD43, is approximately 1:5. This ratio is in accordance with most ratios recommended by researchers in order to guarantee that overfitting does not occur (Masters 1993; Rogers and Dowla 1994). The result of the parametric study for model CHP5-PD43 is shown in Figure 5.3. It can be seen that, in Figures 5.3 (a), (c) and (d), respectively, some unexpected trends are still obtained between the predicted settlement and footing width, average SPT blow count and footing geometry, even though the number of connection weights has been reduced further. Consequently, it can be concluded that eliminating the cascaded connections does not improve model robustness and that the model behaviour shown in Figure 5.1 does not appear to be affected by the number of connection weights nor the type of connection between nodes.

Table 5.5: Structure and performance of ANN models where both direct connections between the input and output nodes and cascaded connections between hidden nodes are prohibited

Model No.	No. hidden Nodes	Performance measures								
		Correlation coefficient, *r*			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-PD31	4	0.89	0.91	0.85	12.0	10.5	13.2	7.7	6.7	9.9
CHP5-PD32	15	0.96	0.96	0.84	7.4	6.8	14.4	5.1	4.6	9.9
CHP5-PD33	13	0.91	0.96	0.86	11.3	7.6	13.1	7.1	4.8	9.4
CHP5-PD34	4	0.89	0.91	0.85	12.3	10.7	13.6	8.0	7.0	10.3
CHP5-PD35	5	0.91	0.92	0.86	11.0	10.1	12.9	7.0	6.2	10.4
CHP5-PD36	8	0.94	0.94	0.88	8.9	9.0	12.5	6.0	6.3	9.4
CHP5-PD37	15	0.94	0.95	0.81	8.9	8.6	15.3	5.9	6.3	10.5
CHP5-PD38	4	0.94	0.92	0.84	8.9	10.4	14.6	5.7	7.5	11.5
CHP5-PD39	5	0.95	0.92	0.89	8.0	10.2	11.7	5.5	5.8	8.7
CHP5-PD40	14	0.97	0.95	0.84	6.1	7.7	14.4	4.4	5.2	10.0
CHP5-PD41	5	0.93	0.93	0.88	9.4	9.4	11.8	6.3	5.7	8.7
CHP5-PD42	20	0.97	0.96	0.81	6.2	6.9	15.6	4.5	4.8	9.9
CHP5-PD43	4	0.91	0.92	0.86	11.6	10.3	13.2	7.3	6.7	9.7
CHP5-PD44	30	0.97	0.96	0.86	6.1	6.6	13.7	4.4	5.0	9.2
CHP5-PD45	5	0.89	0.91	0.85	11.9	10.5	13.6	7.6	6.7	10.4
CHP5-PD46	4	0.87	0.91	0.75	13.2	10.6	17.6	8.6	7.3	12.5
CHP5-PD47	20	0.98	0.97	0.85	4.9	6.7	14.1	3.5	4.5	9.5
CHP5-PD48	4	0.92	0.91	0.88	10.2	10.5	12.1	6.7	6.4	9.5

T = training, S = testing and V= validation

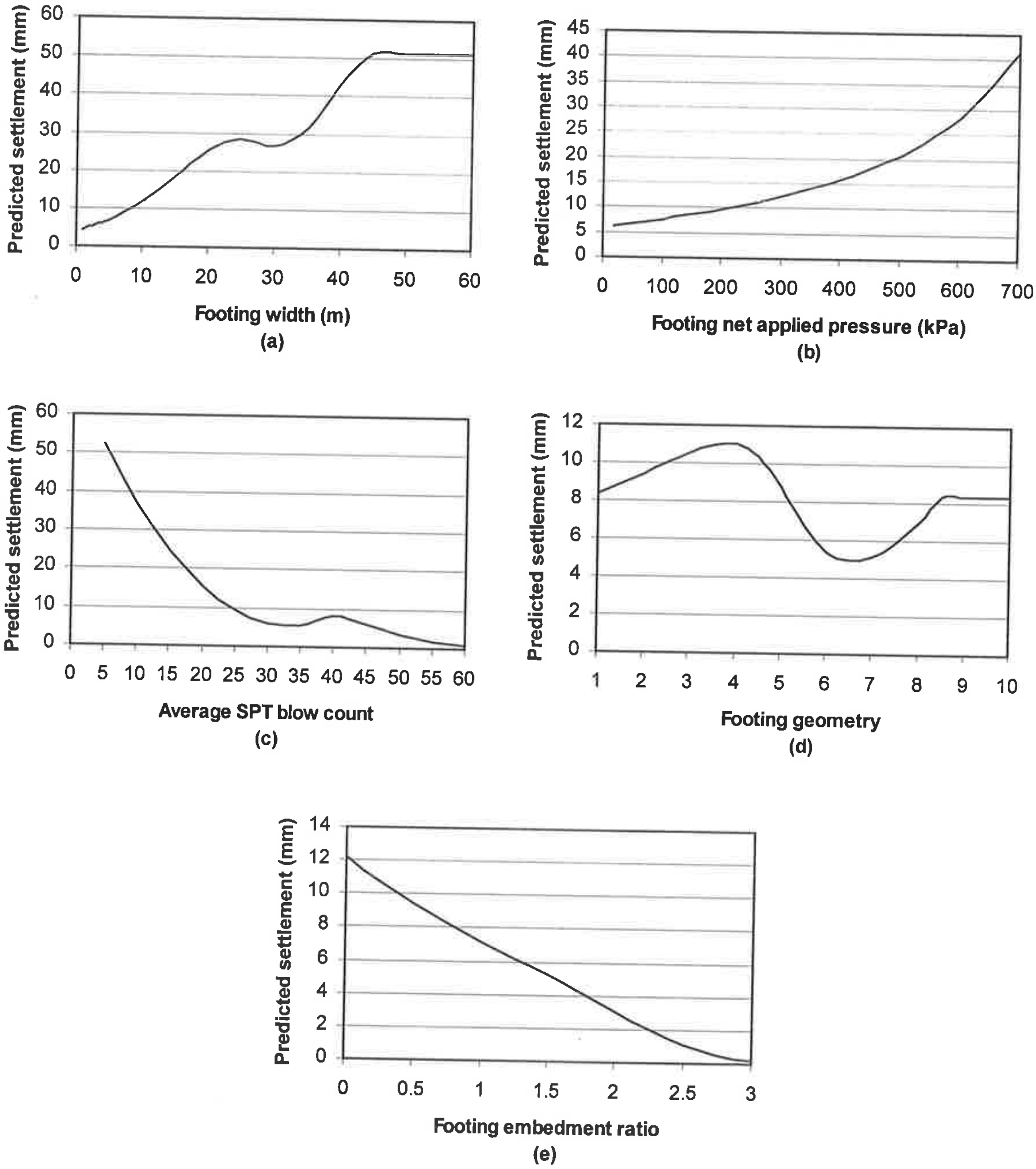


Figure 5.3: Results of parametric study for Model CHP5-PD43

- ***Effect of Data Noise***

In order to investigate the effect of data noise on the robustness of ANN models, a clean set of 500 data samples (see Appendix B) consisting of different combinations of B, q and N, is generated randomly from a uniform distribution. The corresponding settlements are calculated using Meyerhof's equation for settlement prediction as follows (Meyerhof 1965):

$$S_c = \frac{2q}{N}\left(\frac{B}{B+0.3}\right)^2 \tag{5.2}$$

where S_c = calculated settlement (mm), q = footing net applied pressure (kPa), N = average SPT blow count and B = footing width (m). The ANN models are developed using B, q and N as input variables and the calculated settlement from Equation 5.2 as single output variable. The data are divided randomly into 300 cases for training, 100 for testing and 100 for validation and selected to be representative of the same statistical population. The statistics and null hypothesis tests for the training, testing and validation sets are shown in Tables 5.6 and 5.7, respectively. It can be seen that the training, testing and validations sets are statistically consistent and thus representative of the same statistical population.

Table 5.6: Input and output statistics for ANN models of synthetic clean data

Model variables and data sets	Statistical parameters				
	Mean	**Std. Dev.**	**Minimum**	**Maximum**	**Range**
Footing width, B (m)					
Training set	26.1	14.5	0.55	49.9	49.4
Testing set	26.7	13.2	2.2	49.5	47.3
Validation set	28.4	13.6	2.1	49.6	47.5
Footing net applied pressure, q (kPa)					
Training set	393.9	233.2	5.1	799.4	794.3
Testing set	416.7	242.2	14.9	795.7	780.8
Validation set	401.6	224.1	14.4	795.6	781.2
Average SPT blow count, N					
Training set	31.7	16.5	5.1	59.8	54.7
Testing set	31.8	14.8	5.3	58.8	53.5
Validation set	31.0	15.8	5.2	59.3	54.1
Measured settlement, S_m (mm)					
Training set	36.9	44.9	0.11	267.9	267.8
Testing set	37.5	43.1	0.76	253.1	252.3
Validation set	39.4	47.6	0.83	252.1	251.3

Table 5.7: Null hypothesis tests for ANN input and output variables of synthetic clean data

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.37	–1.97	1.97	Accept	1.21	0.73	1.39	Accept
Validation	–1.39	–1.97	1.97	Accept	1.14	0.73	1.39	Accept
q								
Testing	–0.84	–1.97	1.97	Accept	0.93	0.73	1.39	Accept
Validation	–0.29	–1.97	1.97	Accept	1.08	0.73	1.39	Accept
N								
Testing	–0.05	–1.97	1.97	Accept	1.24	0.73	1.39	Accept
Validation	0.37	–1.97	1.97	Accept	1.09	0.73	1.39	Accept
S_m								
Testing	–0.12	–1.97	1.97	Accept	1.08	0.73	1.39	Accept
Validation	–0.47	–1.97	1.97	Accept	0.89	0.73	1.39	Accept

A number of networks are developed using the clean data and the structure and performance results of the developed models are shown in Table 5.8. As with the models developed previously using the actual data, a code is used to identify the names of the different models developed using the synthetic clean data. The code consists of three parts separated by two hyphens. The first part represents an abbreviation to the current chapter, the second represents the first letters of the term "synthetic clean data" and the third part represents the model number that uses *Predict* software. Hence, for example, "CHP5-SCD-PD1" implies Chapter 5, synthetic clean data, Predict Model No. 1. It should be noted that models CHP5-SCD-PD1 to CHP5-SCD-PD10 are trained with the direct connections between the input and output nodes and cascaded connections between hidden nodes, whereas models CHP5-SCD-PD11 to CHP5-SCD-PD20 are trained without these connections. It can be seen from Table 5.8 that model CHP5-SCD-PD6 with 7 hidden layer nodes and model CHP5-SCD-PD17 with 8 hidden layer nodes perform well.

Table 5.8: Structure and performance of ANN models developed using *Predict* and synthetic clean data

Model No.	No. hidden nodes	Performance measures								
		Correlation coefficient, *r*			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-SCD-PD1	14	0.98	0.97	0.98	9.3	9.7	11.9	5.6	5.5	6.9
CHP5-SCD-PD2	18	0.99	0.99	0.99	4.6	3.4	4.9	2.3	1.9	2.3
CHP5-SCD-PD3	11	0.99	0.99	0.99	5.5	5.4	5.5	3.3	3.4	3.5
CHP5-SCD-PD4	10	0.98	0.97	0.98	6.5	9.8	7.9	4.0	5.0	4.6
CHP5-SCD-PD5	13	0.99	0.99	0.99	3.3	3.5	2.9	2.0	2.0	1.9
CHP5-SCD-PD6	7	0.99	0.99	0.99	3.6	3.5	3.3	2.1	2.1	2.0
CHP5-SCD-PD7	15	0.99	0.99	0.99	3.6	2.8	3.1	2.5	2.1	2.3
CHP5-SCD-PD8	9	0.99	0.99	0.99	5.2	3.9	5.6	3.5	2.8	3.7
CHP5-SCD-PD9	12	0.99	0.99	0.99	4.4	3.5	5.9	2.7	2.3	3.3
CHP5-SCD-PD10	15	0.99	0.99	0.99	4.1	3.8	3.7	2.6	2.5	2.6
CHP5-SCD-PD11	12	0.99	0.99	0.99	5.9	5.4	6.9	3.2	2.8	3.3
CHP5-SCD-PD12	15	0.99	0.99	0.99	4.7	3.9	4.4	3.1	2.4	3.0
CHP5-SCD-PD13	18	0.99	0.99	0.99	5.9	7.6	6.7	3.4	4.0	3.8
CHP5-SCD-PD14	12	0.99	0.99	0.99	4.4	4.4	4.3	3.1	3.2	2.9
CHP5-SCD-PD15	14	0.99	0.99	0.99	4.2	3.4	3.3	2.7	2.2	2.3
CHP5-SCD-PD16	9	0.99	0.99	0.99	5.5	4.7	5.5	3.7	3.2	3.6
CHP5-SCD-PD17	8	0.99	0.99	0.99	3.4	3.6	4.2	1.9	2.1	1.9
CHP5-SCD-PD18	10	0.99	0.99	0.99	5.9	6.9	7.2	3.5	3.9	4.0
CHP5-SCD-PD19	14	0.99	0.99	0.99	4.9	4.6	5.2	3.3	3.3	3.5
CHP5-SCD-PD20	8	0.99	0.99	0.99	5.8	4.2	5.1	3.7	3.1	3.3

T = training, S = testing and V= validation

The results of the parametric studies for models CHP5-SCD-PD6 and CHP5-SCD-PD17 are shown in Figure 5.4, which also includes predictions from Meyerhof's equation (Equation 5.2). It can be seen that both models (i.e. models CHP5-SCD-PD6 and CHP5-SCD-PD17) are able to produce similar predictions to those of Meyerhof's equation and both models are able to reflect the underlying physical process of settlement prediction and thus may be considered to be robust. However, the robustness of these models is based on a set of clean data that cannot be replicated in reality.

In order to simulate real-world conditions, some noise is added to the settlements calculated from Equation 5.2 for all 500 data samples. The noise is derived from a normal distribution that has a mean value equal to zero and a standard deviation equal to 20% of the calculated settlement. Details of the noisy data are given in Appendix B and the statistics and null hypothesis tests for the input and output variables are given in Tables 5.9 and 5.10, respectively.

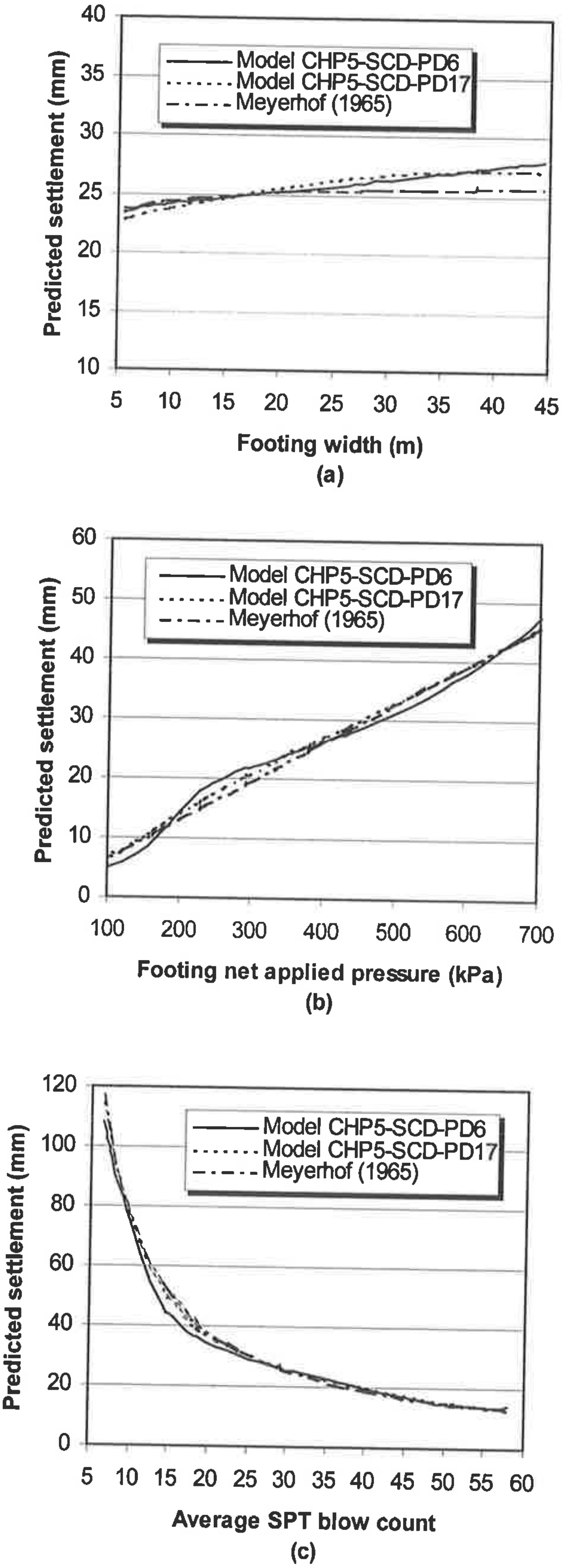


Figure 5.4: Results of parametric study for Models CHP5-SCD-PD6 and CHP5-SCD-PD17

Table 5.9: Input and output statistics for ANN models of synthetic noisy data

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Footing width, *B* (m)					
Training set	26.1	14.5	0.56	49.9	49.3
Testing set	26.7	13.2	2.2	49.5	47.3
Validation set	28.4	13.6	2.1	49.6	47.5
Footing net applied pressure, *q* (kPa)					
Training set	393.9	233.2	5.1	799.4	794.3
Testing set	416.7	242.2	14.9	795.7	780.8
Validation set	401.6	224.1	14.4	795.6	781.2
Average SPT blow count, *N*					
Training set	31.7	16.5	5.1	59.8	54.7
Testing set	31.8	14.8	5.3	58.8	53.5
Validation set	31.0	15.8	5.2	59.3	54.1
Measured settlement, S_m (mm)					
Training set	36.8	47.8	0.12	430.3	430.2
Testing set	37.9	51.5	0.56	285.3	284.7
Validation set	42.2	58.5	0.45	361.1	360.7

Table 5.10: Null hypothesis tests for the ANN input and output variables of synthetic noisy data

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.37	–1.97	1.97	Accept	1.20	0.73	1.39	Accept
Validation	–1.39	–1.97	1.97	Accept	1.14	0.73	1.39	Accept
q								
Testing	–0.84	–1.97	1.97	Accept	0.93	0.73	1.39	Accept
Validation	–0.29	–1.97	1.97	Accept	1.10	0.73	1.39	Accept
N								
Testing	–0.05	–1.97	1.97	Accept	1.24	0.73	1.39	Accept
Validation	0.37	–1.97	1.97	Accept	1.09	0.73	1.39	Accept
S_m								
Testing	–0.19	–1.97	1.97	Accept	0.86	0.73	1.39	Accept
Validation	–0.92	–1.97	1.97	Accept	0.67	0.73	1.39	Reject

Training is repeated and a number of models are developed. The structure and performance results of the developed models are shown in Table 5.11. Again, a code similar to that used previously for identifying the names of the different models of synthetic clean data is also utilised except that the middle part of the previous code (SCD) is replaced by (SND), which denotes the term “synthetic noisy data”. It should be noted that models CHP5-SND-PD1 to CHP5-SND-PD10 are trained with direct connections between the input and output nodes and cascaded connections between hidden nodes, whereas models CHP5-SND-PD11 to CHP5-SND-PD20 are trained without these connections. It can be seen from Table 5.11 that there are two models that perform well with respect to the training, testing and validation data (i.e. models CHP5-SND-PD4 and CHP5-SND-PD14), with 6 and 7 hidden nodes, respectively. The results of the parametric studies for models CHP5-SND-PD4 and CHP5-SND-PD14 are shown in Figure 5.5. It can be seen that the behaviour of models CHP5-SND-PD4 and CHP5-SND-PD14 is not robust, as some abnormal trends are obtained regardless of the type of connection used. This tends to suggest that the behaviour shown in Figure 5.1 may be due to the presence of noise in the data. However, the unexpected behaviour of the ANN models trained with the noisy hypothetical data shown in Figure 5.5 may also be due to the software used, the impact of which is investigated in the next section.

Table 5.11: Structure and performance of ANN models developed using *Predict* for synthetic noisy data

Model No.	No. hidden nodes	Performance measures								
		Correlation coefficient, *r*			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-SND-PD1	16	0.87	0.88	0.89	22.9	23.8	29.2	10.5	12.9	13.2
CHP5-SND-PD2	12	0.87	0.88	0.87	22.9	24.4	29.7	10.9	13.3	13.7
CHP5-SND-PD3	12	0.88	0.90	0.90	23.1	23.0	28.0	10.5	12.2	12.9
CHP5-SND-PD4	6	0.87	0.88	0.89	23.5	24.1	29.3	10.7	12.5	13.1
CHP5-SND-PD5	19	0.87	0.89	0.91	23.8	24.4	29.1	11.1	12.7	13.2
CHP5-SND-PD6	14	0.88	0.90	0.92	23.6	23.8	29.4	11.7	13.0	14.2
CHP5-SND-PD7	14	0.88	0.89	0.91	22.6	23.1	27.7	10.2	12.4	12.5
CHP5-SND-PD8	15	0.87	0.90	0.90	23.3	23.2	29.3	10.6	12.1	12.7
CHP5-SND-PD9	12	0.87	0.88	0.87	23.4	24.3	30.1	11.5	12.9	14.0
CHP5-SND-PD10	36	0.88	0.90	0.90	23.0	22.6	27.1	10.9	12.5	13.2
CHP5-SND-PD11	16	0.87	0.88	0.89	22.9	23.8	29.2	10.5	12.9	13.2
CHP5-SND-PD12	12	0.87	0.88	0.87	22.9	24.4	29.7	10.9	13.3	13.7
CHP5-SND-PD13	15	0.87	0.86	0.85	22.9	25.6	31.8	11.1	13.7	14.8
CHP5-SND-PD14	7	0.88	0.89	0.88	22.8	23.2	29.8	10.5	12.4	13.5
CHP5-SND-PD15	12	0.88	0.89	0.90	23.0	23.1	28.9	10.9	12.4	12.9
CHP5-SND-PD16	12	0.88	0.89	0.89	22.2	23.5	29.2	10.4	12.6	13.2
CHP5-SND-PD17	17	0.88	0.89	0.91	22.5	23.1	27.4	10.4	12.3	12.2
CHP5-SND-PD18	20	0.89	0.89	0.91	24.2	25.6	32.8	12.9	14.1	15.6
CHP5-SND-PD19	15	0.87	0.88	0.88	22.9	24.1	30.3	10.8	13.0	14.1
CHP5-SND-PD20	8	0.86	0.87	0.86	24.5	25.6	31.7	11.1	13.3	13.4

T = training, S = testing and V= validation

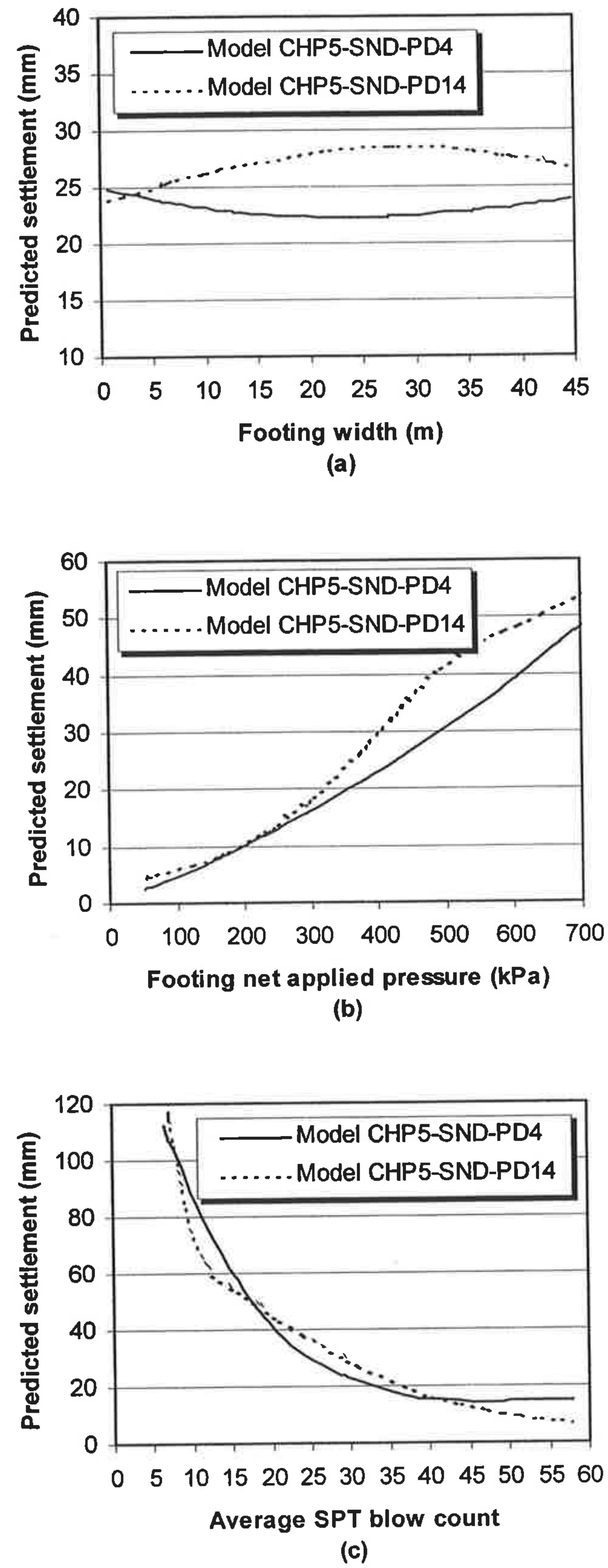


Figure 5.5: Parametric study for Models CHP5-SND-PD4 and CHP5-SND-PD14

- ***Effect of Software Implementation***

In order to investigate the effect of software implementation on the robustness of ANN models, an alternative commercial software system *Neuframe* Version 4.0 (Neusciences 2000), is used. The ANN models that are developed using *Neuframe* are described in the next section.

5.2.2 ANN Models Developed Using *Neuframe*

The model input and output variables, data subsets (i.e. training, testing and validation) and data pre-processing used are the same as those used in the development of model CHP5-PD14 (see §5.2.1). As *Neuframe* Version 4.0 does not support an automatic model-building procedure, the optimum network architecture is obtained using a trial-and-error approach. As mentioned in Chapter 2, a network with one hidden layer can approximate any continuous function, provided that sufficient connection weights are used (Cybenko 1989; Hornik et al. 1989). Consequently, one hidden layer is used in this research.

The general strategy adopted for finding the optimal network architecture and internal parameters that control the training process is as follows. A number of trials are carried out using the default parameters of the software used with one hidden layer and 1, 2, 3, …, 11 hidden layer nodes. It should be noted that 11 is the upper limit for the number of hidden layer nodes needed to map any continuous function for a network with 5 inputs, as discussed by Caudill (1988) and consequently, is used in this work. 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 an attempt to improve model performance. As discussed in Chapter 2, since the back-propagation algorithm uses a first-order gradient descent technique to adjust the connection weights, it may get trapped in a local minimum if the initial starting point in weight space is unfavourable. Consequently, the model that has the optimum momentum term, learning rate and transfer functions is retrained a number of times with different initial weights until no further improvement occurs.

Using the default parameters of the software, a number of networks with different numbers of hidden layer nodes are developed and the results are shown graphically in Figure 5.6 and summarised in Table 5.12. It can be seen from Figure 5.6 that the number of hidden layer nodes has little impact on the predictive ability of the ANN model. Even a network with only one hidden layer node is able to adequately map the underlying relationship. For networks with larger numbers of hidden layer nodes, there is no sign of overtraining, as evidenced by fairly consistent prediction errors. This is to be expected, as cross-validation is used as the stopping criterion. Figure 5.6 shows that the network with 5 hidden layer nodes has the lowest prediction error. However, it is believed that the network with 2 hidden layer nodes is considered optimal, as its prediction error is not far from the network with 5 hidden layer nodes (the error difference being only 0.17 mm) coupled with a smaller number of connection weights. It can also be seen from Table 5.12 that the results obtained for model CHP5-NF2 during validation are generally consistent with those obtained during training and testing, indicating that the model is able to generalise within the range of the data used for training, and can thus be used for predictive purposes.

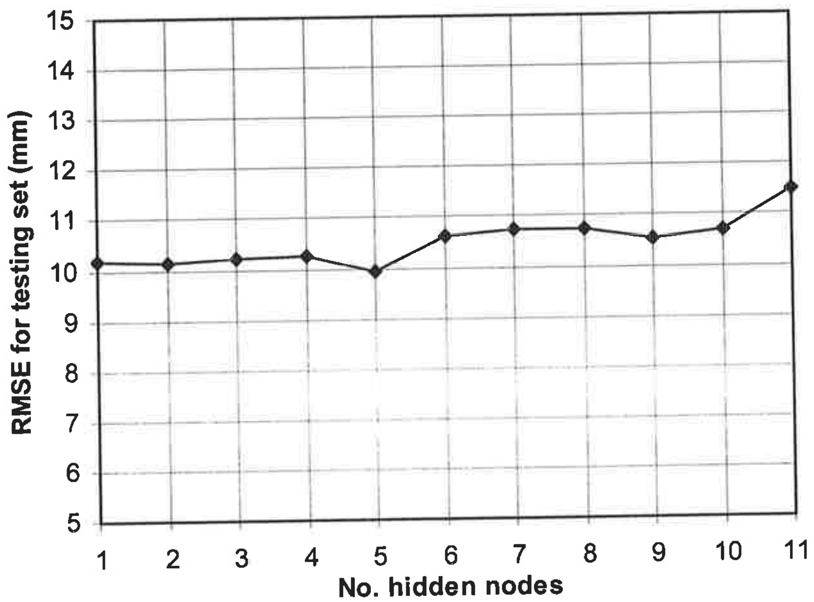


Figure 5.6: Performance of the ANN models developed using *Neuframe* with different hidden layer nodes

(Learning rate = 0.2 and momentum term = 0.8)

Table 5.12: Structure and performance of ANN models developed using *Neuframe*

Parameters effect	Model No.	No. hidden nodes	Learning rate	Momentum term	Transfer function in hidden layer	Transfer function in output layer	Performance measures								
							Correlation coefficient, *r*			RMSE			MAE		
							T	S	V	T	S	V	T	S	V
Default parameters	CHP5-NF1	1	0.2	0.8	Tanh	Sigmoid	0.930	0.928	0.903	10.00	10.16	11.10	6.85	6.42	8.79
	CHP5-NF2	2	0.2	0.8	Tanh	Sigmoid	0.930	0.929	0.905	10.01	10.12	11.04	6.87	6.43	8.78
	CHP5-NF3	3	0.2	0.8	Tanh	Sigmoid	0.931	0.927	0.906	9.90	10.23	10.91	6.76	6.78	8.69
	CHP5-NF4	4	0.2	0.8	Tanh	Sigmoid	0.919	0.929	0.895	10.79	10.27	11.63	7.38	6.89	9.13
	CHP5-NF5	5	0.2	0.8	Tanh	Sigmoid	0.936	0.931	0.899	9.54	9.95	11.34	6.56	6.37	9.06
	CHP5-NF6	6	0.2	0.8	Tanh	Sigmoid	0.916	0.931	0.889	11.15	10.61	12.03	7.85	7.46	9.49
	CHP5-NF7	7	0.2	0.8	Tanh	Sigmoid	0.915	0.928	0.891	11.17	10.74	11.87	7.81	7.46	9.31
	CHP5-NF8	8	0.2	0.8	Tanh	Sigmoid	0.915	0.929	0.890	11.19	10.73	11.93	7.86	7.51	9.38
	CHP5-NF9	9	0.2	0.8	Tanh	Sigmoid	0.916	0.929	0.890	11.08	10.54	11.84	7.64	7.27	9.29
	CHP5-NF10	10	0.2	0.8	Tanh	Sigmoid	0.915	0.928	0.891	11.12	10.70	11.79	7.70	7.38	9.23
	CHP5-NF11	11	0.2	0.8	Tanh	Sigmoid	0.918	0.928	0.899	11.41	11.51	11.39	7.81	7.82	8.62
Momentum terms	CHP5-NF12	2	0.2	0.01	Tanh	Sigmoid	0.916	0.929	0.88	11.06	10.56	11.61	7.44	7.05	8.88
	CHP5-NF13	2	0.2	0.05	Tanh	Sigmoid	0.917	0.929	0.890	11.03	10.58	11.55	7.42	7.07	8.84
	CHP5-NF14	2	0.2	0.1	Tanh	Sigmoid	0.916	0.928	0.891	11.10	10.75	11.47	7.41	7.12	8.68
	CHP5-NF15	2	0.2	0.15	Tanh	Sigmoid	0.917	0.928	0.893	11.07	10.79	11.39	7.38	7.15	8.62
	CHP5-NF16	2	0.2	0.2	Tanh	Sigmoid	0.916	0.929	0.888	10.99	10.45	11.68	7.44	7.02	9.00
	CHP5-NF17	2	0.2	0.4	Tanh	Sigmoid	0.919	0.928	0.898	10.92	10.79	11.19	7.31	7.16	8.55
	CHP5-NF18	2	0.2	0.6	Tanh	Sigmoid	0.920	0.926	0.902	10.83	10.87	10.96	7.19	7.11	8.39
	CHP5-NF19	2	0.2	0.9	Tanh	Sigmoid	0.933	0.919	0.907	9.84	11.05	10.60	6.65	6.91	8.38
	CHP5-NF20	2	0.2	0.95	Tanh	Sigmoid	0.935	0.911	0.907	9.98	11.80	10.67	6.50	7.42	8.23
Learning rates	CHP5-NF21	2	0.005	0.8	Tanh	Sigmoid	0.901	0.924	0.855	11.85	10.49	12.96	7.95	6.81	9.89
	CHP5-NF22	2	0.02	0.8	Tanh	Sigmoid	0.910	0.928	0.875	11.30	10.39	12.16	7.61	6.88	9.26
	CHP5-NF23	2	0.10	0.8	Tanh	Sigmoid	0.920	0.926	0.903	10.84	10.90	10.95	7.19	7.12	8.37
	CHP5-NF24	2	0.15	0.8	Tanh	Sigmoid	0.930	0.926	0.907	9.96	10.41	10.73	6.77	6.50	8.52
	CHP5-NF25	2	0.4	0.8	Tanh	Sigmoid	0.935	0.924	0.902	9.64	10.53	10.95	6.69	6.41	8.65
	CHP5-NF26	2	0.6	0.8	Tanh	Sigmoid	0.936	0.921	0.904	9.57	10.89	10.94	6.60	7.03	8.63
	CHP5-NF27	2	0.8	0.8	Tanh	Sigmoid	0.941	0.915	0.902	9.29	11.45	11.08	6.47	7.47	8.66
	CHP5-NF28	2	0.9	0.8	Tanh	Sigmoid	0.944	0.919	0.896	9.11	11.25	11.51	6.34	7.49	8.91
	CHP5-NF29	2	0.95	0.8	Tanh	Sigmoid	0.949	0.924	0.887	8.80	10.97	12.14	6.17	7.51	9.37
	CHP5-NF30	2	0.99	0.8	Tanh	Sigmoid	0.949	0.924	0.886	8.70	10.96	12.22	6.13	7.52	9.41
Transfer functions	CHP5-NF31	2	0.2	0.8	Sigmoid	Sigmoid	0.917	0.927	0.893	10.92	10.49	11.37	7.29	6.90	8.71
	CHP5-NF32	2	0.2	0.8	Tanh	Tanh	0.918	0.927	0.913	11.43	11.50	10.68	7.96	8.02	8.36
	CHP5-NF33	2	0.2	0.8	Sigmoid	Tanh	0.909	0.928	0.899	11.83	10.96	11.14	8.34	7.62	8.78

T = training, S = testing and V = validation

The effect of the internal parameters controlling the back-propagation algorithm (i.e. momentum term and learning rate) on model performance is investigated for the model with two hidden layer nodes (model CHP5-NF2) resulting in models CHP5-NF12 to CHP5-NF30 (Table 5.12). The effect of the momentum term on model performance is shown graphically in Figure 5.7. It can be seen that the performance of the ANN model is relatively insensitive to momentum, particularly in the range 0.01 to 0.6. The best prediction was obtained with a momentum value of 0.8.

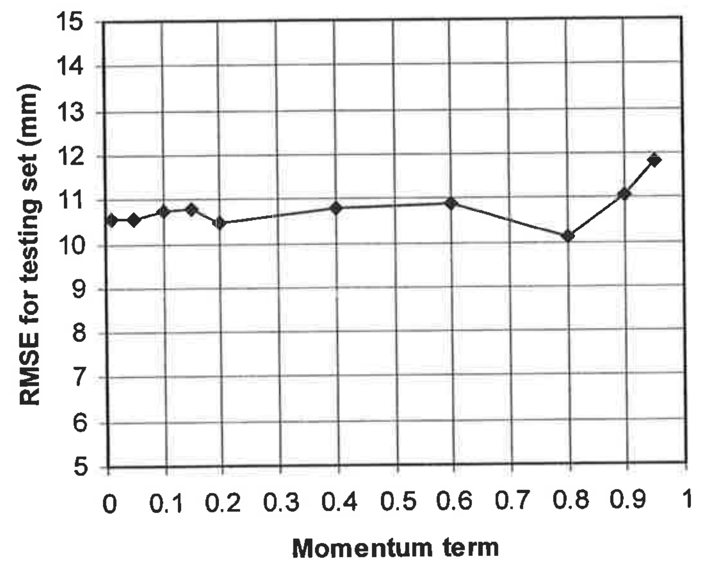


Figure 5.7: Effect of various momentum terms on ANN performance (Hidden nodes = 2 and learning rate = 0.2)

Figure 5.8 shows the effect of different learning rates on model performance. It can be seen that the optimum learning rate was found to be 0.2. At smaller learning rates, prediction errors were higher, probably as a result of the inability of the networks to escape local minima in the error surface due to the small step sizes taken. At larger learning rates, prediction errors increased slightly, possibly as a result of the pseudo-random behaviour of the optimisation algorithm near the local minima in the error surface due to the large step sizes taken in weight space (Maier and Dandy 1998).

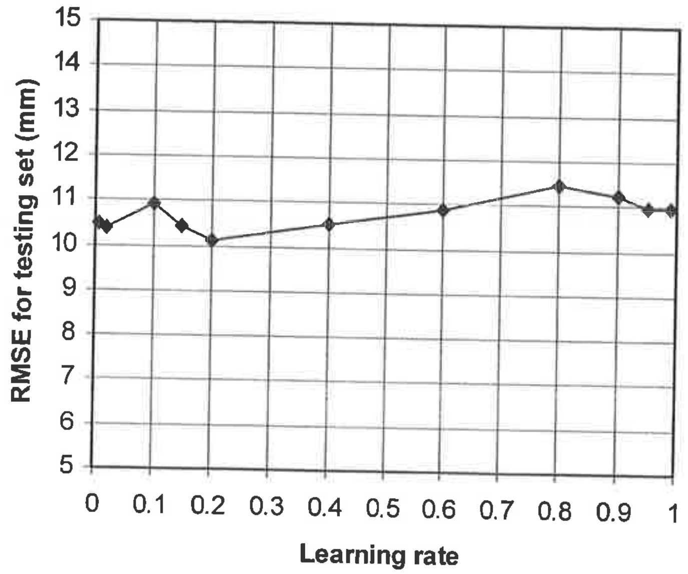


Figure 5.8: Effect of various learning rates on ANN performance (Hidden nodes = 2 and momentum term = 0.8)

The effect of using different transfer functions is shown in Table 5.12 (models CHP5-NF31 to CHP5-NF33). It can be seen that the performance of ANN models is insensitive to transfer functions although a slightly better performance is obtained when the tanh transfer function is used for the hidden layer and the sigmoid transfer function is used for the output layer. The effect of different random starting positions in weight space on prediction error was negligible for the model trained with 2 hidden layer nodes, a momentum value of 0.8, a learning rate of 0.2, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer (i.e. model CHP5-NF2). One possible reason for this is that the error surface in weight space is relatively uncomplicated for the problem under consideration. In addition, as discussed above, by using a learning rate of 0.2, the model is likely to escape local minima in the error surface during training.

The results of the parametric study that investigates the generalisation ability and robustness of model CHP5-NF2 is shown in Figure 5.9. It can be seen that the behaviour of model CHP5-NF2 is as expected, which indicates that the model may be considered to be robust. For example, in Figures 5.9 (a), (b) and (d), respectively, there is an increase in the predicted settlement as footing width, footing net applied pressure and footing geometry increase, respectively. On the other hand, in Figures 5.9 (c) and (e), respectively, the predicted settlement decreases as the average SPT blow count and footing embedment ratio increase, respectively.

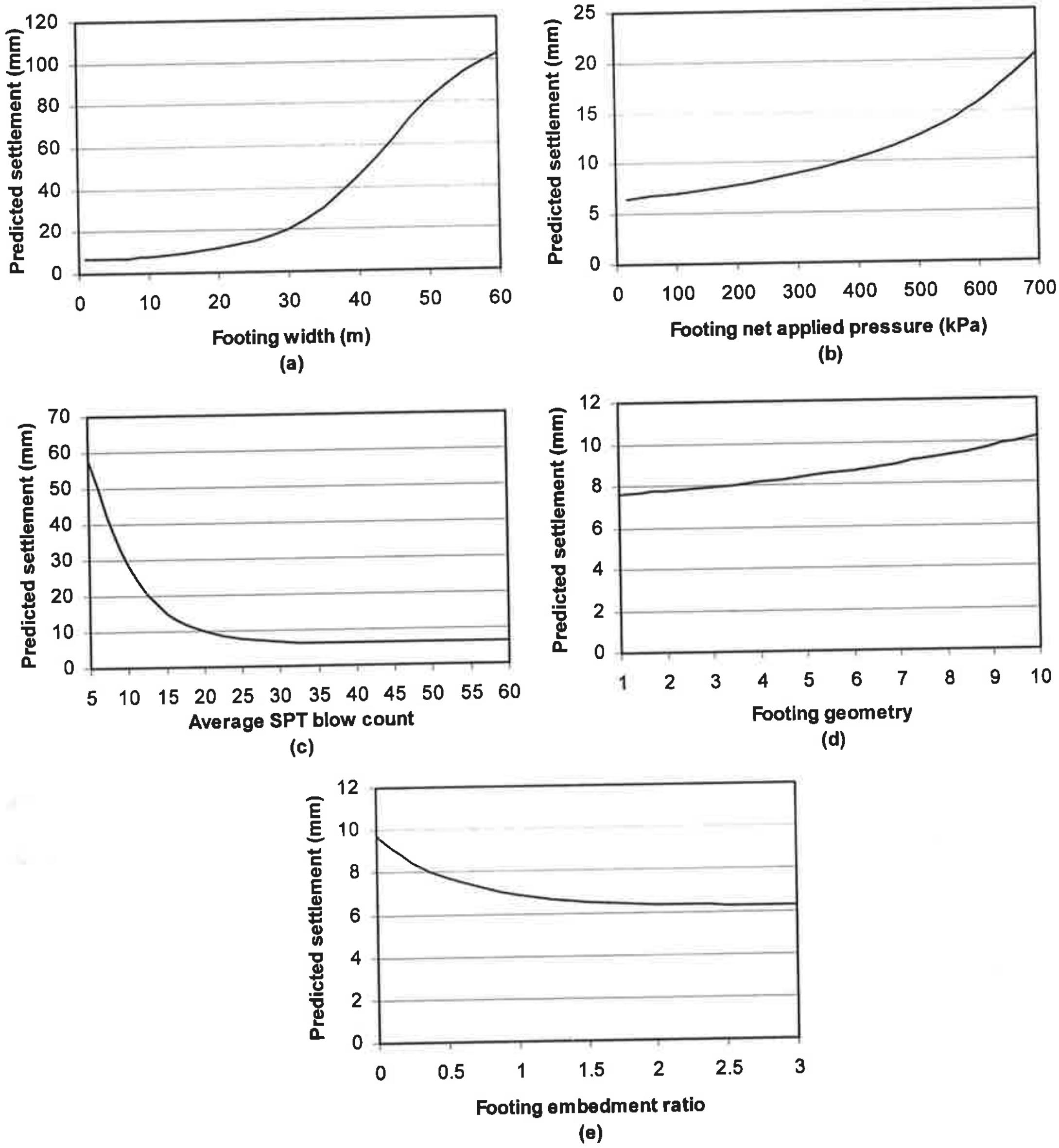


Figure 5.9: Results of parametric study for Model CHP5-NF2

In an attempt to further investigate the effect of the software used on the robustness of ANN models, additional models are developed using *Neuframe* with the same 500 data cases derived from Equation 5.2 for both clean and noisy samples (i.e. the data used in the development of models CHP5-SCD-PD1 to CH5-SCD-PD20 and CHP5-SND-PD1 to CHP5-SND-PD20). The optimal network parameters that are used for developing model CHP5-NF2 (i.e. learning rate of 0.2, momentum term of 0.8, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer) are utilised and networks with 1, 2, 3, ..., 7 hidden layer nodes are trained, resulting in models CHP5-SCD-NF1 to CHP5-SCD-NF7 for clean data and models CHP5-SND-NF1 to CHP5-SND-NF7 for noisy data. The structure and performance results of the developed models are given in Table 5.13. It can be seen that four hidden layer nodes were optimal for both clean and noisy data (i.e. model CHP5-SCD-NF4 for clean data and model CHP5-SND-NF4 for noisy data). The results of the parametric study for models CHP5-SCD-NF4 and CHP5-SND-NF4 are shown in Figure 5.10. It can be seen that regardless of whether the data are clean or noisy, the models obtained using *Neuframe* succeed in interpreting the physical meaning of the settlement problem in a robust fashion. It can also be seen that the models obtained are able to produce similar predictions to those of Meyerhof's equation. This suggests that it is the software used, rather than the degree of noise in the data, that results in the model behaviour exhibited in Figure 5.1.

Table 5.13: Structure and performance of ANN models developed using *Neuframe* of synthetic clean and noisy data

Model No.	No. hidden nodes	Performance measures								
		Correlation coefficient, r			RMSE (mm)			MAE (mm)		
		T	S	V	T	S	V	T	S	V
CHP5-SCD-NF1	1	0.973	0.960	0.967	11.4	11.9	14.8	8.6	8.9	10.0
CHP5-SCD-NF2	2	0.991	0.988	0.991	6.1	7.2	6.5	3.9	4.2	4.1
CHP5-SCD-NF3	3	0.991	0.988	0.991	6.2	6.3	6.7	4.0	4.3	4.2
CHP5-SCD-NF4	4	0.997	0.996	0.998	3.4	3.4	4.1	2.0	2.8	2.0
CHP5-SCD-NF5	5	0.997	0.997	0.998	3.3	3.3	3.9	1.9	2.0	1.9
CHP5-SCD-NF6	6	0.997	0.997	0.998	3.3	3.3	4.0	1.8	1.8	1.9
CHP5-SCD-NF7	7	0.997	0.997	0.998	3.5	3.6	4.4	2.0	2.1	2.1
CHP5-SND-NF1	1	0.877	0.867	0.894	24.0	24.4	32.7	13.4	15.7	16.7
CHP5-SND-NF2	2	0.884	0.889	0.916	23.0	24.7	30.3	11.2	13.3	13.9
CHP5-SND-NF3	3	0.885	0.890	0.918	23.0	24.5	30.1	11.4	13.2	14.1
CHP5-SND-NF4	4	0.885	0.889	0.916	22.3	23.6	26.2	10.5	12.8	12.5
CHP5-SND-NF5	5	0.886	0.889	0.914	22.2	23.6	26.3	10.5	12.6	12.4
CHP5-SND-NF6	6	0.886	0.889	0.915	22.5	24.1	29.2	10.7	13.0	13.4
CHP5-SND-NF7	7	0.886	0.889	0.913	22.3	23.6	26.4	10.6	12.8	12.6

T = training, S = testing and V= validation

It is evident from the previous analyses that model CHP5-NF2 is optimal and can be used successfully as a robust ANN model for settlement prediction of shallow foundations on cohesionless soils. The data used for the development of this model and the predicted settlement obtained for the training, testing and validation sets are summarised in Table 5.14.

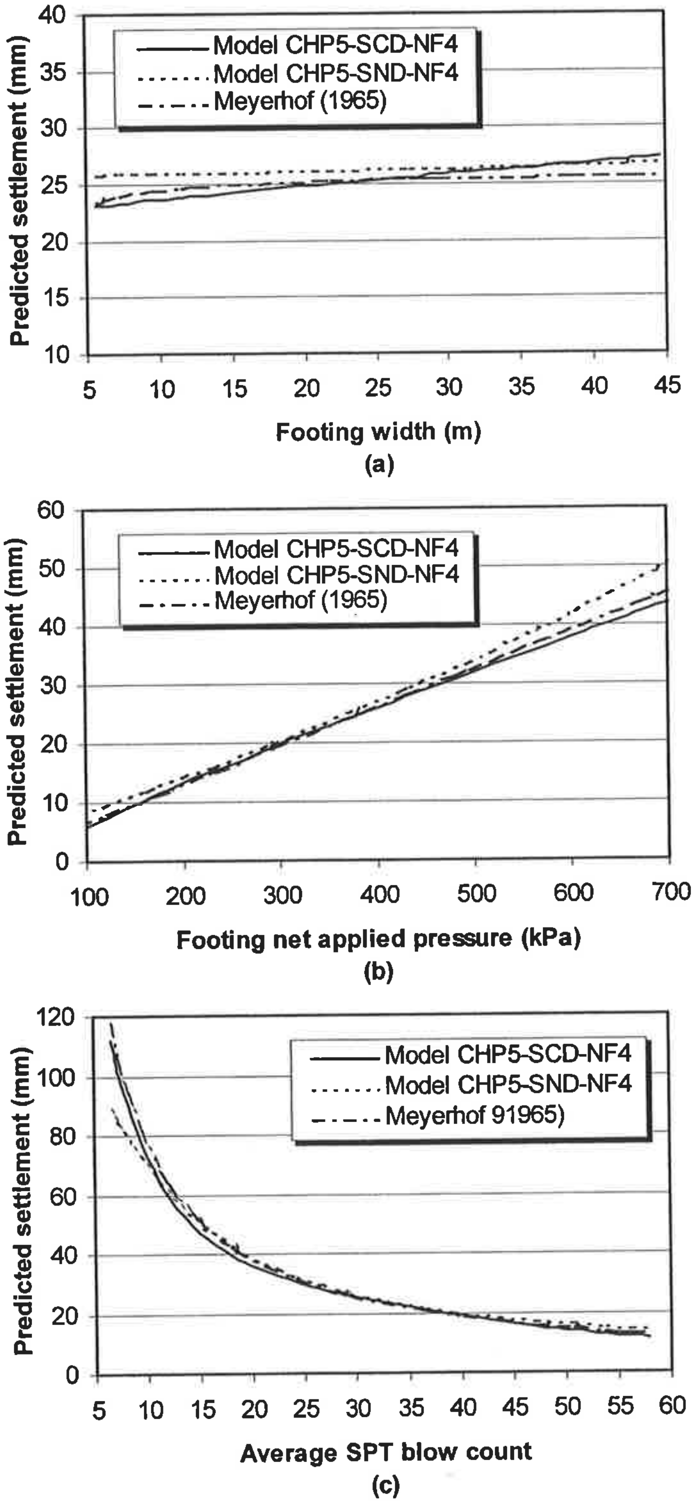


Figure 5.10: Results of parametric study for Models CHP5-SCD-NF4 and CHP5-SND-NF4

Table 5.14: Training, testing and validation data used for the ANN model

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
1*	60	385	47	1	0.09	40	30.1
2*	0.8	78	15	1	0.0	7	11.2
3	2.1	697	50	1	0.71	2.3	7.5
4	14	18.32	15	1.61	0.18	4.2	19.1
5	2.5	284	60	3.8	1.2	1	6.7
6	2.8	142	4	5	0.36	97	60.3
7	1.2	250	25	10.6	0.25	10	10.5
8	0.9	300	20	1	3.44	6.7	7.4
9	25	70	6	1	0.04	121	102.5
10	1.2	150	45	1	0.5	0.6	6.2
11	4.5	195	35	1.3	0.67	3.9	6.3
12	5.5	93	35	2.9	0.52	6.5	6.2
13	4.3	161	20	1.6	0.49	5	8.2
14	4.5	91	12	6.8	0.6	11	16.7
15	15	81	35	4.9	0.2	5.4	6.9
16	4.9	188	20	1.59	0.47	15	8.8
17	4	145	20	1.6	0.5	7.4	7.9
18	2.5	158	21	5.24	0.0	11.7	11.4
19	1.5	77	13	1.0	0.8	2.1	7.8
20	6.0	190	7	1	0.0	74	64.6
21	1.0	284	45	1.0	0.5	4.7	6.5
22	3.3	304	40	1.7	0.9	11.6	6.6
23*	12.2	130	17	1.0	0.09	22	19.5
24	5.2	127.8	58	3.7	0.0	17	6.3
25	3.8	90	12	3.2	0.39	15.5	15.0

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
26	6.7	113	21	1.59	0.51	5.0	7.7
27*	27.4	154	17	1.0	0.0	100	72.3
28	25	75	6	1.0	0.11	87	101.1
29	1.2	199	7	1.0	0.17	13	38.5
30	4.3	102	20	1.6	0.49	7.1	7.5
31	21.7	148	30	1.0	0.14	19.8	9.7
32	5.2	95.8	42	5.3	0.44	9.9	6.1
33	1.0	220	34	1.0	0.0	3.6	6.6
34*	2.5	245	16	1.0	0.0	11	17.8
35	4.9	118.7	22	1.1	0.3	6.4	7.7
36	4.0	512	37	1.8	1.3	12.8	7.2
37	1.5	77	13	1.0	0.8	1.3	7.8
38*	36.6	193	28	1.0	0.0	18.0	40.7
39	14.5	74	6	4.4	0.07	75	90.7
40	30.2	386	18	1.0	0.09	91.6	98.0
41	6.4	71.8	18	1.45	0.23	6.6	9.7
42	4.1	125	20	1	1.2	17.8	6.5
43	3.0	140	38	4.8	0.95	3.0	6.2
44	4.0	225	20	1.6	0.5	9.1	8.9
45	6.4	150	20	1.6	0.5	14.5	8.5
46	4.3	139	20	1.6	0.49	7.1	7.9
47	16.2	154	16	1.6	0.29	15	26.8
48	4.9	123	20	1.6	0.47	6.6	7.9
49	1.2	300	50	1.0	0.42	4.5	6.6
50	4.9	107	20	1.6	0.47	3.6	7.7

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	*q* (kPa)	*N*	*L/B*	D_f/B	S_m (mm)	S_p (mm)
51	22.5	221	20	2.9	0.44	21	29.6
52	2.5	576	18	1.0	0.3	25	27.0
53	3.7	135	20	1.0	1.4	10.1	6.4
54	22.4	64.0	6.0	3.8	0.04	70	102.9
55	4.9	182	20	1.6	0.47	13.8	8.7
56	4.3	134	20	1.0	1.2	15.4	6.6
57	3.0	500	18	1.0	0.29	25	22.2
58	5.1	114.9	42	4.6	0.35	5.8	6.1
59	4.9	97	20	1.6	0.47	4.3	7.7
60	4.3	150	20	1.6	0.49	6.8	8.0
61	1.0	294	40	1.0	0.0	5.0	6.6
62*	22	79	21	1.0	0.23	10.5	17.3
63	5.2	134	22	1.0	0.96	14.7	6.7
64	25	75	6	1.0	0.09	87	101.6
65*	33.5	156	19	1.0	0.0	90	81.1
66	1.2	215	18	1.0	2.2	8.6	6.3
67	6.6	168.1	39	2.0	0.0	15.5	6.5
68	4.3	145	20	1.6	0.49	11	8.0
69	5.2	153.2	44	3.7	0.0	8.9	6.3
70	4.9	161.4	49	2.8	0.0	7.1	6.3
71	22.4	75	6	3.8	0.04	92	103.4
72	5.0	181.9	24	1.7	0.5	11.9	7.4
73	3.4	129	20	1.0	1.5	11.5	6.3
74	11	120	24	3.0	0.45	19.6	8.3
75	20	85	5	1.0	0.15	116	95.5

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	*q* (kPa)	*N*	*L/B*	D_f/B	S_m (mm)	S_p (mm)
76	2.6	147	10	8.5	0.77	12	21.2
77	4.0	507.5	32	1.8	1.3	11.9	7.3
78	4.5	304	40	1.5	0.67	18.3	6.7
79	1.2	268	8	1.0	0.75	12.7	19.2
80	3.7	215	20	1.59	0.49	15	8.7
81	13.1	47.6	25	1.8	0.23	3.6	8.0
82*	33	191	34	1.0	0.16	43.8	12.5
83	8.5	102.5	24	1.0	0.0	16.3	8.7
84	1.5	150	35	1.0	0.4	2.1	6.2
85*	1.0	247.5	16	1.0	0.0	9.9	16.2
86	4.9	112	20	1.7	0.31	7.4	8.4
87	3.7	215	20	1.59	0.49	6.4	8.7
88	4.9	113	20	1.59	0.47	8.9	7.9
89	16	209	14	2.7	0.46	18.6	38.3
90	22	82	21	3.4	0.22	7.7	21.7
91	1.2	215	26	1.0	2.2	1.5	6.3
92*	10	240	60	1.0	0.15	7.0	6.7
93	1.4	230	25	1.0	2.1	3.9	6.3
94	4.3	134	20	1.6	0.49	10.2	7.9
95	4.9	102	20	1.59	0.47	6.9	7.7
96	3.3	52	8	4.2	0.54	35	19.2
97	6	214.5	42	2.7	0.6	4.1	6.5
98	2.1	584	50	1.1	1.4	4.6	7.3
99	1.4	300	50	1.0	2.6	1.5	6.7
100	4.4	93	10	5.5	0.57	8.0	19.9

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
101	1.6	250	25	7.9	0.25	9.3	9.4
102	4.9	199	20	1.6	0.47	11.7	8.9
103	23.6	167	35	1.14	0.13	15.4	8.4
104	1.5	666	18	1.0	0.51	25.0	25.7
105	3.3	52	8	4.2	0.54	20.0	19.2
106	2.5	284	60	3.8	1.2	3.0	6.7
107	19	80	15	1.0	0.0	52.0	42.9
108	22.9	165	30	1.4	0.13	20.4	10.7
109	5.5	139	20	1.6	0.47	9.4	8.3
110	3.0	231	20	1.6	0.5	8.1	8.7
111	3.7	290	20	1.59	0.49	11.2	9.9
112	3.4	247	20	1.6	0.5	12.2	9.0
113	12.2	181	53	1.0	0.25	9.6	6.5
114	7.0	177	22	1.6	0.5	8.3	8.2
115	5.6	112	22	4.3	0.27	15.5	8.8
116	13	193	18	2.4	0.16	22	23.5
117	3.3	98.6	7.0	4.4	0.61	37.1	24.3
118	1.2	320	25	1.0	0.0	2.8	9.2
119	25	63	6	1.0	0.08	84	101.3
120	13	193.8	18	1.7	0.16	18.8	22.0
121	4.6	112	24	5.0	0.43	11.2	7.4
122	6.1	155.6	38	5.0	0.25	16.8	6.5
123	4.6	85.7	39	4.5	0.59	21.1	6.1
124	1.0	564	45	1.0	0.5	4.4	7.3
125	5.8	72.8	17	4.1	0.43	11.9	10.2

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
126	4.6	113	20	1.6	0.5	5.1	7.7
127	3.7	252	20	1.6	0.49	16.5	9.3
128	6.1	144.1	23	5.0	1.1	11.7	6.8
129	3.7	139	20	1.6	0.49	7.4	7.8
130	7.0	131.2	42.0	5.1	0.33	11.9	6.3
131	6.0	158	42	2.7	0.47	7.9	6.3
132	3.7	279	20	1.6	0.49	8.6	9.7
133	16	70	12	1.3	0.09	90	45.1
134	6.0	162	30	2.7	0.6	11.0	6.6
135	0.9	113	6	1.0	1.0	6.4	11.6
136	3.4	81.4	34	6.7	0.0	10.7	6.5
137	4.0	97	20	1.6	0.5	6.1	7.4
138	2.4	190	22	1.6	1.9	8.5	6.3
139	17.6	218	20	4.8	0.61	26	19.4
140	4.3	177	20	1.6	0.49	8.1	8.3
141	3.0	500	18	1.0	0.25	25.0	23.5
142	1.5	150	50	1.0	0.4	1.0	6.2
143	55	233.6	60	1.8	0.18	25.0	7.9
144	5.3	121	17	9.9	0.49	12.0	15.6
145	1.0	196	25	1.0	3.0	6.0	6.2
146*	42.7	166	21	1.0	0.0	80.0	96.5
147	20.0	85.0	5	1.0	0.15	81.0	95.4
148	0.9	300.0	30	1.0	1.3	4.0	6.5
149	20.0	145.0	7	1.0	0.0	120.0	98.5
150	3.5	25.0	12	1.0	0.43	3.0	10.3

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
151	2.1	584.0	50	1.1	1.1	4.4	7.2
152*	24.4	120.0	27	1.0	0.0	14.3	14.8
153	1.2	215.0	29	1.0	2.2	2.5	6.2
154	9.0	115.0	11	8.0	0.5	25.0	38.4
155	4.6	111.1	43	3.5	0.0	23.9	6.2
156	3.6	304.0	40	1.8	0.83	13.3	6.6
157	25.0	76.0	6	1.0	0.08	85.0	101.9
158	3.7	225.0	20	1.6	0.49	7.4	8.8
159	13.0	193.0	18	2.1	0.16	23.5	22.8
160	14.5	253.5	26	1.0	0.24	18.0	10.8
161	41.2	104.0	36	1.0	0.24	10.0	12.2
162	6.0	162.0	30	2.7	0.47	10.5	6.7
163	34.0	270.0	30	1.7	0.23	22.0	25.3
164	3.3	99.0	4	4.4	0.30	37.0	55.9
165	25.0	86.0	6	1.0	0.10	120.0	102.0
166	1.8	575.0	50	1.6	0.83	2.7	7.2
167	15.0	148.0	20	1.3	0.0	40.0	20.1
168	1.0	339.0	45	1.0	0.5	6.0	6.6
169*	15.2	33.0	20	1.0	0.02	2.8	13.7
170	15.0	136.0	55	1.7	0.40	16.2	6.5
171	2.6	293.0	37	4.1	0.38	10.9	6.8
172	6.4	100.5	18	1.0	0.23	7.1	10.1
173	4.6	166.0	20	1.6	0.5	8.1	8.2
174	1.2	150.0	28	1.0	0.5	1.3	6.4
175	6.1	161.0	20	1.6	0.49	10.2	8.6

Table 5.14: Training, testing and validation data used for the ANN model (continued)

Case record	Input variables					Output variable	
	B (m)	q (kPa)	N	L/B	D_f/B	S_m (mm)	S_p (mm)
176	3.0	230.8	50	3.3	1.0	21.1	6.5
177	1.1	78.0	13	1.0	1.09	2.0	7.0
178	1.8	230.0	25	1.0	1.7	3.4	6.3
179	0.9	133.0	5	1.0	0.33	7.6	31.4
180	5.1	116.8	19	3.1	0.24	19.3	10.2
181	0.9	300.0	20	1.0	1.33	2.7	6.9
182	2.25	400.0	8	1.1	1.02	43.0	22.7
183	2.6	196.3	9	8.1	0.77	33.0	27.9
184	2.1	347.0	50	1.9	1.4	1.8	6.7
185	14.5	74.0	6	4.4	0.07	74.0	90.7
186	25.5	175.0	21	1.0	0.1	25.0	37.3
187	1.0	284.0	25	2.2	3.0	10.5	6.4
188	17.2	34.0	17	2.5	0.27	3.6	19.1
189	18.3	41.0	20	1.0	0.02	4.8	17.2

*Circular footings
Training data = 1 to 121, Testing data = 122 to 152 and Validation data = 153 to 189

5.3 Data Division for ANN Models

It is evident from Chapter 3 that ANNs have been applied to many geotechnical engineering problems and have demonstrated some degree of success. In the majority of these applications, data division is carried out on an arbitrary basis. However, as mentioned in Chapter 2, the way the data are divided can have a significant impact on the performance of ANN models. In this section, the issue of data division and its impact on ANN model performance is investigated. Four data division methods are tested: (i) random data division; (ii) data division to ensure statistical consistency of the subsets needed for ANN model development; (iii) data division using self-organising maps (SOMs); and (iv) a new data division method using fuzzy clustering. For the first two methods, the relationship between the statistical properties of the training, testing and validation sets and model performance is investigated. For the second method, the effect of the proportion of data used for training, testing and validation on model performance is also investigated. The last two methods are introduced as alternative approaches for data division that (i) negate the need to choose which proportion of the data to use for training, testing and validation and (ii) ensure that each of the subsets are representative of the available data. The specific objectives of this section are:

1. To investigate the relationship between the statistical properties of the data subsets used to develop ANN models and model performance;
2. To introduce a new approach to data division for ANNs based on fuzzy clustering;
3. To compare the performance of the new approach with that of three existing approaches, including random data division, data division to ensure statistical consistency between the various subsets and data division using a SOM, although the second and third of these have yet to be applied to geotechnical engineering;
4. To investigate the relationship between the proportion of the data in each of the subsets used to develop ANN models and their performance, in relation to the data division method that ensures statistical consistency between data sets; and
5. To investigate the impact of the number of data points used from each cluster for training on model performance in relation to the SOM data division method.

In order to meet the objectives set out above, the four approaches to data division are investigated below.

- **Approach 1: Random**

As mentioned in Chapter 2, a random approach is generally used in the field of geotechnical engineering for dividing the available data into the subsets needed for ANN model development, with no attention given to the statistical consistency of the data subsets. As a result, the performance of the trained model on the validation data is highly dependent on which data are contained in the validation set (e.g. whether the validation set contains extreme data or not), making it impossible to assess the true generalisation ability of the model within the domain of the available data. Another shortcoming of this approach is that the proportion of the data to be used for training, testing and validation needs to be chosen *a priori* by the modeller. However, there are no firm guidelines in the literature to assist with this task, although, as mentioned in Chapter 2, some rules-of-thumb exist, such as using two thirds of the data for model calibration (i.e. training and testing) and one third for model validation.

As part of this approach, the available 189 individual case records are randomly divided into training, testing and validation subsets. In total, 80% of the data (i.e. 152 individual cases) are used for calibration and 20% of the data (i.e. 37 individual cases) are used for validation. The calibration data are further divided into 70% for training (i.e. 106 individual cases) and 30% for testing (i.e. 46 individual cases).

- **Approach 2: Statistically Consistent**

As mentioned in §5.2, part of this approach is to divide the data 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, and thus represent the same statistical population. The major shortcoming of this approach is that it is based on trial-and-error and that the proportion of the data to be used for training, testing and validation needs to be chosen in advance by the modeller, as discussed previously.

As mentioned above, as part of this approach, the 189 individual case records are divided into three statistically consistent subsets. In order to investigate the impact of the proportion of the data used in the various subsets in relation to model performance (see Objective 4), a number of different proportions of the available data are used for training, testing and validation. The different proportions investigated are summarised in Table 5.15.

Table 5.15: Different proportions of data for training, testing and validation

Validation set (%)	Remaining data	
	Training set (%)	**Testing set (%)**
10	70	30
	80	20
	90	10
20	70	30
	80	20
	90	10
30	70	30
	80	20
	90	10

- **Approach 3: Self-organising map (SOM)**

As mentioned in Chapter 2, self-organising maps belong to the class of unsupervised neural networks, and can be used for data clustering. Once clustering has been successfully accomplished, samples are chosen from each cluster to form the training, testing and validation sets. Different approaches for achieving this have been suggested in the literature. Kocjancic and Zupan (2000) suggested using a fixed number of data samples from each cluster to form the data subsets needed for ANN model development. However, this still requires a subjective decision as to what proportion of data points from each cluster to allocate to the different data subsets. Bowden et al. (2002) suggested randomly selecting three samples from each cluster to form the ANN data subsets; one for each of the training, testing and validation sets. In the instance when a cluster contains two data records, one record is chosen for training and the other is chosen for testing. If a cluster contains only one data record, this record is included in the training set. This approach overcomes the problem of having to decide how many data points from each cluster to allocate to the different data subsets. In addition, this approach utilises the minimum number of data points for model development, thus increasing computational efficiency. However, it is unclear if better model performance could be achieved if all data points remaining in a cluster, after removal of the testing and validation values, are used for training rather than just one data point from each cluster. Although Bowden et al. (2002) conducted a preliminary investigation into this issue and found that the inclusion of the additional training samples did not improve model performance, further investigation into this matter is needed.

In summary, the SOM data division method has a number of advantages, including:

1. There is no need to decide which proportion of the data to use for training, testing and validation.
2. The statistical properties of the training, testing and validation data are similar, provided that intra-cluster variation is sufficiently small.
3. Information is provided about whether "outliers" (not necessarily in the statistical sense) exist in the data set. For example, if a cluster contains only one data sample, this sample should be included in the training set. If it were to be included in the

validation set, the trained ANN model could not be expected to perform well, as the validation data may fall outside the range of the training data.

The main disadvantage of this approach is that different parameters that control the learning process in SOMs (i.e. learning rate, neighbourhood size and size and shape of the map) have to be selected in advance. Moreover, as mentioned in Chapter 2, there are no precise rules for the optimum choice of these parameters.

As part of this approach, the PC-based software *Neuframe* Version 4.0 (Neusciences 2000) is used to cluster the data using a SOM. The available data inputs (i.e. B, q, N, L/B and D_f/B) and corresponding output (S_m) are presented to the SOM as inputs (Figure 5.11). As mentioned previously, there is no precise rule for determining the optimum size of the map. Consequently, a number of map sizes are investigated, including 5 × 5, 6 × 6, 7 × 7 and 8 × 8. For all map sizes, the default parameters (e.g. learning rate and neighbourhood size) suggested in the software are used (Neusciences 2000) and training is continued for 10,000 iterations, as the connection weights remain stable after this point. A grid size of 8 × 8 is chosen as it ensures that the maximum number of clusters are found from the training data (Bowden et al. 2002).

In order to investigate the impact of the number of data points used from each cluster for training on model performance (see Objective 5), two different options for choosing training data from each cluster are adopted. As part of the first option, all data records remaining after the selection of the testing and validation data are used for training. As a result, a total of 110 records are used for training, 46 for testing and 33 for validation. As part of the second option, only one data point from each cluster is chosen for training. As a result, 54 records are used for training, 46 for testing and 33 for validation, resulting in a reduction in the data used for training by approximately 50%.

- **Approach 4: Fuzzy Clustering**

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

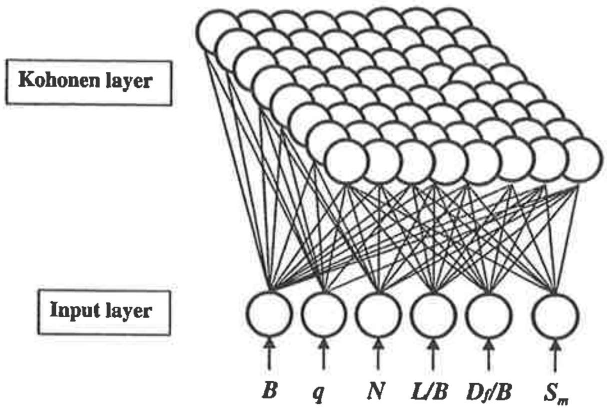


Figure 5.11: SOM for settlement data clustering

$$C = \sum_{\upsilon=1}^{l} \frac{\sum_{i,j=1}^{p} u_{i\upsilon}^2 u_{j\upsilon}^2 d_{ij}}{2\sum_{j=1}^{p} u_{j\upsilon}^2} \tag{5.3}$$

where:

l = number of clusters;

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

$u_{i\upsilon}$ = unknown membership function of data point i to cluster υ.

The sum in the numerator ranges over all pairs of data points (i,j), and the membership functions are subject to the following constrains:

$$u_{i\upsilon} \geq 0 \qquad \text{for } i = 1, \ldots, p;\ \upsilon = 1, \ldots, l \tag{5.4}$$

$$\sum_{\upsilon} u_{i\upsilon} = 1 \qquad \text{for } i = 1, \ldots, p \tag{5.5}$$

The above constraints imply that memberships cannot be negative, and that each data point has a constant total membership value, distributed over the clusters, normalised to unity. For hard clustering, a data point is assigned to the cluster that has the largest membership value.

The basic notion of fuzzy clustering for data division is similar to that underlying the SOM data division approach in the sense that both are used to cluster similar data records together and once data are clustered, samples are chosen from the clusters to form the training, testing and validation sets. However, the fuzzy clustering approach has a number of features that enable it to overcome the shortcomings of the SOM data division approach.

Firstly, an analytical procedure can be used to determine the optimum number of clusters. This is achieved with the aid of the silhouette value $s(i)$, which is a measure of how well individual data points lie within the cluster they have been assigned to at the end of the clustering process, and is given by (Kaufman and Rousseeuw 1990):

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, -1 \leq s(i) \leq 1 \tag{5.6}$$

where:

$a(i)$ = average dissimilarity of data point i to all other points in a cluster A; and

$b(i)$ = the smallest average dissimilarity of data point i to all points in any cluster E different from A.

For an individual data record (i) in cluster A, if $s(i)$ is close to 1, this implies that the "within" dissimilarity $a(i)$ is smaller than the smallest "between" dissimilarity $b(i)$, and therefore data record i can be deemed to have a strong membership to cluster A. By calculating the average silhouette width $\bar{s}(l)$ for the entire data set for different numbers of cluster, the optimum number of clusters can be determined by choosing the number of clusters that maximises the value of $\bar{s}(l)$.

Secondly, guidelines can be developed to determine which data points from each cluster should be used for training, testing and validation. Information about the degree of membership each data point has to the cluster it has been assigned to can be used to ensure that any significant intra-cluster variation is taken into account when assigning data points to their respective subsets. As part of the data division approach introduced in this research, it is suggested to rank the data points in each cluster in order of increasing membership value. Next, each data point is assigned to one of ten equally spaced membership intervals (i.e. 0.0–0.1, 0.1–0.2, ..., 0.9–1.0) and one data point from each membership interval is assigned to the testing set and another data point from that interval is assigned to the validation set while the remaining data points from the same interval are assigned to the training set. By using this approach, the best possible representation of the available data is achieved in each of the three data subsets.

The detailed procedure for using fuzzy clustering for ANN data division introduced in this research is as follows:

1. An initial number of clusters, not less than two, is chosen (the initial number of clusters can be assumed to be equal to 5% of the available data);
2. The available data are clustered using the fuzzy clustering technique and the average silhouette width $\bar{s}(l)$ of the entire data set is calculated;
3. The number of clusters is increased by one and Step 2 is repeated until $\bar{s}(l)$ remains constant or the number of clusters reaches 50% of the available data;
4. The number of clusters that result in the largest value of $\bar{s}(l)$ is considered optimum;
5. 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); and
6. 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 all remaining data samples are chosen for the training set. In the instance when 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.

As part of this method, the software FANNY (Kaufman and Rousseeuw 1990) is used to cluster the data using fuzzy clustering. Using the procedure outlined above, 10 to 94 clusters are tried. The average silhouette width of the entire data $\bar{s}(l)$ is maximised when 16 clusters are used and is equal to 0.3. The membership values obtained for all data records are shown in Appendix C. Using the procedure outlined previously, samples are chosen for the training, testing and validation sets and as a result, a total of 143 data records are used for training, 25 for testing and 21 for validation.

- **Results and Discussion**

The statistics of the training, testing and validation sets obtained when the data are divided in a purely random fashion (Approach 1) and where the statistics of the subsets are taken into account (Approach 2), are shown in Tables 5.16 and 5.17, respectively. It can be seen that when the data are divided in a purely random manner (Approach 1 – Table 5.16), there are some inconsistencies in the statistics between the various data subsets. This is confirmed by the results of the *t*- and *F*-tests (Table 5.18), which show that the hypotheses are rejected for most of the testing and validation sets and consequently, the data in the three subsets generally do not belong to the same statistical population. However, it should be noted that this is not necessarily the case when the data are divided in a random manner, as there are many different possible ways in which the data can be divided into training, testing and validation subsets. The results in Table 5.17 show that when the data are divided in a way that takes into account the statistical properties of the various subsets (Approach 2), the statistics are in much better agreement, as expected. This is confirmed by the outcomes of the *t*- and *F*-tests (Table 5.19), which show that the hypotheses are accepted for all of the testing and validation sets and consequently, the training, testing and validation sets are generally representative of each other.

The structure and performance results of the ANN models developed using the random data division method are given in Table 5.20. It should be noted that these models are developed using a learning rate of 0.2, momentum term of 0.8, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer. It can be seen from Table 5.20 that model CHP5-NF34 with one hidden layer node can be considered optimal.

The performance of the optimum models developed using the data sets whose statistics are shown in Tables 5.16 and 5.17 (i.e. models CHP5-NF34 and CHP5-NF2) are shown in Table 5.21 (columns 2 and 3). It can be seen that there is a direct relationship between the consistency in the statistics between training, testing and validation sets and consistency in model performance. When the training, testing and validation data are not representative of each other, there can be large discrepancies in the model performance obtained using the training, testing and validation sets. Consequently, the results obtained using the validation set may not be truly representative of the performance of the trained model, as the validation set may contain extreme data points that were not used in the model calibration (training) phase. Consequently, the best model given the available data has not been developed. Similarly, if the results obtained using the testing set are not representative of those obtained using the training set, training may be ceased at a sub-optimal time, or a sub-optimal network geometry or learning rate or momentum value may be chosen. However, when the training, testing and validation sets are representative of each other, the performance of the model on each of the three subsets is very similar, indicating that the model has the ability to interpolate within the extremes contained in the available data.

The model performances obtained when different proportions of the available data are used for training, testing and validation, in conjunction with the data division method which takes into account the statistical properties of the data (Approach 2), are shown in Table 5.22. It should be noted that these models are developed using two hidden layer nodes, a learning rate of 0.2, momentum term of 0.8, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer. The statistics and null hypothesis tests for the training, testing and validation sets are given in Appendices D and E, respectively. A code is used to distinguish between the various proportions of the available data used for training, testing and validation. The code consists of three numbers. The first number represents the percentage of the data used in the validation set, whereas the second two numbers, placed between brackets and separated by a hyphen, are the percentages that divide the remaining data into training and testing sets, respectively. It can be seen from Table 5.22 that there is no clear relationship between the proportion of data used for training, testing and validation and model performance. The best result is obtained when 20% of the data are used for validation and the remaining data are divided into 70% for training and 30% for testing (i.e. model CHP5-NF2). The results in Table 5.22 also indicate that there can be significant variation in the results obtained, depending on which proportion of the data is used for training, testing and validation, even when the statistical properties of the data subsets are taken into account. This may be due to the difficulties in obtaining representative data sets for some of the proportions for training, testing and validation investigated for the particular data set used.

Table 5.16: Input and output statistics obtained using random data division

Model variables and data sets	Statistical parameters				
	Mean	**Std. Dev.**	**Minimum**	**Maximum**	**Range**
Footing width, *B* (m)					
Training set	9.4	11.3	0.8	60.0	59.2
Testing set	9.2	10.3	0.9	41.2	40.3
Validation set	6.1	4.3	2.25	25.5	23.25
Footing net applied pressure, *q* (kPa)					
Training set	161.3	98.0	18.3	697.0	678.7
Testing set	267.2	155.2	47.6	666.0	618.4
Validation set	161.2	101.5	71.8	507.6	435.7
Average SPT blow count, *N*					
Training set	21.6	11.8	4.0	60.0	56.0
Testing set	28.6	15.7	4.0	60.0	56.0
Validation set	27.8	13.4	7.0	58.0	51.0
Footing geometry, *L/B*					
Training set	1.9	1.5	1.0	9.9	8.9
Testing set	1.9	1.9	1.0	10.5	9.5
Validation set	3.3	1.9	1.0	8.1	7.1
Footing embedment ratio, D_f/B					
Training set	0.57	0.59	0.0	3.4	3.4
Testing set	0.52	0.64	0.0	3.0	3.0
Validation set	0.41	0.41	0.0	1.8	1.8
Measured settlement, S_m (mm)					
Training set	20.7	28.7	0.6	121.0	120.4
Testing set	23.0	30.5	1.8	120.0	118.2
Validation set	16.1	9.7	4.1	43.0	38.9

Table 5.17: 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
Footing width, B (m)					
Training set	8.3	9.8	0.8	60.0	59.2
Testing set	9.3	10.9	0.9	55.0	54.1
Validation set	9.4	10.1	0.9	41.2	40.3
Footing net applied pressure, q (kPa)					
Training set	188.4	129.0	18.3	697.0	678.7
Testing set	183.2	118.7	25.0	584.0	559.0
Validation set	187.9	114.6	33.0	575.0	542.0
Average SPT blow count, N					
Training set	24.6	13.6	4.0	60.0	56.0
Testing set	24.6	12.9	5.0	60.0	55.0
Validation set	24.3	14.1	4.0	55.0	51.0
Footing geometry, L/B					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.3	1.9	1.0	9.9	8.9
Validation set	2.1	1.8	1.0	8.0	7.0
Footing embedment ratio, D_f/B					
Training set	0.52	0.57	0.0	3.4	3.4
Testing set	0.49	0.52	0.0	3.0	3.0
Validation set	0.59	0.64	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training set	20.0	27.2	0.6	121.0	120.4
Testing set	21.4	26.6	1.0	120.0	119.0
Validation set	20.4	25.2	1.3	120.0	118.7

Table 5.18: Null hypothesis tests for random data division

Variable and Data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	0.10	–1.97	1.97	Accept	1.20	0.59	1.87	Accept
Validation	1.70	–1.97	1.97	Accept	6.90	0.61	1.77	Reject
q								
Testing	–5.10	–1.97	1.97	Reject	0.39	0.59	1.87	Reject
Validation	0.00	–1.97	1.97	Accept	0.93	0.61	1.77	Accept
N								
Testing	–3.00	–1.97	1.97	Reject	0.56	0.59	1.87	Reject
Validation	–2.70	–1.97	1.97	Reject	0.78	0.61	1.77	Accept
L/B								
Testing	0.00	–1.97	1.97	Accept	0.62	0.59	1.87	Accept
Validation	–4.50	–1.97	1.97	Reject	0.62	0.61	1.77	Accept
D_f/B								
Testing	0.47	–1.97	1.97	Accept	0.85	0.59	1.87	Accept
Validation	1.52	–1.97	1.97	Accept	2.10	0.61	1.77	Reject
S_m								
Testing	–0.45	–1.97	1.97	Accept	0.89	0.59	1.87	Accept
Validation	0.95	–1.97	1.97	Accept	8.80	0.61	1.77	Reject

Table 5.19: 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*
B								
Testing	–0.58	–1.97	1.97	Accept	0.81	0.59	1.87	Accept
Validation	–0.61	–1.97	1.97	Accept	0.94	0.61	1.77	Accept
q								
Testing	0.23	–1.97	1.97	Accept	1.18	0.59	1.87	Accept
Validation	0.02	–1.97	1.97	Accept	1.27	0.61	1.77	Accept
N								
Testing	0.00	–1.97	1.97	Accept	1.11	0.59	1.87	Accept
Validation	0.11	–1.97	1.97	Accept	0.93	0.61	1.77	Accept
L/B								
Testing	–0.64	–1.97	1.97	Accept	0.80	0.59	1.87	Accept
Validation	0.00	–1.97	1.97	Accept	0.89	0.61	1.77	Accept
D_f/B								
Testing	0.31	–1.97	1.97	Accept	1.20	0.59	1.87	Accept
Validation	–0.62	–1.97	1.97	Accept	0.79	0.61	1.77	Accept
S_m								
Testing	–0.29	–1.97	1.97	Accept	1.05	0.59	1.87	Accept
Validation	–0.08	–1.97	1.97	Accept	1.17	0.61	1.77	Accept

Table 5.20: Structure and performance of ANN models using random data division

Model No.	No. hidden nodes	Performance measure								
		Correlation coefficient, *r*			RMSE			MAE		
		T	S	V	T	S	V	T	S	V
CHP5-NF34	1	0.944	0.845	0.659	9.35	16.39	10.57	6.23	11.94	8.85
CHP5-NF35	2	0.945	0.834	0.653	9.34	17.12	10.85	6.30	12.42	8.97
CHP5-NF36	3	0.944	0.844	0.660	9.61	16.54	10.96	6.45	12.06	9.34
CHP5-NF37	4	0.945	0.845	0.655	9.39	16.47	10.83	6.29	11.96	9.17
CHP5-NF38	5	0.944	0.848	0.656	9.39	16.27	10.75	6.26	11.82	9.08
CHP5-NF39	6	0.943	0.843	0.663	9.49	16.57	11.12	6.47	12.29	9.46
CHP5-NF40	7	0.943	0.837	0.671	9.51	16.93	11.28	6.47	12.65	9.69
CHP5-NF41	8	0.943	0.832	0.674	9.51	17.22	11.43	6.49	12.91	9.77
CHP5-NF42	9	0.942	0.846	0.647	9.59	16.48	10.97	6.49	12.23	9.13
CHP5-NF43	10	0.944	0.839	0.664	9.47	16.79	11.24	6.45	12.45	9.60
CHP5-NF44	11	0.943	0.852	0.651	9.55	16.12	11.02	6.53	11.94	9.27

T = training, S = testing and V= validation

Table 5.21: Performance of ANN models using data subsets obtained for different approaches of data division

Performance measures and data sets	Random division	Statistical division	SOM	Fuzzy clustering
Training				
Correlation coefficient, *r*	0.944	0.930	0.890	0.912
RMSE (mm)	9.35	10.01	11.58	10.62
MAE (mm)	6.23	6.87	7.93	7.43
Testing				
Correlation coefficient, *r*	0.845	0.929	0.942	0.967
RMSE (mm)	16.39	10.12	10.43	10.48
MAE (mm)	11.94	6.43	7.98	6.92
Validation				
Correlation coefficient, *r*	0.659	0.905	0.958	0.957
RMSE (mm)	10.57	11.04	10.12	9.59
MAE (mm)	8.85	8.78	7.12	6.13

Table 5.22: Performance of ANN models for different data proportions using statistical data division approach

Model No.	Data proportions and sets	Performance measures		
		Correlation coefficient, *r*	RMSE (mm)	MAE (mm)
CHP5-NF45	**10(70-30)**			
	Training set	0.922	9.34	6.55
	Testing set	0.929	11.34	7.35
	Validation set	0.861	17.08	9.49
CHP5-NF46	**10(80-20)**			
	Training set	0.939	9.26	6.63
	Testing set	0.876	13.82	7.96
	Validation set	0.909	12.72	9.07
CHP5-NF47	**10(90-10)**			
	Training set	0.934	9.25	6.04
	Testing set	0.924	13.87	10.43
	Validation set	0.849	18.35	9.95
CHP5-NF2	**20(70-30)**			
	Training set	0.930	10.01	6.87
	Testing set	0.929	10.12	6.43
	Validation set	0.905	11.04	8.78
CHP5-NF48	**20(80-20)**			
	Training set	0.933	9.57	6.63
	Testing set	0.929	10.96	6.94
	Validation set	0.898	11.39	9.01
CHP5-NF49	**20(90-10)**			
	Training set	0.918	10.67	7.51
	Testing set	0.945	10.46	6.89
	Validation set	0.878	12.52	9.49
CHP5-NF50	**30(70-30)**			
	Training set	0.920	11.01	7.88
	Testing set	0.938	10.93	7.28
	Validation set	0.903	10.94	7.76
CHP5-NF51	**30(80-20)**			
	Training set	0.926	10.68	7.12
	Testing set	0.903	11.52	7.71
	Validation set	0.887	11.55	7.83
CHP5-NF52	**30(90-10)**			
	Training set	0.923	10.10	7.38
	Testing set	0.835	16.33	10.78
	Validation set	0.920	10.80	7.53

The difficulties associated with deciding which proportion of the available data to use for training, testing and validation can be overcome by using a SOM (Approach 3) or fuzzy clustering (Approach 4) for obtaining appropriate data subsets. However, as discussed previously, two different approaches for choosing the training data from the clusters obtained for a SOM have been proposed in the literature, and are investigated here. The statistics of the data used from clusters obtained using the two different approaches of SOM are given in Tables 5.23 and 5.24, respectively, whereas the null hypothesis tests are given in Tables 5.25 and 5.26, respectively. It can be seen that the statistics of the data in each of the subsets obtained using the two approaches of SOM are very close to each other (Tables 5.23 and 5.24). This is confirmed by the results of and the t- and F-tests (Tables 5.25 and 5.26), which indicate that the three data sets in Tables 5.23 and 5.24 may be considered to be representative of the same statistical population. The performance of ANN models developed using the aforementioned SOM approaches (model CHP5-NF53 and CHP5-NF54) is shown in Table 5.27. It should be noted that these models are developed using two hidden layer nodes, a learning rate of 0.2, momentum term of 0.8, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer. The results in Table 5.27 indicate that it is better to use all of the data remaining after the testing and validation data have been removed from each cluster for training, rather than choosing only one data point from each cluster, as the RMSE in the testing set increases from 10.43 to 14.43 mm and the MAE increases from 7.98 to 10.21 mm, when the additional training data are discarded. However, there is a slight decrease in the coefficient of correlation, r, from 0.942 to 0.928 when the additional training data are included. Consequently, the subsequent discussion in relation to the SOM data division method (Approach 3) is restricted to the case where all remaining data are used for training.

The statistics of the data in each of the subsets obtained using the fuzzy clustering (Approach 4) data division method are shown in Table 5.28. The t- and F-tests (Table 5.29) of the data indicate that the three data sets may be considered to be representative of each other. The performance of the ANN model (model CHP5-NF55) developed using the fuzzy clustering data division method is shown in Table 5.30. It should be noted that this model is developed using two hidden layer nodes, a learning rate of 0.2, momentum term of 0.8, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer. It can be seen from Table 5.30 that the model performs well with respect to all data sets (i.e. training, testing and validation).

Table 5.23: Input and output statistics for the first approach of SOM data division

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Footing width, *B* (m)					
Training set	7.9	9.0	0.8	60.0	59.2
Testing set	10.8	13.1	0.9	55.0	54.1
Validation set	8.8	8.8	1.1	33.5	32.4
Footing net applied pressure, *q* (kPa)					
Training set	184.6	119.0	18.3	697.0	678.7
Testing set	204.6	133.9	52.0	666.0	614.0
Validation set	170.8	122.3	25.0	584.0	559.0
Average SPT blow count, *N*					
Training set	24.0	12.8	4.0	60.0	56.0
Testing set	26.3	15.4	5.0	60.0	55.0
Validation set	24.0	13.0	6.0	50.0	44.0
Footing geometry, *L/B*					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.1	1.9	1.0	9.9	8.9
Validation set	2.2	1.7	1.0	7.8	6.8
Footing embedment ratio, D_f/B					
Training set	0.57	0.6	0.0	3.4	3.4
Testing set	0.49	0.5	0.0	2.1	2.1
Validation set	0.42	0.4	0.0	1.8	1.8
Measured settlement, S_m (mm)					
Training set	18.7	24.5	0.6	121.0	120.4
Testing set	22.7	27.4	1.3	116.0	114.7
Validation set	23.0	31.8	1.0	120.0	119.0

Table 5.24: Input and output statistics for the second approach of SOM data division

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Footing width, B (m)					
Training set	8.5	10.6	0.8	60.0	59.2
Testing set	10.8	13.1	0.9	55.0	54.1
Validation set	8.8	8.8	1.1	33.5	32.4
Footing net applied pressure, q (kPa)					
Training set	185.2	123.2	18.3	697.0	678.7
Testing set	204.6	133.9	52.0	666.0	614.0
Validation set	170.8	122.3	25.0	584.0	559.0
Average SPT blow count, N					
Training set	23.6	14.4	4.0	60.0	56.0
Testing set	26.3	15.4	5.0	60.0	55.0
Validation set	24.0	13.0	6.0	50.0	44.0
Footing geometry, L/B					
Training set	2.1	1.9	1.0	10.5	9.5
Testing set	2.1	1.9	1.0	9.9	8.9
Validation set	2.2	1.7	1.0	7.8	6.8
Footing embedment ratio, D_f/B					
Training set	0.53	0.65	0.0	3.4	3.4
Testing set	0.49	0.5	0.0	2.1	2.1
Validation set	0.42	0.43	0.0	1.8	1.8
Measured settlement, S_m (mm)					
Training set	22.1	29.9	0.6	121.0	120.4
Testing set	22.7	27.4	1.3	116.0	114.7
Validation set	23.0	31.8	1.0	120.0	119.0

Table 5.25: Null hypothesis tests for the first approach of SOM data division

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–1.64	–1.97	1.97	Accept	0.47	0.62	1.68	Reject
Validation	–0.56	–1.97	1.97	Accept	1.05	0.59	1.84	Accept
q								
Testing	–0.92	–1.97	1.97	Accept	0.79	0.62	1.69	Accept
Validation	0.58	–1.97	1.97	Accept	0.95	0.59	1.83	Accept
N								
Testing	–0.96	–1.97	1.97	Accept	0.69	0.62	1.69	Accept
Validation	0.00	–1.97	1.97	Accept	0.97	0.59	1.83	Accept
L/B								
Testing	0.00	–1.97	1.97	Accept	0.80	0.62	1.69	Accept
Validation	–0.29	–1.97	1.97	Accept	0.99	0.59	1.83	Accept
D_f/B								
Testing	0.76	–1.97	1.97	Accept	1.99	0.62	1.69	Reject
Validation	1.24	–1.97	1.97	Accept	2.28	0.59	1.83	Reject
S_m								
Testing	–0.92	–1.97	1.97	Accept	0.79	0.62	1.69	Accept
Validation	–0.84	–1.97	1.97	Accept	0.59	0.59	1.83	Reject

Table 5.26: Null hypothesis tests for the second approach of SOM data division

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.97	–1.98	1.98	Accept	0.65	0.57	1.77	Accept
Validation	0.14	–1.98	1.98	Accept	1.45	0.57	1.77	Accept
q								
Testing	–0.75	–1.98	1.98	Accept	0.85	0.57	1.77	Accept
Validation	0.53	–1.98	1.98	Accept	1.01	0.57	1.77	Accept
N								
Testing	–0.91	–1.98	1.98	Accept	0.87	0.57	1.77	Accept
Validation	–0.13	–1.98	1.98	Accept	1.22	0.57	1.77	Accept
L/B								
Testing	0.00	–1.98	1.98	Accept	0.98	0.57	1.77	Accept
Validation	–0.25	–1.98	1.98	Accept	1.25	0.57	1.77	Accept
D_f/B								
Testing	0.35	–1.98	1.98	Accept	1.69	0.57	1.77	Accept
Validation	0.86	–1.98	1.98	Accept	2.30	0.57	1.77	Reject
S_m								
Testing	–0.10	–1.98	1.98	Accept	1.19	0.57	1.77	Accept
Validation	–0.13	–1.98	1.98	Accept	0.88	0.57	1.77	Accept

Table 5.27: Performance of ANN models using different approaches of data division for SOM

Performance measures and data sets	SOM (first approach) Model CHP5-NF53	SOM (second approach) Model CHP5-NF54
Training		
Correlation coefficient, r	0.890	0.931
RMSE (mm)	11.58	11.80
MAE (mm)	7.93	9.19
Testing		
Correlation coefficient, r	0.942	0.928
RMSE (mm)	10.43	14.43
MAE (mm)	7.98	10.21
Validation		
Correlation coefficient, r	0.958	0.960
RMSE (mm)	10.12	12.07
MAE (mm)	7.12	9.32

Table 5.28: Input and output statistics obtained using fuzzy clustering

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Minimum	Maximum	Range
Footing width, B (m)					
Training set	8.7	10.1	0.8	60.0	59.2
Testing set	8.6	10.1	1.0	42.7	41.7
Validation set	9.2	10.6	1.2	36.6	35.4
Footing net applied pressure, q (kPa)					
Training set	180.2	120.0	18.3	697.0	678.7
Testing set	209.4	134.4	64.0	584.0	520.0
Validation set	207.0	164.4	47.6	584.0	536.4
Average SPT blow count, N					
Training set	24.6	13.2	4.0	60.0	56.0
Testing set	23.4	12.4	6.0	50.0	44.0
Validation set	25.5	16.8	5.0	60.0	55.0
Footing geometry, L/B					
Training set	2.2	1.9	1.0	10.5	9.5
Testing set	2.0	1.5	1.0	6.7	5.7
Validation set	1.7	1.1	1.0	5.2	4.2
Footing embedment ratio, D_f/B					
Training set	0.54	0.61	0.0	3.4	3.4
Testing set	0.50	0.40	0.0	1.4	1.4
Validation set	0.50	0.51	0.0	2.1	2.1
Measured settlement, S_m (mm)					
Training set	20.0	25.8	0.6	121.0	120.4
Testing set	21.5	26.5	2.1	87.0	84.9
Validation set	22.1	32.5	1.3	120.0	118.7

Table 5.29: Null hypothesis tests for data division 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*
B								
Testing	0.05	–1.97	1.97	Accept	0.65	0.57	1.99	Accept
Validation	–0.21	–1.97	1.97	Accept	1.45	0.55	1.99	Accept
q								
Testing	1.10	–1.97	1.97	Accept	0.85	0.57	1.99	Accept
Validation	–0.91	–1.97	1.97	Accept	1.01	0.55	1.99	Reject
N								
Testing	0.42	–1.97	1.97	Accept	0.87	0.57	1.99	Accept
Validation	–0.28	–1.97	1.97	Accept	1.22	0.55	1.99	Accept
L/B								
Testing	0.50	–1.97	1.97	Accept	0.98	0.57	1.99	Accept
Validation	1.18	–1.97	1.97	Accept	1.25	0.55	1.99	Reject
D_f/B								
Testing	0.32	–1.97	1.97	Accept	1.69	0.57	1.99	Reject
Validation	0.29	–1.97	1.97	Accept	2.30	0.55	1.99	Accept
S_m								
Testing	–0.27	–1.97	1.97	Accept	1.19	0.57	1.99	Accept
Validation	–0.34	–1.97	1.97	Accept	0.88	0.55	1.99	Accept

Table 5.30: performance of ANN model (Model CHP5-NF55) using fuzzy clustering data division method

Performance measure	Data set		
	Training	Testing	Validation
Correlation coefficient, r	0.912	0.967	0.957
RMSE (mm)	10.62	10.48	9.59
MAE (mm)	7.43	6.92	6.13

The success of the SOM and fuzzy clustering data division methods is illustrated in Table 5.21 (columns 4 and 5), which compares the predictive results obtained using the four different approaches to data division investigated. It can be seen that the results obtained for the SOM (Approach 3) and fuzzy clustering (Approach 4) data division methods are very close to those obtained for the statistically consistent data division method (Approach 2) and significantly better than the results obtained for the purely random data division method (Approach 1). It should be noted that the results presented for the data division method that takes into account the statistical properties of the subsets (Approach 2) are for the proportion of training, testing and validation data that gives the best results. Consequently, it appears as though the SOM and fuzzy clustering methods are suitable approaches for dividing data into training, testing and validation subsets. However, as discussed previously, fuzzy clustering data division has the advantage over SOM data division that an optimum number of clusters can be obtained analytically and, consequently, the fuzzy clustering data division approach removes the subjectivity associated with the SOM data division method.

The results of a parametric study carried out to examine the robustness of the ANN models developed using the three successful data division methods (i.e. Approaches 2, 3 and 4) are shown in Figure 5.12. It can be seen that there is a slight difference in settlement prediction among the models developed using the three data division methods. However, this is to be expected, as the optimisation of the three models is based on different initial weights and calibration data sets. In general, the performance of the three models is in good agreement and may be considered to be similar. The direction of the trends illustrated in Figure 5.12 also indicates that the behaviour of the models developed is similar to what one would expect based on a physical understanding of settlement prediction. This indicates that the three models are robust and could be used for predictive purposes with confidence.

It is evident from the previous data division analyses that the performance of the ANN models developed using the three successful data division methods is very similar (see Table 5.21). However, it is also evident that the performance of the model that used data division method based on statistical consistency (model CHP5-NF2) is slightly better. Consequently, this model (i.e. model CHP5-NF2) will be used for the remaining analyses in this chapter and will henceforth be referred to as the ANN model.

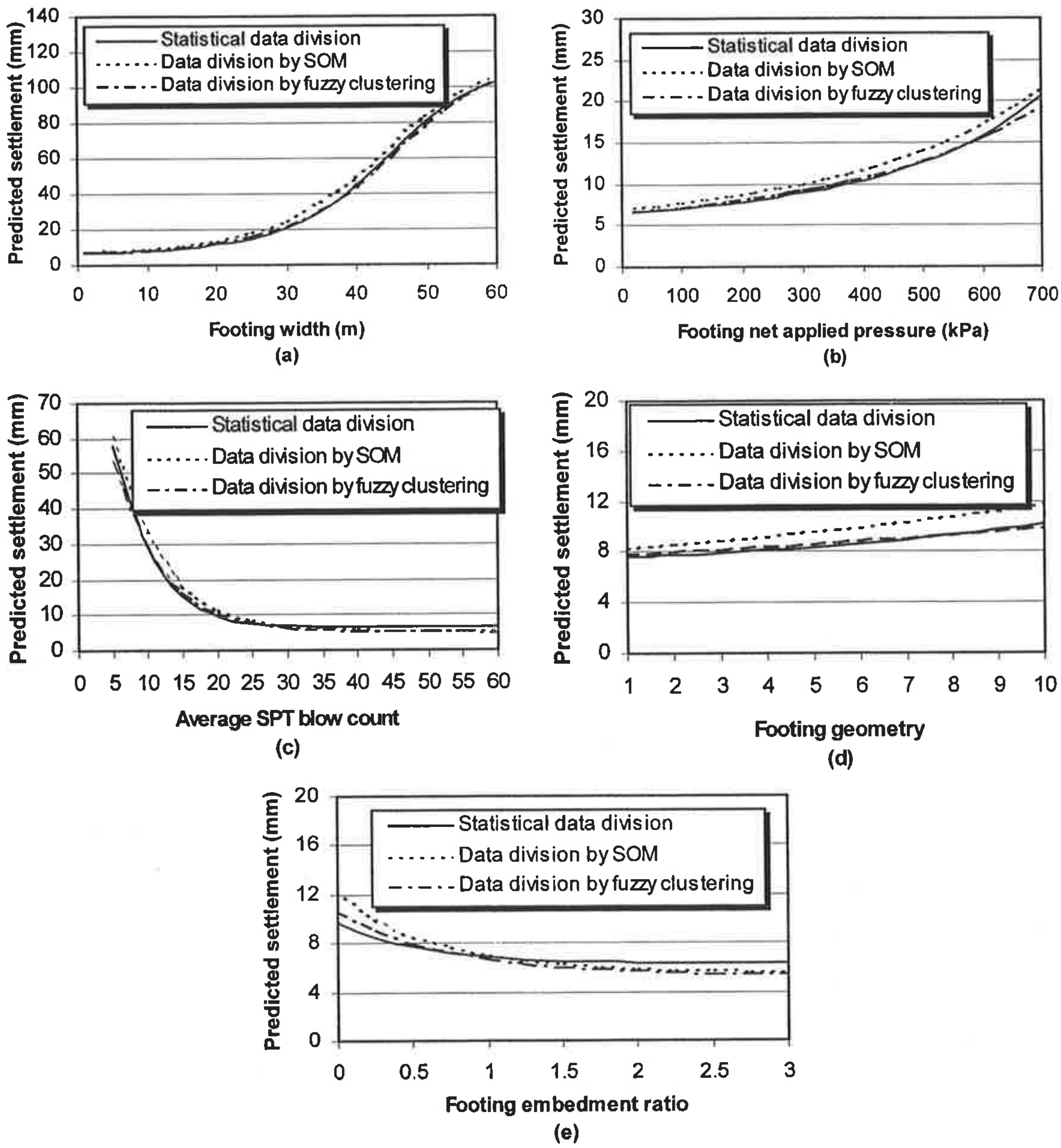


Figure 5.12: Results of parametric study of ANN models using data subsets obtained for different approaches of data division

5.4 Data Transformation of ANN Model Inputs

The main purpose of data transformation is to modify the distribution of the input variables before they are applied to ANNs in order to provide a better mapping to the outputs. It has been suggested that certain transformations of input variables may help to improve the performance of ANN models (Fortin et al. 1997; Shi 2000). However, it has also been shown that input data transformation does not affect the performance of the ANN models in any way (Faraway and Chatfield 1998).

Recently, Shi (2000) proposed a new data transformation method called *distribution transformation* and found that the method reduced the prediction error for a cowboy hat surface by 50%. Shi (2000) also found that the method succeeded in reducing the error of predictions of settlement of tunnels by more than 13%. Consequently, it is worthwhile to apply the distribution transformation method to the ANN model developed in this research. The method transforms a stream of random data distributed in any range to uniformly distributed data points between 0.0 and 1.0. The method requires a probability distribution function to be fitted to each of the input variables and by using the relationship between the probability distribution and cumulative distribution functions, any distribution in any range can be transformed to a uniform distribution between 0.0 and 1.0 (Shi 2000).

The ANN model (model CHP5-NF2) is redeveloped to incorporate the distribution transformation of input variables. The network architecture and internal parameters are maintained. For each input variable, the training data set is used to obtain the distribution transformation function. The software @Risk (Palisade 2000) is used for this purpose. For a set of data values, the software can define the probability distribution function that best fits these values from 38 candidate distributions and provides statistical parameters that describe the distribution. The theoretical distributions that are found to best match the actual distribution of the available data for the ANN model input variables are shown in Table 5.31.

Using the fitted distribution functions shown in Table 5.31, the original data are transformed to uniformly distributed data points between 0.0 and 1.0. The ANN model (model CHP5-NF2) is re-trained using the transformed data and the performance is shown in Table 5.32, which also includes the performance of the ANN model without distribution transformation (i.e. linear transformation or scaling).

Table 5.31: Fitted distributions of input variables for Model CHP5-NF2

Input variable	Fitted distribution	Statistical parameters				
		α	β	γ	$\bar{x}$	λ
B	Exponential	N/A	7.5	N/A	N/A	N/A
q	Inverse-gamma	5.1	963.7	N/A	N/A	N/A
N	Log-logistic	3.7	26.6	-4.8	N/A	N/A
L/B	Exponential	N/A	1.1	N/A	N/A	N/A
D_f/B	Inverse-gaussian	N/A	N/A	N/A	0.65	0.74

α = shape parameter for log-logistic distribution
β = scale parameter for inverse-gamma or decay constant for exponential distributions
γ = location parameter for log-logistic distribution
$\bar{x}$ = mean
λ = shape parameter for Inverse-gaussian distribution
N/A = not applicable

Table 5.32: Performance of Model CHP5-NF2 using linear and distribution transformations of input variables

Performance Measures	Linear transformation			Distribution transformation		
	Training	Testing	Validation	Training	Testing	Validation
r	0.930	0.929	0.905	0.722	0.832	0.616
RMSE (mm)	10.01	10.12	11.04	20.14	15.51	22.23
MAE (mm)	6.87	6.43	8.78	13.75	10.89	15.34

It can be seen from Table 5.32 that the performance of the ANN model on the training, testing and validation sets is significantly worse when distribution transformation of the input variables is used. This may be because of the distortion that might have occurred to the original relationships between the ANN model inputs and the corresponding output as a result of data transformation. This finding adds further weight to the argument that data used by ANN models do not need to be transformed.

5.5 Sensitivity Analysis of the ANN Model Inputs

In an attempt to identify which of the input variables have the most significant impact on settlement predictions, a sensitivity analysis is carried out on the ANN model (model CHP5-NF2). A simple and innovative technique proposed by Garson (1991) is used to interpret the relative importance of the input variables by examining the connection weights of the trained network. For a network with one hidden layer, the technique involves a process of partitioning the hidden output connection weights into components associated with each input node. For model CHP5-NF2, the method is illustrated as follows. The model has five input nodes, two hidden nodes, and one output node with connection weights as follows:

Hidden nodes	Weights					
	(B)	(q)	(N)	(L/B)	(D_f/B)	(S_m)
Hidden 1	0.227354	0.481161	0.229593	−0.017031	0.067341	0.725351
Hidden 2	−2.442513	−1.114891	4.239639	−0.498853	2.500301	−2.984165

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

1. For each hidden node i, obtain the products P_{ij} (where j represents the column number of the weights mentioned above) by multiplying 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. As an example: $P_{11} = 0.227354 \times 0.725351 = 0.164911$.

	(B)	(q)	(N)	(L/B)	(D_f/B)
Hidden 1	0.164911	0.349011	0.166536	0.012353	0.048846
Hidden 2	7.288861	3.327017	12.65178	1.488659	7.461311

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

$$Q_{11} = 0.164911/(0.164911 + 0.349011 + 0.166536 + 0.012353 + 0.048846) = 0.222355$$

	(B)	(q)	(N)	(L/B)	(D_f/B)
Hidden 1	0.222355	0.470582	0.224545	0.016656	0.065859
Hidden 2	0.226238	0.103267	0.392698	0.046206	0.231591

3. For each input node, sum Q_{ij} to obtain S_j. As an example: $S_1 = 0.222355 + 0.226238 = 0.448593$.

	(B)	(q)	(N)	(L/B)	(D_f/B)
Sum	0.448593	0.573849	0.617243	0.062863	0.297451

4. Divide S_j by the sum for all input variables to get the relative importance of all output weights attributed to the given input variable. As an example, the relative importance for input node 1 is equal to:

$$(0.448593 \times 100)/(0.448593 + 0.573849 + 0.617243 + 0.062863 + 0.297451) = 22.4\%$$

	(B)	(q)	(N)	(L/B)	(D_f/B)
Relative importance (%)	22.4	28.6	30.8	3.1	14.8

The results indicate that N has the most significant effect on the predicted settlement followed by q and B, each with a relative importance of 30.8, 28.6 and 22.4%, respectively. The results also indicate that D_f/B has a moderate impact on settlement with a relative importance equals to 14.8%, while L/B has the smallest impact on settlement with 3.1% relative importance. The above results indicate that N, q and B are the most important factors affecting settlement, whereas the effect of L/B and D_f/B may be considered secondary, which agrees well with the discussion in Chapter 4.

5.6 Comparison of ANN Model with Traditional Methods

Comparisons of the results of the validation set obtained using the ANN model (model CHP5-NF2) and the three traditional methods described in Chapter 4 are carried out and presented in Table 5.33 and Figure 5.13. Table 5.33 shows that the ANN method performs better than the traditional methods for all three performance measures considered. The coefficient of correlation, r, the RMSE and MAE obtained using the ANN model are: 0.905, 11.04 mm and 8.78 mm, respectively. In contrast, these measures range from 0.440 to 0.798, from 23.55 mm to 25.72 mm and from 11.81 mm to 16.69 mm, respectively, when the traditional methods are used. Figure 5.13 shows that the ANN model performs reasonably well for the full range of measured settlements considered. In contrast, the traditional methods appear to work well only for small settlements, in the range of 10 – 20 mm. The method of Schmertmann et al. (1978) tends to overpredict larger settlements, the method of Schultze and Sherif (1973) tends to severely underpredict larger settlements and the method of Meyerhof (1965) appears to both over-and under-predict larger settlements, although all settlements in excess of 60 mm are generally under-predicted. It is evident from the above results that ANNs provide more accurate settlement predictions than the traditional methods which indicates that ANNs succeeded to overcome the limitations discussed in Chapter 4 for the traditional methods considered for comparison. This can be attributed to the fact that ANNs are a data driven approach in which the data alone are used to capture the relationship between settlements and the factors affecting them. This appears to result in a more reliable relationship between settlements and the factors affecting them, especially when the theory that governs this relationship is uncertain.

Table 5.33: ANN and traditional methods for settlement prediction

Performance measure	ANN	Meyerhof (1965)	Schultze and Sherif (1973)	Schmertmann et al. (1978)
r	0.905	0.440	0.729	0.798
RMSE	11.04	25.72	23.55	23.67
MAE	8.78	16.59	11.81	15.69

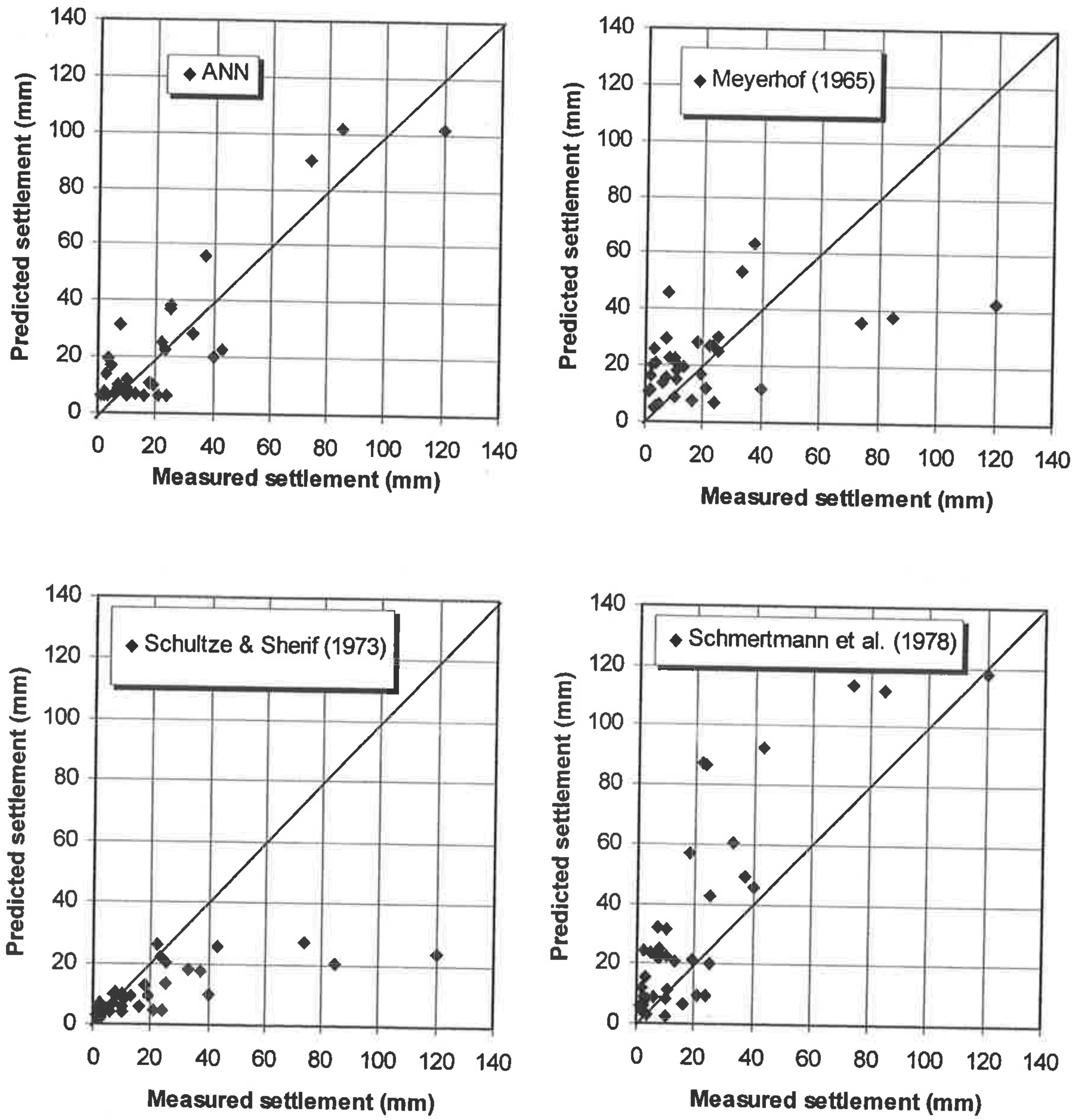


Figure 5.13: Measured vs predicted settlement for ANN and traditional methods

5.7 ANN Model Equation and Design Charts

The small number of connection weights obtained for the optimal ANN model (model CHP5-NF2) enables the network to be translated into a relatively simple formula. To demonstrate this, the structure of the ANN model is shown in Figure 5.14, while its connection weights and threshold levels are summarised in Table 5.34.

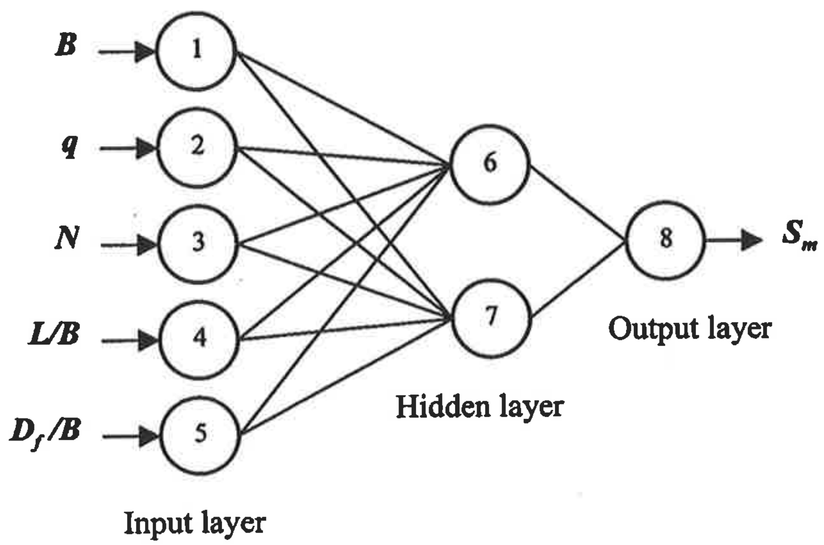


Figure 5.14: Structure of the ANN optimal model

Table 5.34: weights and threshold levels for the ANN optimal model

Hidden layer nodes	w_{ji} (weight from node i in the input layer to node j in the hidden layer)					Hidden layer threshold (θ_j)
	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	
$j=6$	0.227	0.481	0.229	−0.017	0.067	0.124
$j=7$	−2.442	−1.114	4.239	−0.498	2.500	0.188
Output layer nodes	w_{ji} **(weight from node i in the hidden layer to node j in the output layer)**					**Output layer threshold (θ_j)**
	$i=6$	$i=7$	-	-	-	
$j=8$	0.725	−2.984	-	-	-	−0.312

Using the connection weights and threshold levels shown in Table 5.34, the predicted settlement can be expressed as follows:

$$S_p = \frac{1}{1+e^{(0.3-0.7\tanh x_1+2.9\tanh x_2)}} \tag{5.7}$$

where:

$$x_1 = \theta_6 + w_{61}B + w_{62}q + w_{63}N + w_{64}(L/B) + w_{65}(D_f/B) \tag{5.8}$$

and

$$x_2 = \theta_7 + w_{71}B + w_{72}q + w_{73}N + w_{74}(L/B) + w_{75}(D_f/B) \tag{5.9}$$

It should be noted that, before using Equations 5.8 and 5.9, all input variables (i.e. B, q, N, L/B and D_f/B) need to be scaled between 0.0 and 1.0 using Equation 5.1 and the data ranges in the ANN model training (see Table 5.1). It should also be noted that the predicted settlement obtained from Equation 5.7 is scaled between 0.0 and 1.0 and in order to obtain the actual value, this settlement has to be re-scaled using Equation 5.1 and the data ranges in Table 5.1. The procedure for scaling and substituting the values of the weights and threshold levels from Table 5.34, Equations 5.7, 5.8 and 5.9 can be rewritten as follows:

$$S_p = 0.6 + \left[\frac{120.4}{1+e^{(0.312-0.725\tanh x_1+2.984\tanh x_2)}}\right] \tag{5.10}$$

and

$$x_1 = 0.1 + 10^{-3}\left[3.8B + 0.7q + 4.1N - 1.8(L/B) + 19(D_f/B)\right] \tag{5.11}$$

$$x_2 = 10^{-3}\left[0.7 - 41B - 1.6q + 75N - 52(L/B) + 740(D_f/B)\right] \tag{5.12}$$

where:

S_p = predicted settlement (mm);

B = footing width (m);

q = footing net applied pressure (kPa);

N = average SPT blow count;

L/B = footing geometry; and

D_f/B = footing embedment ratio.

It should be noted that Equation 5.10 is valid only for the ranges of values of B, q, N, L/B and D_f/B given in Table 5.1. This is due to the fact that ANNs perform best in interpolation and not extrapolation (Flood and Kartam 1994; Minns and Hall 1996; Tokar and Johnson 1999). An executable computer program of the optimal ANN model is also provided for routine work in practice and the FORTRAN code for the program is given in Appendix F.

A numerical example is provided to better explain the implementation of the settlement formula. A rectangular footing whose dimensions are 2.5 × 4.0 m is founded at a depth equal to 1.5 m below the ground surface. The soil beneath the footing is sand that extends to a depth in excess of twice its width. The net applied load exerted on the footing is 350 kPa and the average SPT blow count over a depth of twice its width is 16.

Solution:

Given the information provided, $B = 2.5$ m; $L = 4.0$ m; $q = 350$ kPa; $N = 16$ and $D_f = 1.5$ m.

From Equation 5.11:

$$x_1 = 0.1 + 10^{-3}\left[3.8 \times 2.5 + 0.7 \times 350 + 4.1 \times 16 - 1.8\left(\frac{4.0}{2.5}\right) + 19\left(\frac{1.5}{2.5}\right)\right] = 0.4286$$

From Equation 5.12:

$$x_2 = 10^{-3}\left[0.7 - 41 \times 2.5 - 1.6 \times 350 + 75 \times 16 - 52\left(\frac{4.0}{2.5}\right) + 740\left(\frac{1.5}{2.5}\right)\right] = 0.8990$$

By substituting x_1 and x_2 in Equation 5.10, the predicted settlement can be obtained as follows:

$$S_p = 0.6 + \left[\frac{120.4}{1 + e^{(0.312 - 0.725 \tanh 0.4286 + 2.984 \tanh 0.8990)}}\right] = 13.2 \text{ mm}$$

In order to facilitate the ANN technique for settlement prediction of shallow foundations on cohesionless soils, the information obtained from the ANN model is translated into a set of design charts suitable for practical use in order to avoid computer or hard calculations. This is carried out by entering synthetic data into the trained ANN model such that the synthetic data lie within the ranges of the data used during the ANN model development. A series of design charts are generated and are summarised in Appendix G.

Figure 5.15 is an illustrative example of the design charts obtained for $L/B = 1.0$ and $D_f/B = 0.0$. It can be seen that, for each graph and at a certain footing net applied pressure, settlement increases as the footing width increases, as expected. It can also be seen that, for each graph and at a certain footing width, settlement increases as the footing net applied pressure increases, also as expected. On the other hand, moving from one graph to another and at the same footing width and footing net applied pressure, the settlement decreases as the SPT blow count increases, again, as expected. These results add more confirmation to the robustness and credibility of the ANN model for settlement prediction. The design charts provide a simple and quick tool of estimating settlement for general use in practice. However, when greater precision is required, the computer model (Appendix F) or formula (Equations 5.10 to 5.12) can be used.

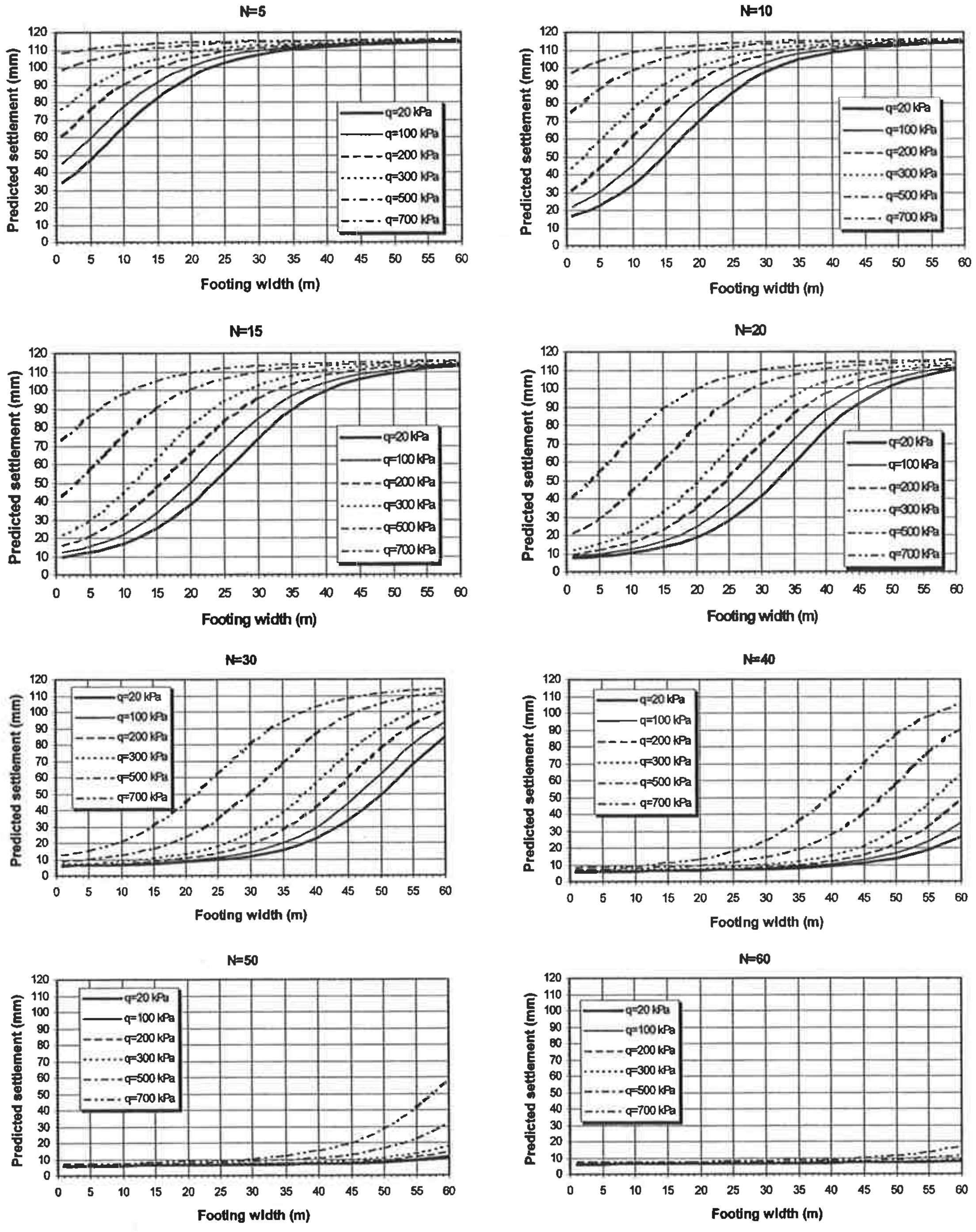


Figure 5.15: Illustrative set of design charts based on the ANN model

(L/B = 1.0 and D_f/B = 0.0)

5.8 Summary and Conclusions

Multi-layer perceptrons (MLPs) trained with the back-propagation algorithm were used to demonstrate the feasibility of employing artificial neural networks (ANNs) to predict the settlement of shallow foundations on cohesionless soils. A database containing 189 case records of actual field measurements for settlement of shallow foundations on cohesionless soils was compiled and used for ANN model development and verification. The use of parametric studies was presented as a way of testing the generalisation ability and robustness of ANN models. The effect of the number and type of connection weights, data noise and software implementation on the robustness of ANN models was investigated. The effect of using various learning rates, momentum terms, transfer functions and initial weights on the results of ANN models was also investigated. The effect of various data division methods on the performance of ANN models was examined and a new approach for data division was presented and evaluated. The effect of data transformation of the input variables on the performance of ANN models was also examined. A sensitivity analysis was carried out on the ANN model to study the relative importance of the factors that affect settlement. The results between the predicted and measured settlements obtained using ANNs were compared with those obtained using the traditional methods. Finally, the ANN model was translated into a relatively simple practical equation and a series of design charts from which the settlement of shallow foundations on cohesionless soils can be easily obtained.

The analyses carried out in this chapter have yielded the following results and conclusions:

1. ANNs have the ability to predict the settlement of shallow foundations on cohesionless soils with a high degree of accuracy for predicted settlements ranging from 0.6 to 121.0 mm.

2. The results of the robustness studies carried out in this chapter lead to the following findings:

- Good performance of ANN models on training, testing and validation sets does not guarantee the robustness of the predictive ability of the models over a range of data similar to that used for training.
- If cross-validation is used as the stopping criterion, reducing the number of ANN connection weights or changing the type of connection between nodes (e.g. cascaded and direct connections) does not appear to improve their robustness.
- The results of the parametric study carried out on both noise-free and noisy hypothetical data suggest that the degree of noise in the data used to calibrate ANN models may affect their robustness. However, ANN models trained with noisy hypothetical data using two different commercial software systems [i.e. NeuralWorks *Predict* release 2.1 (NeuralWare 1997) and *Neuframe* version 4.0 (Neusciences 2000)] lead to the following conclusion. ANN models developed using *Predict* failed to interpret the underlying physical meaning of the relationships between settlements and the factors affecting them, whereas ANN models developed using *Neuframe* succeeded in interpreting these relationships. This indicates that it is the software used, rather than the degree of noise in the data, that affects ANN models robustness. As some geotechnical engineering researchers have used, and continue to use *Predict*, one must question the validity of the subsequent models.
- It is recommended that the capability of the software used for developing ANN models be considered carefully and that a parametric study, such as the one presented in this chapter, be conducted in order to ensure that a model can be used for predictive purposes with confidence.

3. The ANN models developed in this chapter to study the impact of the internal network parameters on model performance indicate that ANN performance is relatively insensitive to the number of hidden layer nodes, momentum term or transfer functions. On the other hand, the impact of the learning rate on model predictions is more pronounced, with step sizes that are too small or too large resulting in reduced model performance. Overall, the optimum model (model CHP5-NF2) is obtained using 2 hidden layer nodes, a momentum term of 0.8, learning rate of 0.2, tanh transfer function in the hidden layer and sigmoid transfer function in the output layer.

4. There is a direct relationship between the consistency of the statistics between the training, testing and validation sets and the consistency in model performance. Consequently, the statistical properties of the various data subsets should be taken into account as part of any data division procedure to ensure that the best possible model is developed, given the available data.

5. The proportion of the data used for training, testing and validation appears to have an effect on model performance. However, there appears to be no clear relationship between the proportion of the data used in each of the subsets and model performance, although in the trials conducted, the optimal model performance was obtained when 20% of the data are used for validation and 70% of the remaining data were used for training and 30% for testing.

6. The data division approaches using a SOM and the proposed approach using fuzzy clustering appear to be applicable, as they have a number of advantages, including:
   - There is no need to decide which proportion of the available data to use for training, testing and validation.
   - The statistical properties of the resulting training, testing and validation data are similar, provided that the clusters are small enough.
   - Information is provided about where extreme values exist in the data set and consequently, they can be included in the training set. If they were to be included in the validation set, the trained ANN model could not be expected to perform well, as the validation data would fall outside the range of the training data. A potential disadvantage of the SOM approach is that the parameters that control the learning process need to be selected, potentially affecting the results obtained. This disadvantage is overcome using the fuzzy clustering technique.

7. The distribution transformation method for input variables does not appear to improve the performance of ANN models.

8. The sensitivity analysis indicated that the SPT blow count, the footing net applied pressure and the footing width are the most important factors affecting settlement, each with a relative importance of 30.8, 28.6 and 22.4%, respectively. The footing

embedment ratio and the footing geometry have less impact on settlement with a relative importance equal to 14.8 and 3.1%, respectively.

9. The ANN method outperforms the traditional methods considered for an independent validation set with $r = 0.905$, RMSE = 11.04 mm and MAE = 8.78 mm, while these measures were: $r = 0.440$, 0.729 and 0.798; RMSE = 25.72, 23.55 and 23.67 mm and MAE = 16.59, 11.81 and 15.69 mm when the method proposed by Meyerhof (1965), Schultze and Sherif (1973) and Schmertmann et al. (1978) are used, respectively.

10. Due to its parsimonious nature, the ANN model was able to be translated into a simple and practical formula from which settlement may be calculated, as shown in Equation 5.10. In addition, the ANN model was translated into a computer program (Appendix F) and a series of design charts (Appendix G), facilitating settlement prediction.

In the following chapter, neurofuzzy networks will be examined to investigate their ability to predict settlement of shallow foundations on cohesionless soils and to assist with providing a better understanding of the relationship between settlement and the factors affecting it.

Chapter 6

Settlement Prediction by Neurofuzzy Networks

6.1 Introduction

As mentioned in Chapter 2, neurofuzzy networks can be trained to provide input/output data mappings and to extract knowledge regarding the relationships between model inputs and the corresponding outputs. Neurofuzzy networks enable the knowledge that has been learnt in the network to be expressed in the form of a fuzzy rule base. Ni et al. (1996) have already applied a neurofuzzy network approach in geotechnical engineering to the evaluation of slope failure potential. However, a review of the literature indicates that neurofuzzy networks are new tools in the field of geotechnical engineering. In this chapter, the feasibility of adopting neurofuzzy networks for predicting the settlement of shallow foundations on cohesionless soils is tested. In addition, the ability of neurofuzzy networks to assist with providing a better understanding of the relationship between settlement and the factors affecting settlement is investigated.

6.2 Development of Neurofuzzy Models

The database used for the development of the multi-layer perceptron (MLP) models in Chapter 5 is also used to develop the neurofuzzy models in this chapter. The type of neurofuzzy network that is used in this work is the B-spline network trained with the adaptive spline modelling of observation data (ASMOD) algorithm described in Chapter 2. The software *Neuframe* Version 4.0 (Neusciences 2000) is used to simulate B-spline neurofuzzy network operation. The criteria adopted in Chapter 5 for choosing the MLP model inputs and outputs are considered for the development of neurofuzzy models in this chapter. As a consequence, five input variables are used as potential neurofuzzy model inputs. These include the footing width (B), footing net applied pressure (q), average SPT blow count (N) as a measure of soil compressibility (or density), footing geometry (L/B) and footing embedment ratio (D_f/B). The only output variable is the measured settlement (S_m). As mentioned in Chapter 2, the ASMOD algorithm automatically optimises model architecture and selects the input variables that have the most significant impact on model outputs. The ASMOD algorithm also uses stopping criteria [e.g. Bayesian Information Criterion (BIC), Akaike's Information criteria (AIC) and Final Prediction Error (FPE)] that require the data to be divided into two sets. A training set to build the model and an independent validation set to test the predictive ability of the model in real-world situations. The training and testing sets used in Chapter 5 to develop the optimum MLP model (i.e. model CHP5-NF2) are combined to form the training set for neurofuzzy networks, whereas the validation set is kept the same and thus, a fair comparison between neurofuzzy and MLP models can be carried out. Using this procedure, 152 (80%) of the available data records are used for training and 37 (20%) are used for validation. In an attempt to obtain an optimum neurofuzzy model, the BIC, AIC and FPE stopping criteria are examined and the results are given below.

6.3 Results and Discussion

A summary of the structure of the neurofuzzy models developed in this chapter, and the number of fuzzy rules produced for each model, is given in Table 6.1. The performance results of the models obtained are given in Table 6.2. A code is used to identify the names of the different models developed. The code consists of two parts separated by a hyphen. The first part is an abbreviation that denotes the current chapter, whereas the second part is an abbreviation that denotes the stopping criterion used. It can be seen from Table 6.1 that all models select only three input variables (i.e. B, q and N) as the most significant inputs, whereas the footing geometry (L/B) and footing embedment ratio (D_f/B) are not selected in any model. This is in agreement with the results of the sensitivity analysis carried out in Chapter 5 on the optimum MLP model (model CHP5-NF2). The sensitivity analysis in Chapter 5 indicated that B, q and N have the most significant impact on settlement, D_f/B has a moderate impact on settlement and L/B has the smallest impact on settlement. The neurofuzzy models obtained in this chapter are assessed in terms of prediction accuracy, model parsimony and model transparency and the optimum model is described in more detail in §6.4. The following conclusions can be drawn:

Table 6.1: Summary of the neurofuzzy models developed

Model No.	No. of significant inputs	Significant inputs	No. of connection weights	No. of fuzzy rules
CHP6-BIC	3	B, q, N	8	10
CHP6-AIC	3	B, q, N	7	8
CHP6-FPE	3	B, q, N	8	16

Table 6.2: Performance of the neurofuzzy models developed

Model No.	Performance measures					
	Correlation coefficient, r		RMSE (mm)		MAE (mm)	
	Training	Validation	Training	Validation	Training	Validation
CHP6-BIC	0.889	0.881	12.33	12.36	8.08	9.36
CHP6-AIC	0.879	0.863	12.82	13.37	8.29	10.10
CHP6-FPE	0.910	0.875	11.16	13.00	6.82	9.49

- In terms of prediction accuracy, all models are comparable, although model CHP6-BIC performs slightly better than the other models with respect to the validation set (Table 6.2).
- In terms of model parsimony, all models are comparable, even though model CHP6-AIC is found to be more parsimonious than the other models, as it has the lowest number of connection weights (Table 6.1). This is because the AIC penalises complex models.
- In terms of model transparency, models CHP6-BIC and CHP6-AIC are comparable, even though model CHP6-AIC is found to be more transparent, as it describes the relationship between model inputs and outputs using a smaller number of fuzzy rules (Table 6.1). On the other hand, model CHP6-FPE is found to be the worst model in terms of model transparency as it has the highest number of fuzzy rules (Table 6.1), which is almost twice that obtained for models CHP6-BIC and CHP6-AIC. This is because the FPE stopping criterion does not penalise larger models as much as the BIC and AIC. The BIC and AIC stopping criteria penalise complex models to ensure that more parsimonious models are chosen (Schwarz 1978).

It can be seen from Tables 6.1 and 6.2 that model CHP6-BIC is able to strike a balance between model accuracy and model parsimony and transparency. In terms of model accuracy, the performance of model CHP6-BIC is slightly better than the other models with respect to the validation set (Table 6.2). In terms of model parsimony and transparency, model CHP6-BIC has as the same number of input variables as models CHP6-AIC and CHP6-FPE combined, with a number of connection weights and fuzzy rules that are approximately equal to the average of those obtained for models CHP6-AIC and CHP6-FPE (Table 6.1). Overall, model CHP6-BIC can be considered to be optimal.

6.4 Description of the Optimum Neurofuzzy Model

A schematic view of the optimum neurofuzzy model (i.e. model CHP6-BIC) is given in Figure 6.1. It can be seen that the model uses only 3 of the 5 potential input variables as the most significant inputs. The chosen inputs are the footing width (B), footing net applied pressure (q) and the average SPT blow count (N) as a measure of soil density. It can also be seen from Figure 6.1 that the model has one 1D and one 2D subnetwork. In each of the subnetworks obtained, triangular membership functions of order 2 are used for all input variables, as shown in Figure 6.2. It can be seen from this figure that the membership functions of B, q and S_m are presented over a two-valued linguistic universe (i.e. small and large for B, light and heavy for q, and low and high for S_m). On the other hand, the membership functions of the soil density, which is represented herein by the average SPT blow count, N, is presented over a four-valued linguistic universe (i.e. loose, medium, dense and very dense). As a result, the first subnetwork contains 8 rules while the second subnetwork contains 2 rules, resulting in a model with 10 fuzzy rules, as listed in Table 6.3. It should be noted that the number between brackets in Table 6.3 represents the rule confidence described in Chapter 2.

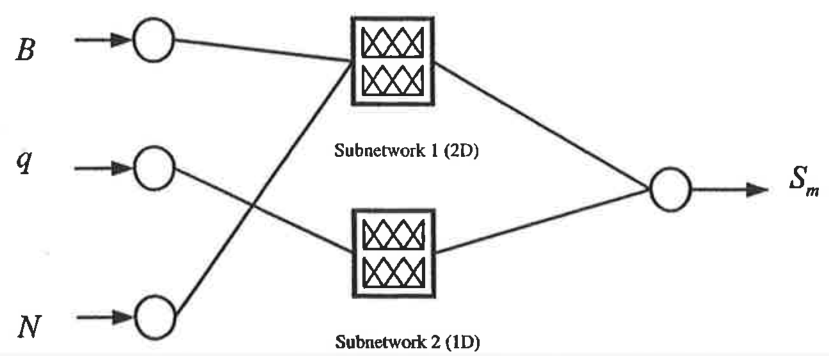


Figure 6.1: Schematic representation of the neurofuzzy model

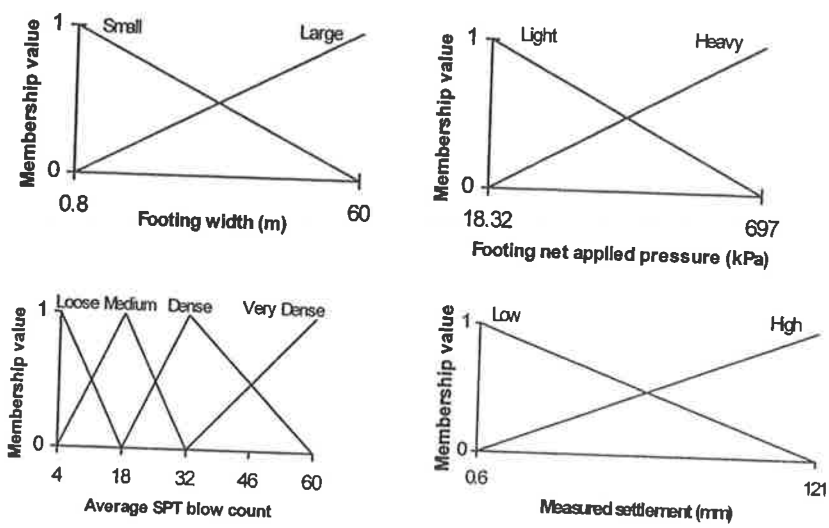


Figure 6.2: Membership functions of input variables used by the neurofuzzy model

Table 6.3: Fuzzy rules extracted by the neurofuzzy model

Subnetwork No.	Rule No.	Rule
1	1	IF "Footing width, B" is *Small* AND "Soil" is *Loose* THEN "Settlement, S_m" is *Low* (0.84) OR "Settlement, S_m" is *High* (0.16)
	2	IF "Footing width, B" is *Large* AND "Soil" is *Loose* THEN "Settlement, S_m" is *High* (1.00)
	3	IF "Footing width, B" is *Small* AND "Soil" is *Medium* density THEN "Settlement, S_m" is *Low* (0.99) OR "Settlement, S_m" is *High* (0.01)
	4	IF "Footing width, B" is *Large* AND "Soil" is *Medium* density THEN "Settlement, S_m" is *Low* (0.44) OR "Settlement, S_m" is *High* (0.56)
	5	IF "Footing width, B" is *Small* AND "Soil" is *Dense* THEN "Settlement, S_m" is *Low* (0.96) OR "Settlement, S_m" is *High* (0.04)
	6	IF "Footing width, B" is *Large* AND "Soil" is *Dense* THEN "Settlement, S_m" is *Low* (0.86) OR "Settlement, S_m" is *High* (0.14)
	7	IF "Footing width, B" is *Small* AND "Soil" is *Very dense* THEN "Settlement, S_m" is *Low* (1.00)
	8	IF "Footing width, B" is *Large* AND "Soil" is *Very dense* THEN "Settlement, S_m" is *Low* (0.87) OR "Settlement, S_m" is *High* (0.13)
2	9	IF "Footing net applied pressure, q" is *Light* THEN "Settlement, S_m" is *Low* (0.96) OR "Settlement, S_m" is *High* (0.04)
	10	IF "Footing net applied pressure, q" is *Heavy* THEN "Settlement, S_m" is *Low* (0.87) OR "Settlement, S_m" is *High* (0.13)

The fuzzy rules in Table 6.3 are considered to be a valuable source of information from which knowledge about the relationships between settlement and the factors affecting settlement can be extracted. The knowledge that can be derived from Table 6.3 is as follows:

- *Small* footings are most likely to result in *low* settlement regardless of the density of the soil they are founded on (Rules 1, 3, 5 and 7);
- *Large* footings are most likely to be susceptible to *high* settlement when they are founded on *loose* soils (Rule 2), and they are most likely to result in *low* settlement when they are founded on *dense* or *very dense* soils (Rules 6 and 8); and
- *Large* footings are equally likely to be susceptible to either *low* or *high* settlement when they are founded on soils of *medium* density (Rule 4).

It can also be seen from Table 6.3 that Rules 9 and 10 seem to contradict what one would expect based on the underlying physical meaning of the settlement problem. Rules 9 and 10 indicate that settlement is most likely to be *low* regardless of whether the applied load is *light* or *heavy*. The most likely reason for this is that the footings contained in the database used were designed so that the applied load does not result in high settlement or bearing capacity failure. Another reason is that there were insufficient training data to cover the full range of possible high settlement conditions. A review of the data used indicates that almost 90% of the data records have settlement that is described to be low settlement, as categorised by the settlement membership functions of the neurofuzzy model. It should be noted that the range of applicability of the fuzzy rules in Table 6.3 is a function of the quality of the data used in the model calibration phase. Consequently, it is unlikely that these fuzzy rules provide a general representation of the relationship between settlement and the factors affecting it. However, in general, the fuzzy rules obtained are in agreement with what one would expect based on the underlying physical meaning of the settlement problem. The above results indicate that neurofuzzy networks have the ability to extract rules from data that make physical sense, which may be used to gain understanding in situations where data are available but physical relationships are not well understood.

One advantage of neurofuzzy networks is that available engineering knowledge can be incorporated into the trained network to optimise model performance and to enhance the interpretation of a constructed model. In this work, this is done by optimising the membership functions of B and N so as to include available geotechnical engineering knowledge. The membership functions of B are optimised to be presented over a three-valued linguistic universe (i.e. small, medium and large) so that small footings are limited to footings of maximum width of 5 m (see Figure 6.3). On the other hand, the membership functions of N are optimised to be presented over a four-valued linguistic universe (i.e. loose, medium, dense and very dense) so that the classification of soil density given by Terzaghi and Peck (1948) can be incorporated, as shown in Figure 6.3 and given in Table 6.4. It should be noted that the most probable values of N that are incorporated in the membership functions for medium and dense soils are taken to be equal to the average of the range given by Terzaghi and Peck (1948). On the other hand, the most probable values of N that are incorporated in the membership functions for loose and very dense soils are taken to be equal to the minimum and maximum values of N, respectively, that are found in the database used. By doing this procedure, model CHP6-BIC is retrained and the performance results of the new model, referred to as "Optimised CHP6-BIC", is given in Table 6.5 together with the performance results of model CHP6-BIC, and the fuzzy rules of the Optimised CHP6-BIC model are listed in Table 6.6. It should be noted that this model again has three inputs (i.e. B, q and N) and the number of connection weights is equal to 9.

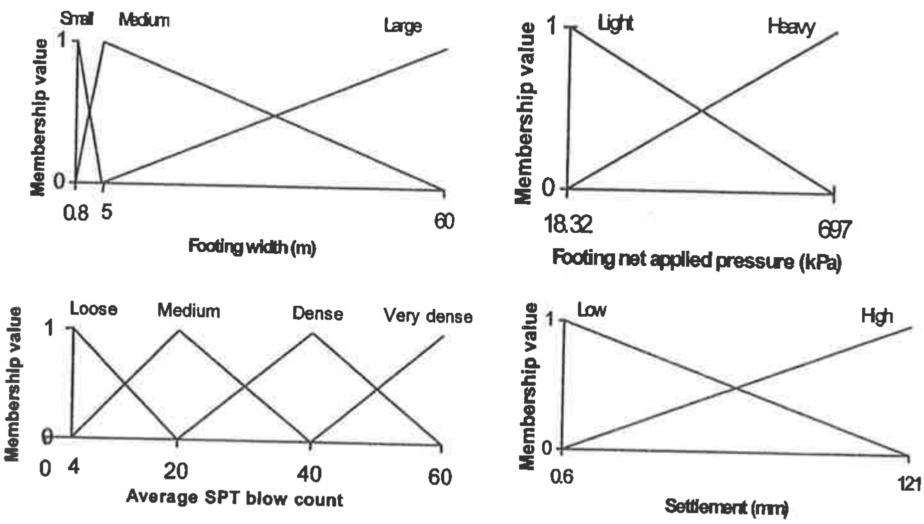


Figure 6.3: Optimised membership functions of the neurofuzzy model

Table 6.4: Optimisation of membership functions of N to incorporate the classification of soil density of Terzaghi and Peck (1948)

Terzaghi and Peck (1948)		Most probable value of N incorporated in the membership functions of the neurofuzzy model
Soil density	Possible values of N	
Loose	<10	4
Medium	10–30	20
Dense	30–50	40
Very dense	>50	60

Table 6.5: Performance of models CHP6-BIC and Optimised CHP6-BIC

Model No.	Performance measure					
	Correlation coeff., r		RMSE (mm)		MAE (mm)	
	Training	Validation	Training	Validation	Training	Validation
Optimised CHP6-BIC	0.893	0.892	12.11	11.74	7.87	8.93
CHP6-BIC	0.889	0.881	12.33	12.36	8.08	9.36

Table 6.6: Fuzzy rules extracted by the Optimised CHP6-BIC model

Subnetwork No.	Rule No.	Rule
1	1	IF "Footing width, B" is *Small* AND "Soil" is *Loose* THEN "Settlement, S_m" is *Low* (0.86) OR "Settlement, S_m" is *High* (0.14)
	2	IF "Footing width, B" is *Medium* AND "Soil" is *Loose* THEN "Settlement, S_m" is *Low* (0.64) OR "Settlement, S_m" is *High* (0.36)
	3	IF "Footing width, B" is *Large* AND "Soil" is *Loose* THEN "Settlement, S_m" is *High* (1.00)
	4	IF "Footing width, B" is *Small* AND "Soil" is *Medium* density THEN "Settlement, S_m" is *Low* (0.91) OR "Settlement, S_m" is *High* (0.09)
	5	IF "Footing width, B" is *Medium* AND "Soil" is *Medium* density THEN "Settlement, S_m" is *Low* (0.90) OR "Settlement, S_m" is *High* (0.10)
	6	IF "Footing width, B" is *Large* AND "Soil" is *Medium* density THEN "Settlement, S_m" is *Low* (0.46) OR "Settlement, S_m" is *High* (0.54)
	7	IF "Footing width, B" is *Small* AND "Soil" is *Dense* THEN "Settlement, S_m" is *Low* (0.94) OR "Settlement, S_m" is *High* (0.06)
	8	IF "Footing width, B" is *Medium* AND "Soil" is *Dense* THEN "Settlement, S_m" is *Low* (0.89) OR "Settlement, S_m" is *High* (0.11)
	9	IF "Footing width, B" is *Large* AND "Soil" is *Dense* THEN "Settlement, S_m" is *Low* (0.84) OR "Settlement, S_m" is *High* (0.16)
	10	IF "Footing width, B" is *Small* AND "Soil" is *Very Dense* THEN "Settlement, S_m" is *Low* (1.00)
	11	IF "Footing width, B" is *Medium* AND "Soil" is *Very Dense* THEN "Settlement, S_m" is *Low* (0.90) OR "Settlement, S_m" is *High* (0.10)
	12	IF "Footing width, B" is *Large* AND "Soil" is *Very Dense* THEN "Settlement, S_m" is *Low* (0.81) OR "Settlement, S_m" is *High* (0.19)
2	13	IF "Footing net applied pressure, q" is *Light* THEN "Settlement, S_m" is *Low* (0.89) OR "Settlement, S_m" is *High* (0.11)
	14	IF "Footing net applied pressure, q" is *Heavy* THEN "Settlement, S_m" is *Low* (0.78) OR "Settlement, S_m" is *High* (0.22)

It can be seen from Table 6.5 that the new model (i.e. Optimised CHP6-BIC) performs well and its performance is slightly better than the previously developed model (i.e. model CHP6-BIC) with respect to the training and validation sets. It can also be seen from Table 6.6 that the model has 14 fuzzy rules. The knowledge that can be extracted from this model is as follows (Table 6.6):

- *Small* footings are most likely to result in *low* settlement regardless of the density of the soil they are founded on (Rules 1, 4, 7 and 10);
- *Medium* size footings are approximately equally likely to be susceptible to either *low* or *high* settlement when they are founded on *loose* soils (Rule 2), and they are most likely to result in *low* settlement when they are founded on *medium*, *dense* or *very dense* soils (Rules 5, 8 and 11);
- *Large* footings are most likely to be susceptible to *high* settlement when they are founded on *loose* soils (Rule 3), and they are most likely to result in *low* settlement when they are founded on *dense* or *very dense* soils (Rules 9 and 12); and
- *Large* footings are equally likely to be susceptible to either *low* or *high* settlement when they are founded on soils of *medium* density (Rule 6).

Also, Rules 13 and 14 again indicate that settlement is most likely to be *low* regardless of whether the applied load is *light* or *heavy*, which contradicts what one would expect. As mentioned previously, the most likely reason for this is that the footings contained in the data set were designed so that the applied load does not result in high settlement and that there were insufficient training data to cover the full range of possible high settlement conditions. In general, the knowledge obtained in Table 6.6 is in agreement with what one would expect, based on the underlying physical meaning of the settlement problem, and are in agreement with the knowledge obtained previously from model CHP6-BIC (Table 6.3). However, the fuzzy rules in Table 6.6 describe the relationship between settlement and the factors affecting settlement in more rational fashion. The above results suggest that it is beneficial to add available expertise to neurofuzzy models, as it can improve model performance and enhance the interpretation of the constructed models.

In order to test the robustness of the neurofuzzy model (i.e. Optimised CHP6-BIC), a parametric study on the input variables is carried out, as suggested in §5.2.1, and the results are presented in Figure 6.4. It can be seen from Figures 6.4 (a) and (b) that the settlement increases as the footing width and footing net applied pressure increase. On the other hand, Figure 6.4 (c) shows that the settlement decreases as the average SPT blow count increases. These results indicate that the behaviour of the neurofuzzy model is similar to what one would expect based on the underlying physical sense of settlement prediction. Consequently, this model can be considered to be robust and hence, can be used for settlement prediction.

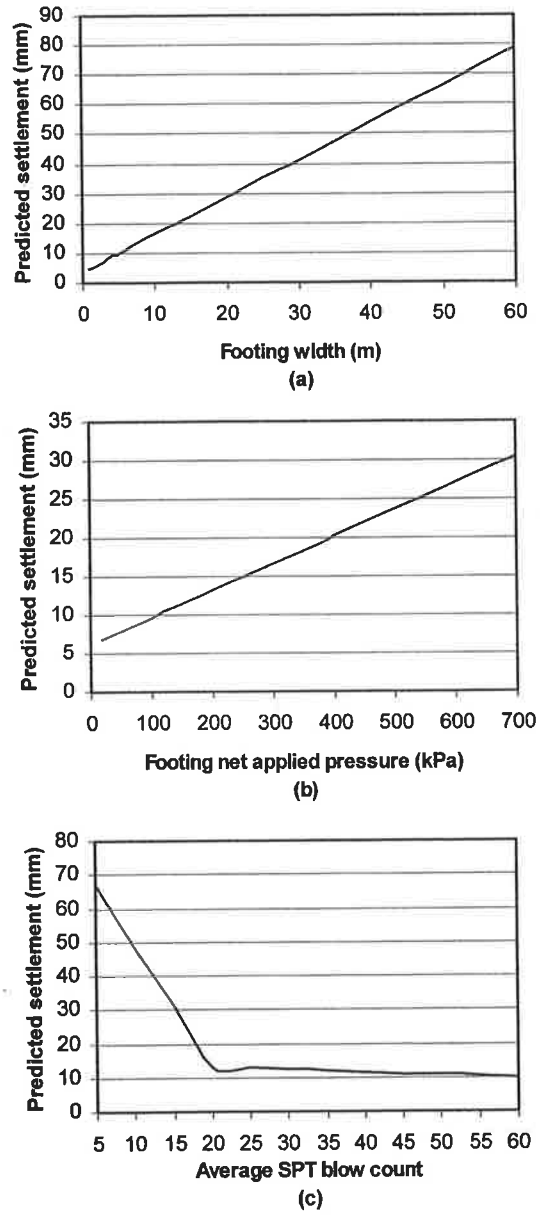


Figure 6.4: Robustness tests for the Optimised CHP6-BIC model

6.5 Comparison of the Neurofuzzy and MLP Models

A comparison between the optimum neurofuzzy model (i.e. Optimised CHP6-BIC) and the best back-propagation MLP model (i.e. model CHP5-NF2) is carried out in terms of model accuracy, model parsimony and model transparency. A summary of the number of inputs and connection weights used by each model is given in Table 6.7. The performance results of the two models with respect to the validation set are also given in Table 6.7. In terms of model accuracy, it can be seen that the two models are comparable, although the MLP model performs slightly better than the neurofuzzy model. This suggests that the two models provide similar settlement prediction. However, although the robustness tests carried out on the two models (Figure 6.5) show that both models are robust, there is a marked difference in the predicted settlements between the two models for a range of data similar to those used for model training. As can be seen from Figure 6.5, the trends of the predicted settlement using the neurofuzzy model are linear, whereas they are non-linear for the MLP model. One possible reason for this behaviour is that optimisation of the neurofuzzy model is based on linear membership functions of order 2, whereas it is based on non-linear transfer functions (sigmoidal or tanh transfer functions) for the MLP model.

In order to investigate the aforementioned reason, an attempt to obtain a non-linear behaviour from the neurofuzzy model is carried out by retraining the Optimised CHP6-BIC model with membership functions of order 3, and then comparing its robustness behaviour with this of the MLP model, as shown in Figure 6.6. It can be seen that the trends of the neurofuzzy model are changed to non-linear, which confirms the reason proposed above. However, the unexpected deviations in the trends of the settlement predicted by the retrained neurofuzzy model suggest that the model in not robust. One would expect the settlement to increase with increase in footing width and footing net applied pressure and that it decrease with the increasing average SPT blow count. It seems that the robustness behaviour of the neurofuzzy model shown in Figure 6.6 is a result of data overfitting. As mentioned in Chapter 2, increasing the order of the membership functions of the B-spline neurofuzzy models results in smoother model outputs, but can lead to overfitting of the data (Brown and Harris 1994).

Table 6.7: Comparison between the neurofuzzy and MLP models

Model type	No. of inputs	No. of connection weights	Model performance on the validation set		
			Correlation coefficient, r	RMSE (mm)	MAE (mm)
Neurofuzzy	3	9	0.892	11.74	8.93
MLP	5	12	0.905	11.04	8.78

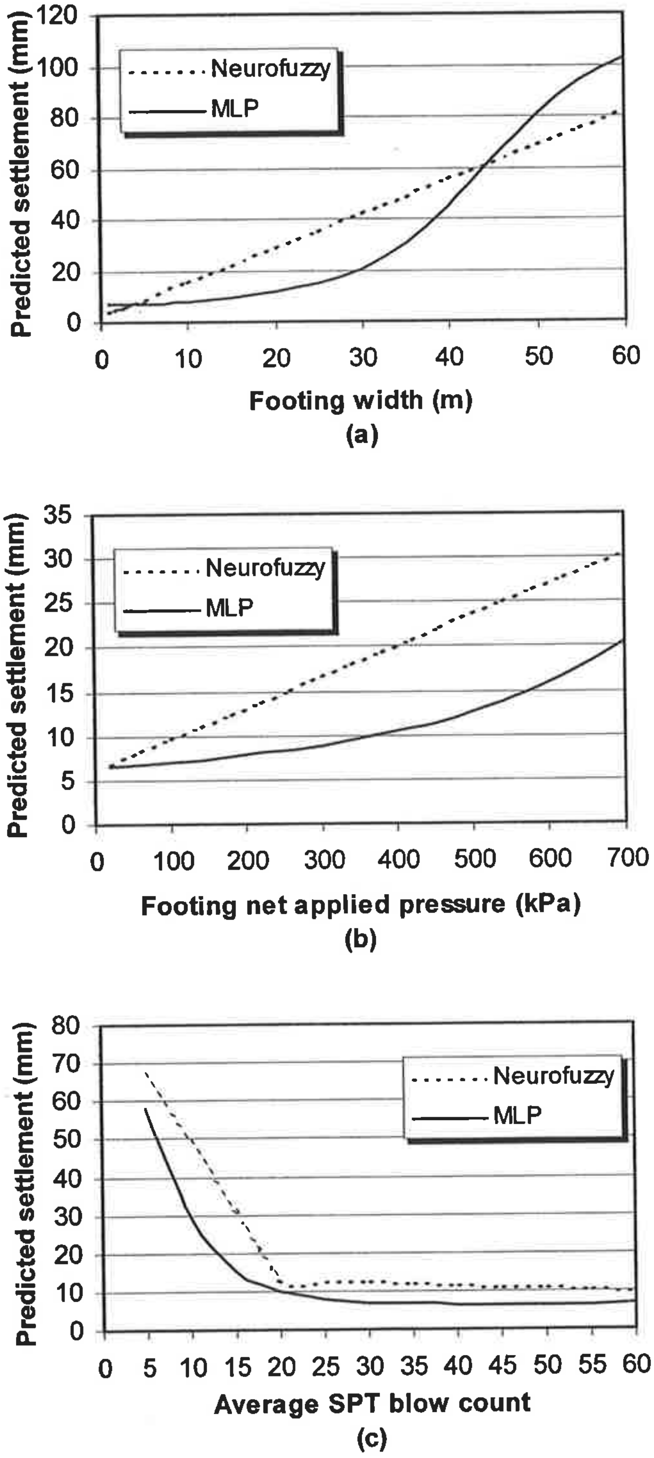


Figure 6.5: Robustness tests for the neurofuzzy and MLP models

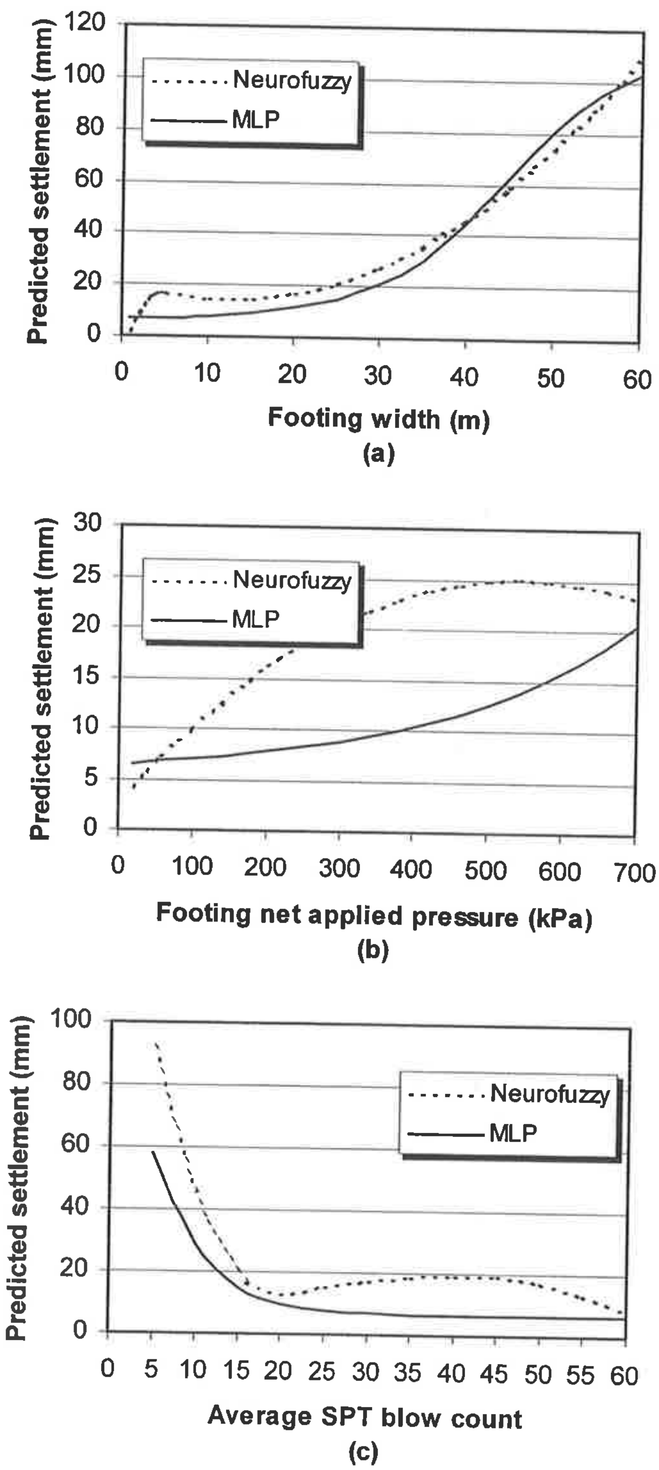


Figure 6.6: Robustness tests for the neurofuzzy model of membership functions of order 3 and MLP model

In order to further investigate the reason stated above, an attempt to obtain linear behaviour from the MLP model is carried out by retraining model CHP5-NF2 using linear transfer functions in the MLP hidden and output layers, and then comparing its robustness behaviour with that of the Optimised CHP6-BIC model, as shown in Figure 6.7. It can be seen that the trends of the MLP model have changed to linear and that the prediction of the MLP and neurofuzzy models becomes closer to each other, as shown in Figures 6.7 (a) and (b). This confirms the reason proposed previously regarding the difference in robustness behaviour between the neurofuzzy and MLP models. However, as shown in Figure 6.7 (c), there still exists some unexpected deviation (negative values) in the trend that relates the predicted settlement with the average SPT blow count for the retrained MLP model, which suggests that the model is not robust. The above results indicate that it is the functions used (i.e. membership basis functions for the neurofuzzy model and the transfer functions for the MLP model) that result in the difference in robustness behaviour between the neurofuzzy and MLP models shown in Figure 6.5.

It is still necessary to decide which model to use as an optimum ANN model. The non-linear robustness behaviour of the predicted settlement by the MLP model seems to be more realistic, as it is unlikely that settlement is linear, which is confirmed by traditional methods and underlying geotechnical engineering knowledge. In addition, the more accurate performance of the MLP model, with respect to the validation set, indicates that the MLP model can provide more accurate settlement predictions in real-world situations. Consequently, the MLP model (i.e. model CHP5-NF2) developed in Chapter 5 will be considered to be the best ANN model and hence will be used for the analyses in the next chapter (Chapter 7).

In terms of model parsimony, the neurofuzzy model is found to be more parsimonious than the back-propagation MLP model, as it has a smaller number of model inputs and connection weights. In terms of model transparency, the neurofuzzy model is found to provide a more explicit interpretation of the relationships between model inputs and the corresponding output in the form of a set of linguistic fuzzy rules that describe the model in a more transparent fashion (Table 6.6). However, as shown in Chapter 5, the small number of hidden layer nodes of the back-propagation MLP model enabled the translation of the model into a relatively simple equation that provides a valuable insight into the relationships between the model inputs and the corresponding outputs. For large MLP models with a greater number of inputs and hidden layer nodes, a derivation of such an equation could be difficult and consequently, the use of neurofuzzy models would be better in such situations.

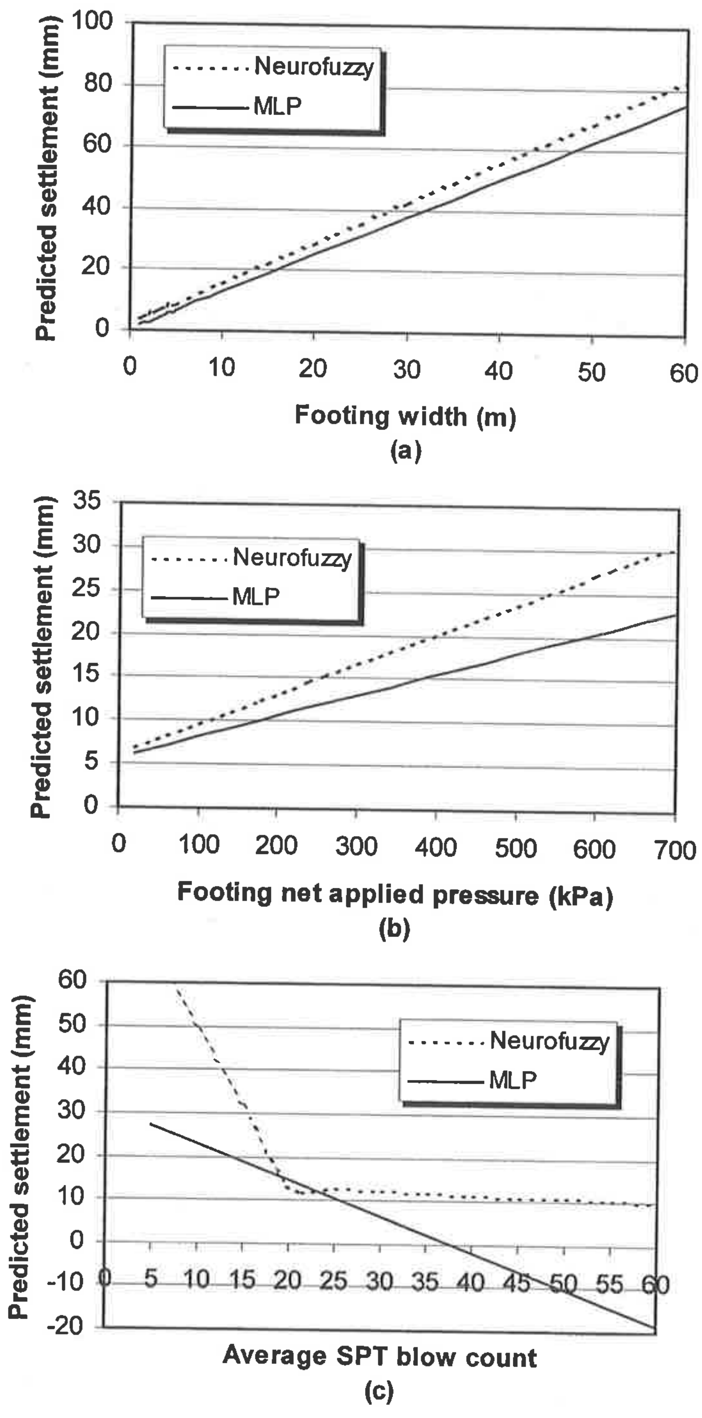


Figure 6.7: Robustness tests for the optimised neurofuzzy model and MLP model of linear transfer functions

6.6 Summary and Conclusions

B-spline networks trained with the ASMOD algorithm were used to demonstrate the feasibility of neurofuzzy models to predict the settlement of shallow foundations on cohesionless soils and to assist with providing a better understanding of the relationship between settlement and the factors affecting it. The ASMOD is an algorithm that automatically optimises model architecture and selects input variables that have the most significant impact on settlement. Five potential input variables (i.e. B, q, N, L/B and D_f/B) were presented to the neurofuzzy models and settlement was the single output. The sensitivity of the neurofuzzy models to a number of stopping criteria, i.e. Bayesian Information Criterion (BIC), Akaike's Information Criterion (AIC) and Final Prediction Error (FPE), was investigated. The models were assessed in terms of prediction accuracy, model parsimony and model transparency. The optimum neurofuzzy model obtained was compared with the best back-propagation MLP model obtained in Chapter 5.

This chapter has yielded the following results and conclusions:

1. Neurofuzzy models have the ability to accurately predict the settlement of shallow foundations on cohesionless soils and are able to extract rules from the data that make physical sense, which may be used to gain a better understanding in situations where data are available but physical relationships are not well understood. In addition, neurofuzzy networks can be modified by incorporating available engineering knowledge to improve model performance and enhance the interpretation of the constructed model.

2. The footing width (B), footing net applied pressure (q) and average SPT blow count (N) were found to be the most significant factors affecting settlement. This is in agreement with the results of the sensitivity analysis carried out on the MLP model in Chapter 5.

3. All neurofuzzy models were found to be comparable in terms of prediction accuracy, even though the model that uses the BIC was found to perform marginally better than the other models on an independent validation set.

4. All neurofuzzy models were found to be comparable in terms of model parsimony, even though the model that uses the AIC was found to be more parsimonious than the other models with the lowest number of connection weights. This was attributed to the fact that AIC penalises complex models.

5. The neurofuzzy models that use the BIC and AIC were found to be more transparent than the model that uses the FPE, as they have fewer fuzzy rules. This was attributed to the fact that the BIC and AIC penalise complex models more.

6. The optimum neurofuzzy and MLP models were found to be comparable in terms of model accuracy, although the MLP model was found to perform slightly better than the neurofuzzy model.

7. The optimum neurofuzzy model was found to be more parsimonious than the back-propagation MLP with fewer model inputs and connection weights.

8. The optimum neurofuzzy model was found to be more transparent than the back-propagation MLP model as it was able to describe the relationship between the model inputs and corresponding output using a set of fuzzy rules. On the other, as shown in Chapter 5, the MLP model was able to be translated into a relatively simple equation that gives valuable insight into the relationships between the model inputs and corresponding outputs.

In the following chapter, the use of stochastic simulation in the analysis of ANN settlement prediction will be examined.

Chapter 7

Stochastic Analysis of Settlement Prediction

7.1 Introduction

Settlement prediction, as with many geotechnical engineering problems, is often affected by a considerable level of uncertainty. Such uncertainty may produce an unreliable estimation of the magnitude of settlement, while reliable settlement prediction is essential for design purposes. Uncertainty affecting settlement prediction is generally caused by one or more of the following (Krizek et al. 1977; Cherubini and Greco 1991):

1. Parameter uncertainty; and
2. Uncertainty associated with the model used for settlement prediction.

There are a number of major and minor factors that contribute to parameter uncertainty. The major factors include (i) poor knowledge of soil properties and (ii) uncertainty in forecasting the magnitude of the imposed loads. Uncertainty associated with poor knowledge of the soil properties is due to the natural spatial variability of soil, which is caused by variations in the mineral composition and characteristics of soil strata during and after soil formation. It is also due to insufficient description of soil characteristics as a result of limited spatial sampling. Uncertainty associated with this source can also be due to errors associated with the technique used to measure the actual soil properties.

Theoretically, the uncertainty associated with the loads acting on foundations can be determined with an acceptable degree of accuracy (Greco and Cherubini 1993). However, realistically, it is unlikely that an accurate estimation of the magnitude of design loads can be made and thus loads should be treated as random variables (Corotis 1972; Peir and Cornell 1973; Melchers 1987).

The minor factors that contribute to parameter uncertainty include (i) footing dimensions and (ii) footing embedment depth. These sources of parameter uncertainty are due to discrepancies between footing dimensions or the footing embedment depth implemented on site and those that appear in construction drawings, as a result of human error.

The second source of uncertainty is caused by the inherent error associated with the modelling technique used to characterise settlement prediction and is usually called *model uncertainty* (Frey 1998). This type of uncertainty is due to the simplified nature of models that are used to describe soil behaviour, which are generally based on a number of assumptions. Unfortunately, model uncertainty is difficult to measure physically and in most instances, the model used to describe a certain phenomenon is assumed to be a perfect predictor (Fenton 2002). However, if sufficient measured and predicted data are available, then the overall uncertainty associated with the prediction method used can be quantified. Consequently, the overall uncertainty associated with a certain prediction method is the sum of the parameter and model uncertainties. This type of uncertainty can be referred to as the prediction method uncertainty.

Most deterministic modelling methods for settlement prediction of shallow foundations on cohesionless soils disregard the above uncertainties in their analysis and simulation. Despite the relative advantage, that has been shown in Chapter 5, for the ANN approach over traditional methods, it does not take into account the considerable level of uncertainty that may affect the magnitude of the predicted settlement. ANNs, like more traditional methods of settlement prediction, are based on deterministic approaches that ignore the above uncertainty and thus provide single values of settlement with no indication of the level of risk associated with these values. An alternative stochastic approach is essential to provide more rational estimation of settlement. Stochastic simulation has a significant benefit over deterministic methods in the sense that the degree of risk (i.e. uncertainty) associated with the model output can be quantified (Jaksa 1995; Barthur 1997; Whitman 2000). In this chapter, stochastic analysis is applied to the ANN model in order to obtain a stochastic model of ANN settlement prediction of shallow foundations on granular soils. This chapter has three main objectives. The first objective is to present and compare practical stochastic approaches that incorporate parameter uncertainty and prediction method uncertainty in the analysis of ANN settlement prediction of shallow foundations on cohesionless soils. The second objective is to examine the effect of varying the parameter uncertainty on the magnitude of the predicted settlement. Finally, and most significantly, the third objective is to develop and provide a set of stochastic design charts that are based on the ANN method for routine use in practice. The charts are useful in the sense that they enable the designer to make informed decisions regarding the level of risk associated with predicted settlements and consequently provide a more realistic indication of what the actual settlement might be. In order to demonstrate the first two objectives set out above, a numerical example is provided.

7.2 Overview of Stochastic Settlement Prediction

Over the last decade, interest in applying more rational stochastic analyses in the field of geotechnical engineering, rather than the less accurate traditional deterministic solutions, has increased rapidly (see Tang 1993). For example, in the area of settlement prediction of shallow foundations, Padilla and Vanmarcke (1974) developed a stochastic approach for settlement prediction of a one-dimensional model based on a first-order probabilistic description of loads and soil properties. Fraser and Wardle (1975) used a first order probabilistic analysis to develop a model for the determination of total and differential settlement of raft foundations resting on layered cross-anisotropic elastic soils, taking into account the uncertainty associated with the imposed loads and supporting soil modulus. Cherubini and Greco (1991) presented a probabilistic approach for settlements predicted using the method proposed by Arnold (1980) to estimate the settlement of spread footings on sand, taking into account the uncertainty associated with the reliability of the technique used for settlement prediction. Brzakala and Pula (1996) also combined finite element analysis with stochastic simulation to provide a probabilistic solution for the estimation of settlement of shallow foundations, considering three basic sources of input parameter uncertainty: random shape of the subsoil (location of an interface between two strata), random material parameters and random loads. Fenton et al. (1996) estimated probabilistic measures of total and differential settlement of spread footings on elastic soils using a two-dimensional finite element model combined with Monte Carlo simulation, taking into account the variability of the soil modulus of elasticity. Sivakugan and Johnson (2002) applied probabilistic analysis to settlement prediction of four deterministic traditional methods including the methods proposed by Terzaghi and Peck (1967), Schmertmann et al. (1978), Burland and Burbidge (1985) and Berardi and Lancellotta (1994), considering the prediction method uncertainty.

7.3 Basic Statistical Definitions

When data are available, it is useful to quantify their statistical properties. The following definitions are some of the most commonly used statistical parameters for data presentation and are defined in many fundamental publications (e.g. Seber 1974; Smith 1986; Raghavarao 1988), and are described here as they form part of the analyses that follow.

- **Mean (μ):** the mean is the average of the sample data (x_1, x_2, …, x_n) of size n and is defined as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (7.1)$$

- **Standard deviation (σ):** the standard deviation is a measure of deviation or spread of the sample data about their mean and is defined as:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (7.2)$$

- **Coefficient of variation (COV):** the coefficient of variation is the expression of the degree of spread of the data in terms of the mean. It is useful for comparing groups with different means and is defined as:

$$COV = \frac{\sigma}{\bar{x}} \qquad (7.3)$$

- **Graphical presentation of data:** the histogram is the most common graphical method for data presentation (Lee et al. 1983). It is a grouping of data into categories of given numerical intervals showing the proportional frequency of observations in each category. The hatched area in Figure 7.1 is an example of a histogram representing the distribution of soil compaction data. The scaling ordinates of the histogram are called the frequency distribution and can be obtained by dividing the number of observed values within a specific interval by the total number of observations.

Another method of data presentation is the probability distribution (usually referred to as 'PDF' = probability density function), of which the solid line in Figure 7.1 provides an example. The PDF has the crucial property that the area under the curve between any two values gives the probability of obtaining an observation between those values. Another graphical method of data presentation is the cumulative frequency distribution function, which calculates successive sums of frequencies up to each interval point and connects these points. The dashed line and right-hand ordinate in Figure 7.2 provide an example of a cumulative frequency distribution curve of the soil compaction data shown in Figure 7.1.

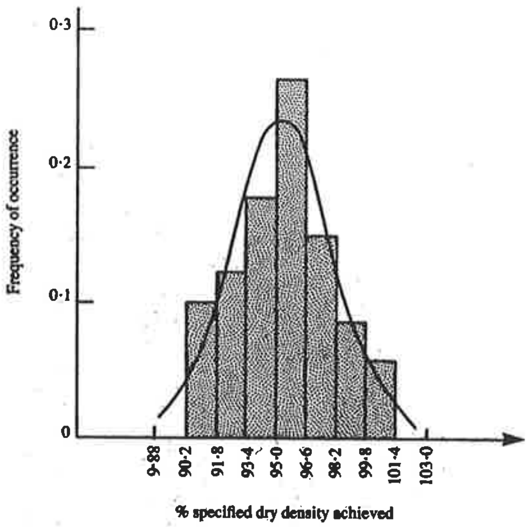


Figure 7.1: Histogram and probability distribution for soil compaction data (Lee et al. 1983)

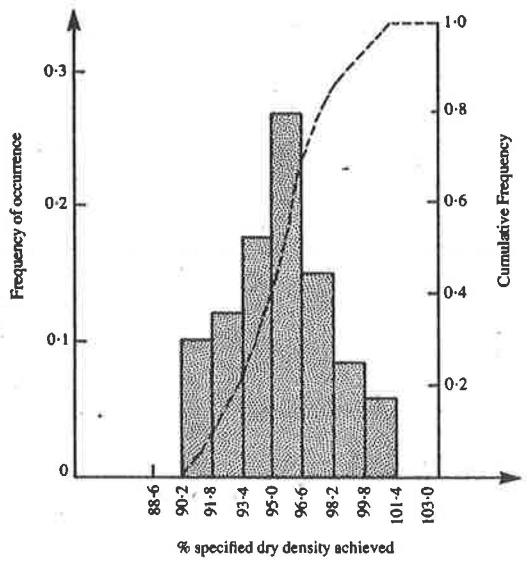


Figure 7.2: Cumulative probability distribution of soil compaction data (Lee et al. 1983)

7.4 Stochastic Analysis of Settlement Prediction

Stochastic analysis is a procedure of handling mathematical problems where some of the parameters are uncertain and thus can be presented as random variables (Prekopa 1995). There are various stochastic approaches available for geotechnical engineering applications. For models involving random input variables with known or assumed probability distributions, Monte Carlo simulation can be used to estimate the probability distribution of the model output. Monte Carlo simulation is the technique that will be used to produce the stochastic solutions for settlement prediction given later and consequently, it will be described briefly in this chapter.

Monte Carlo simulation attempts to generate a random set of values from known or assumed probability distributions of some input variables involved in a certain problem to determine the probability distribution of the output variable. The steps in the Monte Carlo simulation, for a set of input variables (x_1, x_2, x_3, ..., x_n) and the corresponding output variable y, where, $y = f(x_1, x_2, x_3, \ldots, x_n)$, are as follows:

1. Values of each input variable (x_1, x_2, x_3, ..., x_n) are generated randomly by knowing or assuming their probability density function and statistical parameters (e.g. mean, standard deviation);

2. The output variable y is obtained from its deterministic function, $y = f(x_1, x_2, x_3, \ldots, x_n)$;

3. The above steps are repeated, usually thousands of times; and

4. Finally, the output values are used to obtain the mean, standard deviation and probability density function from which the probability of occurrence or risk associated with a certain prediction can be estimated.

Further details about the Monte Carlo technique are given by many authors (e.g. Hammersley and Handscomb 1964; Rubinstein 1981; Ang and Tang 1984).

As mentioned earlier, this chapter applies stochastic simulation to ANNs in order to incorporate parameter uncertainty (i.e. uncertainty associated with the model inputs as soil properties and imposed loads) and prediction method uncertainty (i.e. parameter uncertainty + uncertainty associated with the ANN modelling technique). The approach used to obtain the uncertainties in settlement prediction associated with each of the two types of uncertainty considered is given below.

7.4.1 Inclusion of parameter uncertainty

For an individual case of settlement prediction, the procedure for obtaining the stochastic solution that incorporates the parameter uncertainty is as follows:

1. The values of each input variable (i.e. B, q, N, L/B, D_f/B) are generated randomly by knowing or assuming their probability density function (PDF) and any correlation that exists between them;

2. The deterministic predicted settlement is obtained from the best ANN model obtained in this research (i.e. model CHP5-NF2);

3. The above two steps are repeated, thousands of times as part of Monte Carlo simulation; and

4. Finally, the subsequent settlements are used to obtain the cumulative distribution function (CDF) or to plot the cumulative probability distribution from which the probability of non-exceedance ($P_{N/E}$), or level of risk, associated with a certain prediction can be estimated.

Among the five inputs to the ANN model, two, the footing net applied pressure, q, and average SPT blow count, N, are likely to include more than marginal parameter uncertainty, and thus, in this work, are assumed to be random variables. As mentioned earlier, footing dimensions contribute to parameter uncertainty to a lesser degree and are thus assumed to be deterministic for practical purposes. In addition, the input variable of footing embedment depth, D_f, is also assumed to be deterministic. A number of studies have attempted to characterise the uncertainty associated with q and N, as discussed below.

According to Melchers (1987), loads acting on structures can be divided into two broad groups: natural loads (e.g. wind and earthquake) and human-imposed loads (e.g. dead loads and live loads) and the magnitude of each varies with time and location. Consequently, estimation of total loads imposes uncertainty. As a guide, Auvinet and Rossa (1991) showed that the coefficient of variation, COV (i.e. standard deviation/mean), of permanent loads for Mexico City buildings is 8%. Melchers (1987), stated that dead loads are commonly assumed to be closely approximated by a normal distribution with a COV of 5 to 10%. Rao (1992) stated that dead loads are usually described by a normal distribution with COV of 10%. Krizek et al. (1977) demonstrated that the uncertainty associated with the estimation of the total imposed loads follows a normal distribution and Fraser and Wardle (1975) illustrated that the COV of the total imposed loads that can be encountered in practice is equal to 14%. Padilla and Vanmarcke (1974) also showed that, if dead and live loads were assumed to be stochastically independent, the resulting variability of their sum would have a COV of 10%.

Uncertainty associated with the SPT blow count, *N*, as a measure of soil compressibility is considerable due to the many factors that affect SPT results (Orchant et al. 1987). Fletcher (1965) identified thirteen factors that affect the SPT and can be categorised into the following two major groups: (i) equipment effects (e.g. hammer, hammer drop system, drill rods, and sampler) and (ii) procedural/operator effects (e.g. height of hammer drop, seating of the sampler, errors in counting blows, and cleaning of borehole). Lee et al. (1983) reported that the uncertainty associated with the average SPT blow count can be assumed to follow a normal distribution with COV ranges from 27 to 85% and recommended a value equal to 30%.

7.4.2 Inclusion of prediction method uncertainty

The stochastic solution that incorporates the prediction method uncertainty is based on an assumption that previous measured settlements of foundations may be employed to predict the settlements of other foundations in similar conditions (Cherubini and Greco 1991). The uncertainty of the prediction method can be examined by calculating the settlement ratio, *k* (Cherubini and Greco 1991; Sivakugan and Johnson 2002), which is defined as the ratio of the predicted settlement to the actual measured settlement. If a set of predicted and measured settlements is available, the settlement ratios can be calculated and used to obtain the PDF of *k*. A Monte Carlo simulation can then be conducted to estimate the uncertainty associated with the predicted settlements. The detailed procedure is as follows:

1. The PDF of *k* is estimated using a set of predicted and measured settlements;
2. Random values of *k* are generated from this PDF;
3. For each generated value of *k*, the deterministic settlement is calculated using the ANN model;

4. From the definition of k, the deterministic predicted settlement is divided by the generated random value of k in order to obtain the corresponding actual settlement;

5. Steps 2 to 4 are repeated for many iterations (Monte Carlo simulation); and

6. The settlements obtained as part of the Monte Carlo simulation are used to estimate the CDF or to plot the cumulative probability distribution from which the uncertainty or level of risk associated with a certain settlement prediction can be estimated.

A number of studies have attempted to characterise the prediction method uncertainty. For example, Greco and Cherubini (1993) demonstrated that the distribution of k could be approximated by a lognormal distribution for settlement predictions obtained from the methods proposed by Arnold (1980) and Papadopoulos (1992). Sivakugan and Johnson (2002) showed that k can be represented by beta distribution for settlements predicted using four traditional methods (i.e. Terzaghi and Peck 1967; Schmertmann et al. 1978; Burland and Burbidge 1985; Berardi and Lancellotta 1994). In this work, the distribution of k is obtained using the 189 data records used in Chapter 5 for the development of the ANN model, as shown later.

7.5 Numerical Example

In order to demonstrate the approach outlined in the previous section, the following numerical example is examined, which is identical to the one presented in §5.7. A rectangular footing whose dimensions are 2.5 × 4.0 m is founded at a depth equal to 1.5 m below the ground surface. The soil beneath the footing is sand that extends to a depth in excess of twice its width. The net applied load exerted on the footing is 350 kPa and the average SPT blow count is 16.

7.5.1 Estimation of parameter uncertainty

The deterministic solution of settlement prediction is first obtained from the ANN model in §5.2.2 (model CHP5-NF2) and is found to be 13.3 mm, as given previously in §5.7. The statistical data representing the uncertainty associated with settlement parameters are taken to be equal to those commonly encountered in practice and recommended in the literature, as described earlier, and are shown in Table 7.1. In addition, the database used in Chapter 5 for the development of the ANN model is utilised to determine the coefficient of correlation between the footing net applied pressure, q, and average SPT blow count, N, which was found to equal 0.4.

Table 7.1. Statistics for parameter uncertainty used in the numerical example

Settlement parameter	Mean	St. deviation	COV (%)	PDF
Footing width, B (m)	2.5	Deterministic	deterministic	N/A
Net applied footing load, q (kPa)	350	35	10	Normal
Average SPT blow count, N	16	4.8	30	Normal
Footing geometry, L/B	1.6	Deterministic	deterministic	N/A
Footing embedment ratio, D_f/B	0.6	Deterministic	deterministic	N/A

N/A = not applicable

In order to apply the stochastic approach that incorporates parameter uncertainty, the statistical data in Table 7.1 are used to generate sample values for q and N. These values are generated so as to be within (i) the range of data that can be expected in practical applications and (ii) the ranges of the input data used for training of the ANN model. Again, the PC-based software @Risk (Palisade 2000) is used for this purpose. The stochastic procedure outlined previously for incorporating parameter uncertainty is applied until a convergence criterion is achieved. In order to determine whether convergence has been achieved, the statistics describing the distribution of the predicted settlements are calculated at fixed numbers of simulations and compared with the same statistics at previous simulations. Convergence is deemed to have occurred if the change in the statistics describing the distribution of predicted settlement is 1% or less. It was found that 1,300 simulations are sufficient to achieve convergence. The predicted settlements obtained for the 1,300 simulations are used to plot the cumulative probability distribution curve from which different probabilities of non-exceedance are obtained. The results are shown in Figure 7.3 and are summarised in Table 7.2 (columns 1 and 3). The results shown in columns 2 and 4 of Table 7.2 will be explained later in this section.

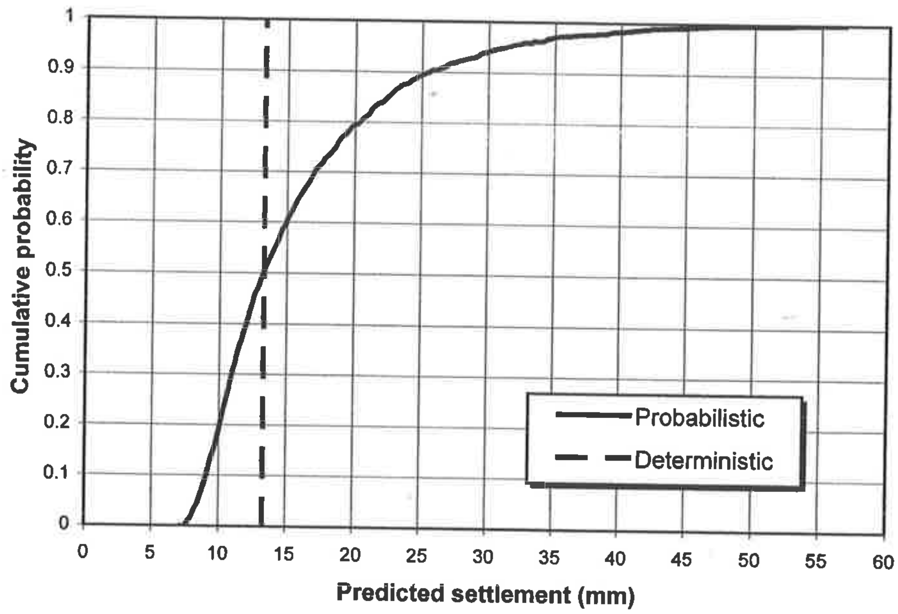


Figure 7.3. Cumulative probability distribution incorporating parameter uncertainty for the numerical example

Table 7.2: Predicted settlements accounting for parameter uncertainty of different q and N for the numerical example

Probability of non-exceedance (%)	COV for q and N (%)		
	$q = 5$ and $N = 27$	$q = 10$ and $N = 30$	$q = 14$ and $N = 85$
75	18.0	18.6	19.5
80	19.6	20.4	23.6
85	21.3	22.4	28.1
90	24.2	25.4	34.2
95	29.2	31.3	42.6

It can be seen from Figure 7.3 that there is a probability of approximately 50% that the settlement could be higher than the deterministic estimation of 13.3 mm. This result indicates that the uncertainty associated with q and N can considerably affect settlement and thus, should not be neglected in the analysis and simulation of settlement prediction. In addition, there are probabilities of 75%, 80%, 85%, 90% and 95% (i.e. probability levels that may be needed for design purposes) that the settlement will not exceed 18.6, 20.4, 22.4, 25.4 and 31.3 mm, respectively (Table 7.2).

Once the stochastic simulation has been performed, a sensitivity analysis is carried out to determine the relative impact of the input variables on the uncertainty associated with predicted settlement. This is done by calculating the correlation coefficient between the predicted settlements and the values generated for each input variable. The higher the correlation between the input and settlement, the more significant the input is in determining the stochastic predicted settlement. As expected, the results of the sensitivity analysis show that the uncertainty associated with the average SPT blow count has a considerable impact on the uncertainty associated with predicted settlement, as it has a high correlation coefficient of –0.935. The negative sign of the correlation coefficient indicates that, as expected, there is an inverse relationship between the average SPT blow count and settlement prediction. On the other hand, the results show that the uncertainty associated with the net applied footing load has a moderate impact on the uncertainty associated with settlement prediction, as it has a correlation coefficient of 0.169.

As discussed earlier, uncertainty estimation of q and N varies considerably (i.e. the COV varies from 5 to 14% for q and from 27 to 85% for N). Consequently, it is worthwhile to carry out a parametric study to examine the effect of changing the COV for q and N on the magnitude of settlement prediction for the numerical example. Using the ANN-based stochastic approach that incorporates parameter uncertainty, two different combinations of the values of the COV for q and N are examined. The minimum values recommended in the literature for the COV of q and N (i.e. 5% for q and 27% for N) are used for one trial and the maximum values (i.e. 14% for q and 85% for N) are used for the other. The probability of non-exceedance for the predicted settlement using the three different combinations of the COV for q and N are shown in Table 7.2.

It can be seen from Table 7.2 that the predicted settlement becomes more conservative as the COV for q and N increases. For example, if the probability of non-exceedance required in the design of a footing under consideration is 95%, the predicted settlement will not exceed 29.2 mm when the COVs for q and N are equal to 5 and 27%, respectively. In other words, there is a 5% chance that the predicted settlement will exceed 29.2 mm when the COVs for q and N are equal to 5 and 27%, respectively. However, to achieve the same level of risk, this settlement will exceed 31.3 mm when the COVs for q and N are increased to 10 and 30%, respectively. Moreover, for a level of risk of 5%, the predicted settlement will exceed 42.6 mm when the COVs for q and N are increased further to be equal to 14 and 85%, respectively. The results in Table 7.2 also illustrate that, to achieve a level of risk of only 5%, the predicted settlement obtained using the maximum combination of the COV for q and N (i.e. 14% for q and 85% for N) is approximately 45% more than the predicted settlement obtained when the minimum combination of the COV for q and N (i.e. 5% for q and 27% for N) is used. This suggests that estimating the correct values of the uncertainty associated with q and N is very important, as uneconomical design of footings results from increasing values of the COVs for q and N.

7.5.2 Estimation of prediction method uncertainty

As discussed earlier, in order to obtain a stochastic solution for settlement prediction that incorporates the prediction method uncertainty, the PDF of k is needed. In this research, the PDF of k for the ANN method is obtained from the 189 data records that are used in Chapter 5 for the development of the ANN model. The values of k are found to fall within the range 0.3 to 10.4. The mean value of k is 1.4 and the standard deviation is 1.3. As mentioned above, the stochastic solution that incorporates the prediction method uncertainty relies on the estimation of the PDF of k. Consequently, false estimation of the PDF of k will affect the final results of the stochastic predicted settlements. If the data used to estimate the PDF of k contain outliers, the distribution of k will be severely affected. As a result, it is necessary to exclude any possible outliers from the data used to estimate the PDF of k. The box plot method (Kotzais et al. 1990), as proposed by Cherubini (2000), is used for this purpose in this research. As part of the method, the central tendency of k is indicated by the median, whereas its spread is indicated by the lower (Q_L) and upper (Q_U) quartiles. Points whose values are either less than (Q_L–1.5IQD) or greater than (Q_U+1.5IQD), where IQD is the interquartile distance and is equal to (Q_U– Q_L), are considered to be outliers. Such a plot is shown in Figure 7.4 for the available data. It can be seen that some points are greater than (Q_U+1.5IQD). Consequently, these data points may be considered to be outliers and are omitted from the data used to estimate the PDF of k. The number of outliers are found to be 20 out of 189 data records, resulting in 169 data records that are used to estimate the PDF of k. The software @Risk is again used to determine the PDF that provides the best fit to the remaining 169 data points. As mentioned in Chapter 5, for a given set of data values, @Risk can identify the probability distribution that best fits these values from 38 candidate distributions and provides the statistical properties that describe the distribution. The theoretical distribution that is found to best match the actual distribution of k is the Weibull distribution (Figure 7.5). The statistical properties of the Weibull distribution obtained are given in Table 7.3. It can be seen that removing the outliers from the analysis of k resulted in a reduction in the mean and standard deviation of k from 1.4 and 1.3 to 1.06 and 0.53, respectively.

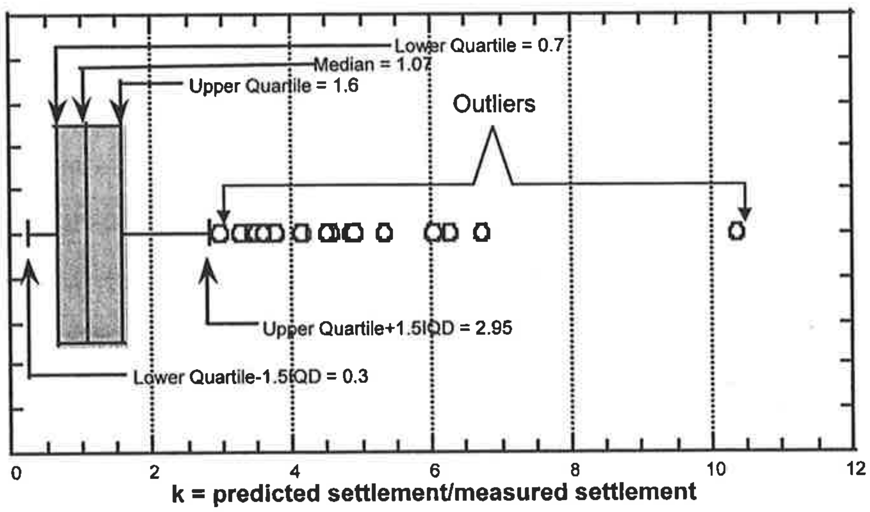


Figure 7.4: Box plot for 189 data records of k = predicted settlement/measured settlement

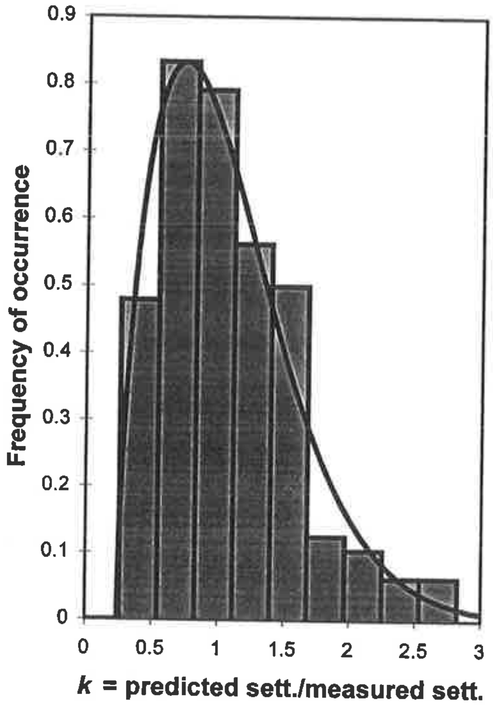


Figure 7.5: Weibull distribution of k

Table 7.3: Weibull distribution parameters of k

Statistical parameter	Value of k (without outliers)	Value of k (with outliers)
Minimum	0.25	0.30
Maximum	2.80	10.4
Mean	1.06	1.40
Standard deviation	0.53	1.30
Shape parameter (α)	1.59	N/A
Scale parameter (β)	0.91	N/A

N/A = not applicable

The numerical example is re-calculated using the procedure outlined previously for incorporating prediction method uncertainty. The Monte Carlo simulation is repeated 1,400 times until a convergence limit equal to 1% is achieved. The settlements obtained are used to plot the cumulative probability distribution curve from which different probabilities of non-exceedance or levels of risk for predicted settlement are obtained. The results are summarised in Table 7.4 and shown in Figure 7.6, which also includes the obtained by considering only the parameter uncertainty.

Table 7.4: Predicted settlements accounting for prediction method uncertainty for the numerical example

Probability of occurrence (%)	Predicted settlement (mm)
75	15.4
80	16.1
85	17.0
90	18.0
95	19.3

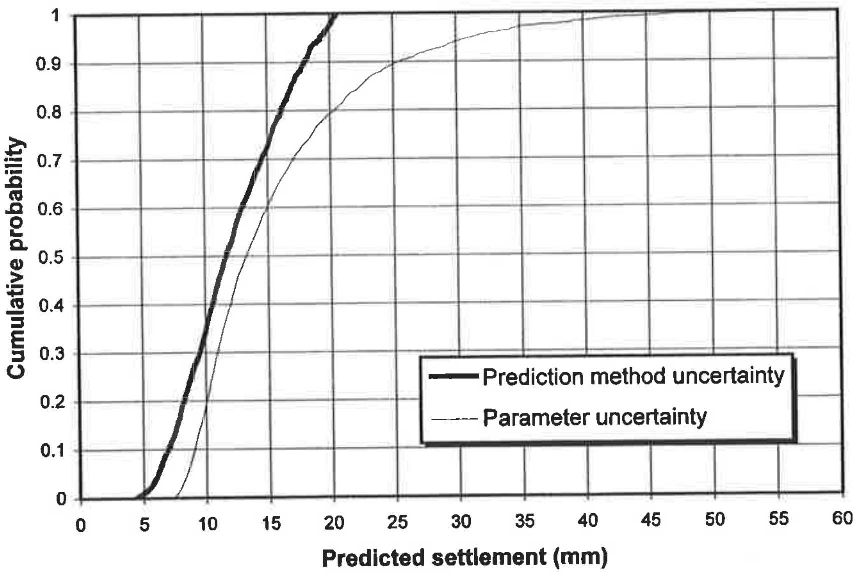


Figure 7.6: Cumulative probability distribution incorporating parameter and prediction method uncertainties for the numerical example

It can be seen from Table 7.4 and Figure 7.6 that the stochastic solution that incorporates the prediction method uncertainty is less conservative than that obtained when parameter uncertainty was considered. For example, there is a 5% level of risk that the predicted settlement will exceed 19.3 mm when the stochastic solution that incorporates the prediction method uncertainty is used. However, to achieve the same level of risk, this settlement will exceed 29.2 mm (for COVs of q and N equal to 5 and 27%), 31.3 mm (for COVs of q and N equal to 10 and 30%), and 42.6 mm (for COVs of q and N equal to 14 and 85%), respectively, when the stochastic solution that incorporates the parameter uncertainty is used, resulting in an increase in the predicted settlement of approximately 50%, 60% and 120%, respectively. This result is surprising, as the solution that incorporates the prediction method uncertainty includes both parameter and model uncertainties and would thus be expected to produce more conservative results. This result may be attributed to the fact that the estimation obtained from the literature related to parameter uncertainty (i.e. uncertainty associated with q and N), is most likely to be conservative as a result of including the soil spatial variation from one site to another in the evaluation of the COVs of q and N. Even the lower end of the variability of q and N appears to be conservative. Consequently, for an individual case of settlement prediction, different results (Table 7.2) are most likely to be obtained from different evaluators depending on how differently q and N are evaluated. On the other hand, the statistical properties and distribution used to estimate prediction method uncertainty (i.e. uncertainty associated with k) are based on actual measured settlements.

The above results imply that collecting as much data as possible, which enables accurate characterisation of parameter or prediction method uncertainties, is very important, as it has significant implications on the design settlement obtained. The results also imply that the stochastic solution using k is probably preferable, as it is easier to collect data on actual and predicted settlements, which enables the uncertainty associated with k to be characterised, rather than data which enable the uncertainty associated with q and N to be quantified. In addition, the stochastic approach using k already includes parameter uncertainty, as it sums the model and parameter uncertainties.

7.6 Stochastic Settlement Prediction Design Charts

The stochastic simulation that incorporates the prediction method uncertainty is used to develop a generic set of stochastic design charts based on the ANN model for routine use in practice. The procedure that is used to develop the charts is as follows:

1. A random synthetic value of predicted settlement is generated between the ranges used in Chapter 5 for the development of the ANN model;

2. The procedure for obtaining the corresponding PDF and CDF of predicted settlements as a result of the uncertainties associated with the prediction method outlined previously is applied;

3. From the above PDF or CDF, the 75%, 80%, 85%, 90%, and 95% probabilities of non-exceedance are determined;

4. Another random synthetic value of predicted settlement is generated by increasing the value generated in Step 1 by 5% of the total range between the minimum and maximum values used for the development of the ANN model;

5. Steps 2 to 4 are repeated until the maximum synthetic value of predicted settlement is reached; and

6. For each probability level of non-exceedance, the synthetic deterministic settlements are plotted against stochastic settlements and a set of design charts are obtained, as shown in Figure 7.7.

For any individual case of settlement prediction within the ranges of the data used for the ANN model development, the deterministic settlement can be obtained from the ANN model and the corresponding stochastic settlement can be obtained readily from Figure 7.7, accounting for a certain desired probability of non-exceedance. For example, if the deterministic ANN model predicts a settlement of 22 mm and reliability levels (i.e. probabilities of non-exceedance) of 90% and 95% are required, the corresponding design settlements are 30 mm and 32 mm, respectively.

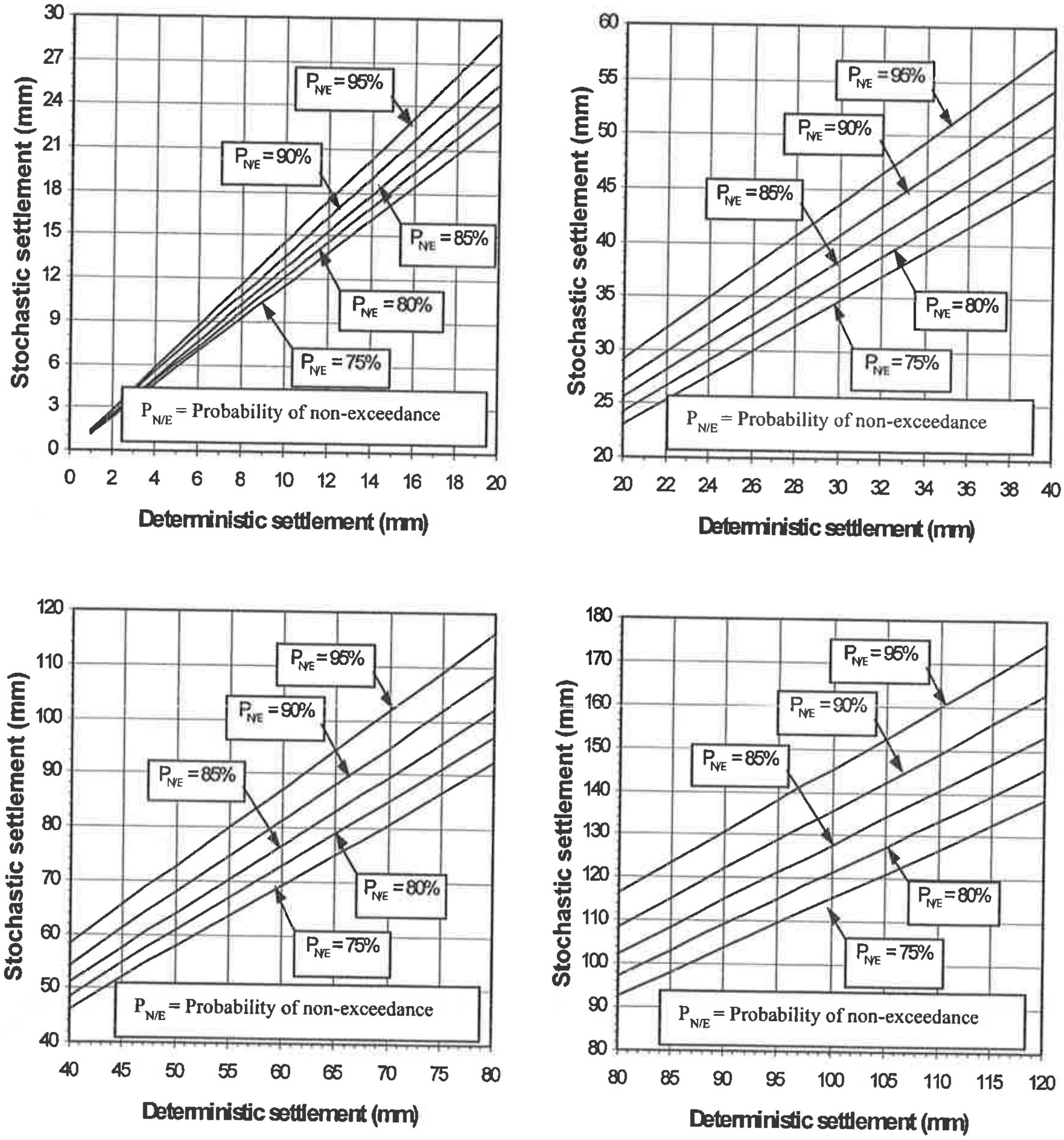


Figure 7.7: Stochastic ANN-based design charts for settlement prediction

7.7 Summary and Conclusions

Stochastic approaches that utilise the Monte Carlo technique were used to generate stochastic settlement prediction of shallow foundations on granular soils from an artificial neural network (ANN) model. The proposed stochastic approaches incorporate either parameter uncertainty or prediction method uncertainty (parameter uncertainty + model uncertainty) and enable the uncertainty associated with predicted settlements to be quantified in the form of a cumulative probability distribution function that provides the designer with the level of risk associated with exceeding a given predicted settlement. A parametric study was also carried out to examine the effect of varying the uncertainty associated with the factors affecting settlement (i.e. coefficient of variation, COV, for the imposed load, q, and soil property, N) on the uncertainty of the predicted settlements. The proposed stochastic approaches compared were applied to a numerical example of settlement prediction. A series of ANN-based design charts that incorporate prediction method uncertainty were developed for routine use in practice.

The results of the numerical example that incorporates parameter uncertainty indicated that there was a probability of approximately 50% that the settlement could be higher than the deterministic estimation with COVs of 10% and 30% for q and N, respectively. The results also indicated that over the range of COVs for q and N suggested in the literature, the design settlements ranged from 29.2 mm to 42.6 mm for a non-exceedance probability of 95%. These results indicated that the uncertainties associated with q and N can considerably affect settlement and thus, they should not be neglected in the analysis and simulation of settlement prediction. This also implied that it is important to collect sufficient data to characterise the uncertainty associated with q and N, as the results obtained were very sensitive to these variables. The ANN-based stochastic approach that incorporates prediction method uncertainty resulted in less conservative settlement prediction, despite the fact that this approach includes both parameter and model uncertainties and thus would be expected to produce more conservative results. This was attributed to the fact that parameter uncertainties were obtained from a subjective evaluation of the COVs of q and N published in the literature, which are likely to be conservative. On the contrary, prediction method uncertainty (i.e. uncertainty associated with k) was obtained from measured data and not from values published in the literature. Furthermore, it is easier to obtain data to characterise the uncertainty associated with k than with q and N. Finally, the charts developed in this work can be used to predict settlements for a certain desired reliability level given the deterministic settlement predicted from the ANN model developed in Chapter 5, which will be a useful tool in the design of shallow foundations on cohesionless soils.

Chapter 8

Summary, Conclusions and Recommendations

8.1 General

Settlement analysis of shallow foundations on cohesionless soils, as with many situations in geotechnical engineering, is a complex problem that is not well understood. For most mathematical models that attempt to solve this problem, the lack of physical understanding is usually supplemented by either simplifying the problem or incorporating several assumptions into the models. Conventional mathematical models for settlement prediction of shallow foundations on cohesionless soils rely on assuming the form of the model in advance, and the unknown model parameters are determined by minimising an error function between model predictions and the known measured values. Consequently, prior knowledge regarding the relationship between model inputs and the corresponding outputs is needed. In the case of settlement of shallow foundations on cohesionless soils, such knowledge is not yet entirely understood. Consequently, model performance may be potentially compromised, as the form of the model chosen may be sup-optimal. In contrast, as shown in this thesis, artificial neural networks (ANNs) use the data alone to determine the structure of the model as well as the unknown model parameters. The ANN modelling philosophy is similar to most available methods for settlement prediction in the sense that both are attempting to capture the relationship between a set of model inputs and their corresponding outputs. However, unlike most available methods, ANNs do not need prior knowledge about the nature of the relationship between model inputs and their corresponding outputs as ANNs use the data alone to capture this relationship, as mentioned earlier. This is an essential benefit that enables ANNs to overcome the limitations of existing methods. Moreover, ANNs can always be updated to obtain better results by presenting new calibration data records, as they become available. One limitation of ANNs, however, is that often the relationship between the input parameters and the output, are complex and cannot be described in a tractable fashion. In addition, the success of ANNs in finding this relationship is not always guaranteed.

In this research, the feasibility of using ANNs for predicting the settlement of shallow foundations on cohesionless soils has been investigated. In the following, summary, contributions of this thesis, recommendations for future work and the main conclusions of this research are presented.

8.2 Summary

This study has investigated the feasibility of using artificial neural networks (ANNs) for settlement prediction of shallow foundations on cohesionless soils. An ANN model, which was found to outperform the most commonly used traditional methods, has been developed for routine use in practice. In addition, stochastic analysis has been applied to the ANN model and a set of stochastic design charts that incorporate the uncertainty associated with the ANN method has been developed and provided.

Chapter 2 detailed the more important features associated with ANNs. These include the structure and operation of ANNs, classification of different types of ANNs and development of ANN models. For the structure and operation of ANNs, it was shown that ANNs consist of a number of processing elements or nodes that are arranged in layers: an input layer, an output layer and one or more intermediate layers called hidden layers. It was also shown that ANNs learn by presenting training data from which the ANN network adjusts its weights until it can find a set of weights that produce the optimum input/output data mappings. In relation to the classification of ANNs, it was shown that ANNs can be categorised on the basis of two major categories: (i) the learning rule used and (ii) the connections between nodes. In relation to learning rules, it was shown that ANNs can be divided into supervised and unsupervised networks. In relation to connections between nodes, it was shown that ANNs can be divided into feed-forward and feedback networks. For the development of ANNs, many factors that affect the development of ANN models were addressed. These include the determination of model inputs, division of data, data pre-processing, determination of model architecture, model optimisation, stopping criteria and model validation.

Chapter 3 provided an overview of some of the more relevant ANN applications in geotechnical engineering. These include prediction of pile capacity, predicting the settlement of foundations, modelling soil properties and behaviour, determination of soil liquefaction, site characterisation, modelling earth retaining structures, evaluating stability of slopes and design of tunnels and underground openings.

Chapter 4 discussed the causes of settlement of shallow foundations. The chapter described the factors affecting settlement of shallow foundations on cohesionless soils, which were divided into primary and secondary factors. The primary factors were the footing width, footing net applied pressure and soil compressibility. The secondary factors were the depth of the water table, time dependence, footing geometry, depth of footing embedment and the thickness of the soil layer. This chapter also described and discussed some of the most commonly used and more relevant methods for settlement prediction of shallow foundations on cohesionless soils. These include the methods proposed by Meyerhof (1965), Schultze and Sherif (1973) and Schmertmann et al. (1978).

Chapter 5 detailed the analysis of data and development of ANN models for settlement prediction of shallow foundations on cohesionless soils. This chapter also presented a comparison between the results obtained using ANN and traditional methods. The analyses carried out in this chapter yielded the following results and conclusions:

- ANNs have the ability to predict the settlement of shallow foundations on cohesionless soils with a high degree of accuracy for predicted settlements ranging from 0.6 to 121.0 mm.

- Good performance of ANN models on training, testing and validation sets does not guarantee the robustness of the predictive ability of the models over a range of data similar to that used for training.

- If cross-validation is used as the stopping criterion, reducing the number of ANN connection weights or changing the type of connection between nodes (e.g. cascaded and direct connections) does not appear to improve ANN model robustness.

- It is recommended that the capability of the software used for developing ANN models be considered carefully and that a parametric study to check model robustness be used in order to ensure that a model can be used for predictive purposes with confidence.

- The optimum model developed in this work was obtained using 2 hidden layer nodes, a momentum term of 0.8, a learning rate of 0.2, the tanh transfer function in the hidden layer and the sigmoid transfer function in the output layer.

- There is a direct relationship between the consistency of the statistics between training, testing and validation sets and the consistency in model performance. Consequently, the statistical properties of the various data subsets should be taken into account as part of any data division procedure to ensure that the best possible model is developed, given the available data.

- The proportion of the data used for training, testing and validation appears to have an effect on model performance. However, there appears to be no clear relationship between the proportion of the data used in each of the subsets and model performance, although in the trials conducted in this work, the optimal model performance was obtained when 20% of the data were used for validation and 70% of the remaining data were used for training and 30% for testing.

- The data division approach using a SOM and the new approach using fuzzy clustering introduced in this research appear to be suitable methods of data division.

- The distribution transformation method for input variables does not appear to improve the performance of ANN models.

- The SPT blow count, footing net applied pressure and footing width have the most significant impact on settlement with relative importance levels equal to 30.8, 28.6 and 22.4%, respectively. On the other hand, the footing embedment ratio and footing geometry have less impact on settlement with relative importance levels equal to 14.8 and 3.1%, respectively.

- The ANN method outperforms the traditional methods considered for an independent validation set with r = 0.905, RMSE = 11.04 mm and MAE = 8.78 mm, while these measures were: r = 0.440, 0.729 and 0.798; RMSE = 25.72, 23.55 and 23.67 mm and MAE = 16.59, 11.81 and 15.69 mm when the method proposed by Meyerhof (1965), Schultze and Sherif (1973) and Schmertmann et al. (1978), respectively, are used.

- The ANN model could be translated into a relatively simple and practical formula from which settlement can be calculated as follows:

$$S_p = 0.6 + \left[\frac{120.4}{1 + e^{(0.312 - 0.725 \tanh x_1 + 2.984 \tanh x_2)}}\right]$$

and

$$x_1 = 0.1 + 10^{-3}\left[3.8B + 0.7q + 4.1N - 1.8(L/B) + 19(D_f/B)\right]$$

$$x_2 = 10^{-3}\left[0.7 - 41B - 1.6q + 75N - 52(L/B) + 740(D_f/B)\right]$$

where:

S_p	= predicted settlement (mm);
B	= footing width (m);
q	= net applied footing load (kPa);
N	= average SPT blow count;
L/B	= footing geometry; and
D_f/B	= footing embedment ratio.

- The ANN model could be used to generate a series of design charts (Appendix G) from which settlement can be obtained easily.

Chapter 6 detailed the development of neurofuzzy networks for settlement prediction of shallow foundations on cohesionless soils. Neurofuzzy networks were used in this chapter to investigate their capability for predicting the settlement of shallow foundations on cohesionless soils and to assist with providing a better understanding of the relationship between settlement and the factors affecting it. The sensitivity of the neurofuzzy models to a number of stopping criteria [i.e. Bayesian Information Criterion (BIC), Akaike's Information Criterion (AIC) and Final Prediction Error (FPE)] was investigated. The obtained models were assessed in terms of model accuracy, model parsimony and model transparency. The optimum neurofuzzy model obtained was compared with the best back-propagation MLP model obtained in Chapter 5. A number of conclusions were derived from this chapter. These include:

- Neurofuzzy models have the ability to accurately predict the settlement of shallow foundations on cohesionless soils and are capable of extracting rules from the data that make physical sense, which could be used to gain understanding in situations where data are available but physical relationships are not well understood. In addition, neurofuzzy networks can be modified by incorporating available engineering knowledge to improve model performance and enhance the interpretation of the constructed model.

- The footing width, footing net applied pressure and average SPT blow count were found to be the most significant factors affecting settlement. This is in agreement with the results found in Chapter 5.

- All neurofuzzy models were found to be comparable in terms of prediction accuracy, even though the model that uses the FPE was found to perform marginally better than the other models.

- All neurofuzzy models were found to be comparable in terms of model parsimony, even though the model that uses the AIC was found to be more parsimonious than the other models with the lowest number of connection weights.

- The neurofuzzy models that use the BIC and AIC were found to be comparable in terms of model transparency and more transparent than the model that uses the FPE, as they have fewer number of fuzzy rules. This was attributed to the fact that the BIC and AIC penalise complex models to a greater extent.

- The optimum neurofuzzy and MLP models were found to be comparable in terms of model accuracy, although the MLP model was found to perform slightly better than the neurofuzzy model.

- The optimum neurofuzzy model was found to be more parsimonious than the back-propagation MLP with fewer model inputs and connection weights.

- The optimum neurofuzzy model was found to be more transparent than the back-propagation MLP model as it was able to describe the relationship between the model inputs and corresponding output using a set of fuzzy rules.

Finally, Chapter 7 examined the use of stochastic simulation in the analysis of ANN settlement prediction and produced a set of stochastic design charts for routine use in practice. It was found that the stochastic analysis was essential to incorporate the uncertainties associated with the predicted settlements. It was also found that the charts developed are a useful tool for the design of shallow foundations on cohesionless soils, as they provide the designer with the level of risk associated with predicted settlements and thus can give a more realistic indication of what the actual settlement might be.

8.3 Original Contributions of the Research

To the author's best knowledge, this thesis has made the following original contributions:

1. The suitability of using two different types of artificial neural networks (ANNs), i.e. multi-layer perceptrons (MLP) trained with the back-propagation algorithm and B-spline neurofuzzy networks trained with the adaptive spline modelling of observation data (ASMOD) algorithm, for predicting the settlement of shallow foundations on cohesionless soils has been assessed.

2. New guidelines to assist in the development of ANN models have been provided and in particular, information is given on:

- How to test the robustness of ANN models. This is a new test that has never before been proposed in any field of ANN application.
- How to test the capability of the software used for developing ANN models.
- The effect of using different internal network parameters on ANN model performance, including the number of hidden layer nodes, learning rate, momentum term and transfer function.
- The effect of the statistical consistency of the training, testing and validation data sets on ANN model performance.
- The effect of the proportion of the data used for training, testing and validation on ANN model performance.
- Four different methods for ANN data division, including random data division, data division to ensure statistical consistency of the subsets needed for ANN model development, data division using self-organising maps (SOMs) and a new data division method using fuzzy clustering. The second and third data division methods have been used to a limited extent in water engineering but never in geotechnical engineering, whereas the fourth method (fuzzy clustering) has never before been applied to the development of ANN models in any field.
- The effect of the distribution transformation method for input variables on ANN model performance.
- How to test the relative importance of ANN model inputs using sensitivity analysis.

3. An ANN model that has been found to outperform the most commonly used traditional methods for settlement prediction of shallow foundations on cohesionless soils has been developed and a database that contains a total of 189 case records of measured settlements has been provided.

4. A simple and practical formula that is based on the ANN model for settlement prediction of shallow foundations on cohesionless soils has been introduced and a series of design charts that are based on the ANN model for settlement prediction have been generated.

5. It has been shown that neurofuzzy networks can be used to extract rules from data that make physical sense, which could be used to gain understanding in situations where data are available but physical relationships are not well understood. In addition, it has also been shown that neurofuzzy networks can be modified by incorporating available expertise to improve model performance and enhance the interpretation of constructed model.

6. It has been demonstrated that ANNs can be used to provide valuable information about the relationship between model inputs and their corresponding outputs in the form of a relatively simple equation (in case of the back-propagation MLP) or a set of fuzzy rules (in case of neurofuzzy networks), and thus ANNs do not have to be treated as a "*black boxes*".

7. The effect of the parameter uncertainty and prediction method uncertainty in the analysis of ANN settlement prediction of shallow foundations on cohesionless soils has been investigated, which, until now, has never been applied to ANN models in geotechnical engineering. In addition, the effect of varying the amount of the parameter uncertainty on the magnitude of the predicted settlement has been investigated.

8. A set of useful stochastic design charts that are based on the ANN method has been developed and provided for routine use in practice.

8.4 Recommendations for Future Work

1. Despite the good performance of ANNs in this work and in many situations in geotechnical engineering, they suffer from a number of shortcomings, notably, the lack of theory to help with their development, the fact that success in finding a good solution is usually obtained by trial-and-error and their limited ability to explain the way they use the available information to arrive at such solutions. Consequently, there is a need for a comprehensive set of guidelines to assist in the development of ANNs, even though, this thesis has provided a significant contribution in this regard.

There is also a need for more future research into methods that provide a comprehensive explanation of how ANNs arrive at a prediction.

2. The guidelines proposed in this research for the development of ANN models need to be applied to other case studies to investigate whether they can be considered as generic or whether they are specific only to the case study considered in this thesis.

3. The ANN model developed in this research is based on 189 data records and is suitable for use in an interpolative sense within the range of the data used for model calibration. Consequently, like all empirical models, the range of applicability of the developed ANN model is constrained by the data used in the model calibration phase. In order to update the model and make it more accurate in the future, it is desirable to include additional data so that the model can accurately predicts settlements across a wider range of footing sizes and soil conditions.

4. Although the ANN model developed in this research was found to outperform all of the traditional methods for settlement prediction examined, it is based on in-situ measurements of soil compressibility that use the SPT. Whilst the SPT is not the most accurate in-situ method for soil compressibility, a model based on the CPT, for example, is likely to produce even better predictions. However, before such a model can be developed, it is necessary to obtain an extensive set of CPT-based data and corresponding settlement measurements. At present, in the published literature, such data are very limited indeed.

8.5 Conclusions

From the analyses presented in this thesis, it can be concluded that:

- Artificial neural networks (ANNs) have the ability to predict the settlement of shallow foundations on cohesionless soil with a high level of accuracy and outperform the most commonly used traditional methods.

- The new proposed approach for data division that is based on fuzzy clustering appears to be applicable and provides a useful contribution to the development of ANN models.

- ANNs need not to be treated as a "*black boxes*" as it has been demonstrated that ANNs can be translated into a relatively simple equation (in the case of the back-propagation MLP model) or a set of fuzzy rules (in the case of neurofuzzy networks) that give valuable information regarding the relationships between the model inputs and their corresponding outputs.

- The stochastic analysis for ANN models proposed in this study allows the level of risk associated with predicted settlement to be quantified in the form of a set of stochastic design charts that provide the designer with more rational settlement prediction. The charts developed in this work can be used to predict settlements for a desired reliability level given the deterministic settlement predicted from the ANN model. This will be a useful tool in the design of shallow foundations on cohesionless soils.

References

Abu-Kiefa, M. A. (1998). "General regression neural networks for driven piles in cohesionless soils." *J. Geotech. & Geoenv. Eng.*, 124(12), 1177-1185.

Agrawal, G., Chameau, J. A., and Bourdeau, P. L. (1997). "Assessing the liquefaction susceptibility at a site based on information from penetration testing." *Artificial neural networks for civil engineers: Fundamentals and applications*, N. Kartam, I. Flood, and J. H. Garrett, eds., ASCE, New York, 185-214.

Agrawal, G., Weeraratne, S., and Khilnani, K. (1994). "Estimating clay liner and cover permeability using computational neural networks." *Proc. First Congress on Computing in Civil Eng.*, Washington, June 20-22.

Ali, H. E., and Najjar, Y. M. (1998). "Neuronet-based approach for assessing liquefaction potential of soils." *Transportation Research Record No. 1633*, 3-8.

Alpan, I. (1964). "Estimating the settlements of foundations on sand." *Civil Engineering & Public Works Review*, 59(Nov.), 1415-1418.

Altrock, C. V. (1995). *Fuzzy logic and neurofuzzy applications explained*, Prentice-Hall, Englewood Cliffs, New Jersey.

Amar, S., and Baguelin, F. (1984). "Experimental investigation of the short- and long-term behaviour of shallow foundations." *Proc. 8th Regional Conf. for Africa on Soil Mech. & Found. Eng.*, Harare.

Amari, S. I., Murata, N., Muller, K. R., Finke, M., and Yang, H. H. (1997). "Asympotic statistical theory of overtraining and cross-validation." *IEEE Transactions on Neural Networks*, 8(5), 985-996.

American Petroleum Institute. (1984). *RP2A: Recommended practice for planning, designing and constructing fixed offshore platforms*, Washington, DC.

Ang, A. H., and Tang, W. H. (1984). *Probability concepts in engineering planning and design*, John Wiley & Sons, New York.

Arnold, M. (1980). "Prediction of footing settlements on sand." *Ground Engineering*, 3, 40-49.

Arnold, M. A. (1999). "Artificial neural networks applied to the prediction of settlements of shallow foundations in granular soils," MEngSc. thesis, James Cook University, Queensland.

ASCE Task Committee of Artificial Neural Networks in Hydrology (2000). "Artificial neural networks in hydrology. I: Preliminary concepts." *J. Hydrologic Eng.*, 5(2), 115-123.

Auvinet, G., and Rossa, O. (1991). "Reliability of foundations on soft soils." *6th Int. Conf. on Applications of Statistics and Probability in Civil Eng.*, Mexico, 2, 769-774.

Barthur, R. (1997). "Design of residential footings built on expansive soil using probabilistic methods," MSc. thesis, The University of Adelaide, Adelaide.

Basheer, I. A. (1998). "Neuromechanistic-based modeling and simulation of constitutive behaviour of fine-grained soils," PhD dissertation, Kansas State University, Manhattan, KS.

Basheer, I. A., and Najjar, Y. M. (1995). "A neural-network for soil compaction." *Proc. 5th Int. Symp. Numerical Models in Geomechanics*, Roterdam: Balkema, 435-440.

Basheer, I. A., and Najjar, Y. M. (1998). "Modeling cyclic constitutive behavior by neural networks: Theoretical and real data." *Proc. 12th Engineering Mechanics Conf.*, La Jolla, California, May 17-20, 952-955.

Basheer, I. A., Reddi, L. N., and Najjar, Y. M. (1996). "Site characterization by neuronets: An application to the landfill siting problem." *Ground Water*, 34, 610-617.

Bazaraa, A. R. S. S. (1967). "Use of the standard penetration test for estimating settlements of shallow foundations on sand," PhD thesis, University of Illinois, Urbana, Illinois.

Bebis, G., and Georgiopoulos, M. (1994). "Feed-forward neural networks: Why network size is so important." *IEEE Potentials*, October/November, 27-31.

Berardi, R., and Lancellotta, R. (1994). "Prediction of settlements of footings on sand: Accuracy and reliability." ASCE Geotechnical Special Publication 40, 640-651.

Berke, L., and Hajela, P. (1991). "Application of neural networks in structural optimization." NATO/AGARD Advanced Study Institute, 23(I-II), 731-745.

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.

Braddock, R. D., Kremmer, M. L., and Sanzogni, L. (1998). "Feed-forward artificial neural network model for forecasting rainfall run-off." *Environmetrics*, 9, 419-432.

Breiman, L. (1994). "Comment on 'Neural networks: A review from a statistical' by B. Cheng and D. M. Titterington." *Statistical Science*, 9(1), 38-42.

Briaud, J. L., and Gibbens, R. (1999). "Behaviour of five large spread footings in sand." *J. Geotech. & Geoenv. Eng.*, 125(9), 787-796.

Broms, B. B., and Lim, P. C. (1988). "A simple pile driving formula based on stress-wave measurements." *Proc. 3rd Int. Conf. on the Application of Stress-Wave Theory to Piles*, Vancouver, 591-600.

Brown, M., and Harris, C. (1994). *Neurofuzzy adaptive modelling and control*, Prentice-Hall, Englewood Cliffs, New Jersey.

Brown, M., and Harris, C. J. (1995). "A perspective and critique of adaptive neurofuzzy systems used for modelling and control applications." *Int. J. Neural Systems*, 6(2), 197-220.

Brzakala, W., and Pula, W. (1996). "A probabilistic analysis of foundation settlements." *Computers and Geotechnics*, 18(4), 291-309.

Burbidge, M. C. (1982). "A case study review of settlements on granular soil," MSc. thesis, Imperial College of Science and Technology, University of London, London.

Burke, L. I., and Ignizio, J. P. (1992). "Neural networks and operations research: An overview." *Computer and Operations Research*, 19(3/4), 179-189.

Burland, J. B. (1973). "Shaft friction of piles in clay." *Ground Engineering*, 6(3), 1-15.

Burland, J. B., and Burbidge, M. C. (1985). "Settlement of foundations on sand and gravel." *Proc. Institution of Civil Engineers*, London, 78-Part 1, 1325-1381.

Cai, S., Toral, H., Qiu, J., and Archer, J. S. (1994). "Neural network based objective flow regime identification in air-water two phase flow." *Canadian J. Chemical Eng.*, 72, 440-445.

Cal, Y. (1995). "Soil classification by neural-network." *Advances in Engineering Software*, 22(2), 95-97.

Campolo, M., Soldati, A., and Andreussi, P. (1999). "Forecasting river flow rate during low-flow periods using neural networks." *Water Resources Research*, 35(11), 3547-3552.

Castellano, G., Fanelli, A. M., and Pelillo, M. (1997). "An iterative pruning algorithm for feedforward neural networks." *IEEE Transactions on Neural Networks*, 8(3), 519-531.

Caudill, M. (1988). "Neural networks primer, Part III." *AI Expert*, 3(6), 53-59.

Chan, L. W., and Fallside, F. (1987). "An adaptive training algorithm for back-propagation networks." *Computer Speech and Language*, 2, 205-218.

Chan, W. T., Chow, Y. K., and Liu, L. F. (1995). "Neural network: An alternative to pile driving formulas." *J. Computers and Geotechnics*, 17, 135-156.

Cherubini, C. (2000). "General regression neural networks for driven piles in cohesionless soils: Discussion by Claudio Cherubini." *J. Geotech. & Geoenviron. Eng.*, 126(9), 855-856.

Cherubini, C., and Greco, V. R. (1991). "Prediction of spread footing settlements on sand." *6th Int. Conf. on Application of Statistics and Probability in Civil Eng.*, Mexico, 2, 761-767.

Chester, D. L. (1990). "Why two hidden layers are better than one." *Int. Joint Conf. on Neural Networks*, 1, 265-268.

Christian, J. T., and Carrier, W. D. (1978). "Janbu, Bjerrum and Kjaernsli's chart reinterpreted." *Canadian Geotech. J.*, 15(1), 123-128.

Christian, J. T., Ladd, C. C., and Baecher, G. B. (1994). "Reliability applied to slope stability analysis." *J. Geotech. Eng.*, 120(12), 2180-2207.

Coduto, D. P. (1994). *Foundation design principles and practices*, Prentice-Hall, Englewood Cliffs, New Jersey.

Coduto, D. P. (1999). *Geotechnical engineering principles and practices*, Prentice-Hall, New Jersey.

Consoli, N. C., Schnaid, F., and Milititsky, J. (1998). "Interpretation of plate load tests on residual soil site." *J. Geotech. & Geoenv. Eng.*, 124(9), 857-867.

Corotis, R. B. (1972). "Statistical analysis of live load in column design." *J. Struct. Div.*, 98(ST8), 1803-1815.

Coyle, H. M., and Castello, R. R. (1981). "New design correlations for piles in sand." *J. Geotech. Eng.*, 107(7), 965-986.

Cybenko, G. (1989). "Approximation by superpositions of a sigmoidal function." *Mathematics of Control, Signals, and Systems*, 3, 303-314.

D'Appolonia, D. J., and D'Appolonia, E. (1970). "Use of SPT to estimate settlement of footings on sand." *Proc. Symposium of Foundation Interbedded Sands*, Division of Applied Geomechanics, Commonwealth Scientific & Industrial Research Organisation, Australia and the Australian Geomechanics Society, Perth, 16-22.

D'Appolonia, D. J., D'Appolonia, E., and Brissette, R. F. (1968). "Settlement of spread footings on sand." *J. Soil Mechanics & Found. Div.*, 94(SM3), 735-760.

Ellis, G. W., Yao, C., and Zhao, R. (1992). "Neural network modeling of the mechanical behavior of sand." *Proc. Engineering Mechanics*, 421-424.

Ellis, G. W., Yao, C., Zhao, R., and Penumadu, D. (1995). "Stress-strain modelling of sands using artificial neural networks." *J. Geotech. Eng.*, 121(5), 429-435.

Fahlman, S. E. (1988). "Fast learning variations on back-propagation: An empirical study." *Proc. 1988 Connectionist Models Summer School*, Pittsburgh, 38-51.

Fahlman, S. E., and Lebiere, C. (1990). "The cascade-correlation learning architecture." *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, ed., Morgan Kaufmann, San Mateo, California, 524-532.

Fang, H. (1991). *Foundation engineering handbook*, Van Nostrand Reinhold, New York.

Faraway, J., and Chatfield, C. (1998). "Time series forecasting with neural networks: A comparative study using the airline data." *Applied Statistics*, 47(2), 231-250.

Fausett, L. V. (1994). *Fundamentals neural networks: Architecture, algorithms, and applications*, Prentice-Hall, Englewood Cliffs, New Jersey.

Fenton, G. A. (2002). Personal communication.

Fenton, G. A., Paice, G. M., and Griffiths, D. V. (1996). "Probabilistic analysis of foundation settlement." *ASCE Geotechnical Special Publication*, 58/1(Jul 31-Aug 3), 657-665.

Finnhoff, W., Hergert, F., and Zimmermann, H. G. (1993). "Improving model selection by nonconvergent methods." *Neural Networks*, 6, 771-783.

Fletcher, G. F. A. (1965). "The standard penetration test, its uses and abuses." *J. Soil Mechanics and Found. Eng. Div.*, 91(SM4), 67-75.

Flood, I. (1991). "A Gaussian–based neural network architecture and complementary training algorithm." *Proc. Int. Conf. on Neural Networks*, New York, 1, 171-176.

Flood, I., and Kartam, N. (1994). "Neural networks in civil engineering I: Principles and understanding." *J. Computing in Civil Eng.*, 8(2), 131-148.

Fortin, V., Ouarda, T. B. M. J., and Bobee, B. (1997). "Comment on 'The use of artificial neural networks for the prediction of water quality parameters' by H. R. Maier and G. C. Dandy." *Water Resources Research*, 33(10), 2423-22424.

Fraser, R. A., and Wardle, L. J. (1975). "Raft foundations – Case study and sensitivity analysis." *Proc. 2nd Int. Conf. on Applications of Statistics and Probability in Soil and Structural Eng.*, Aachen, 89-117.

Frey, C. (1998). "Quantitative analysis of variability and uncertainty in energy and environmental systems." *Uncertainty modeling and analysis in civil engineering*, B. M. Ayyub, ed., CRC Press, Boca Raton, 381-423.

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

Ghaboussi, J., and Sidarta, D. E. (1998). "New nested adaptive neural networks (NANN) for constitutive modeling." *J. Computers and Geotechnics*, 22(1), 29-52.

Gibbs, H. J., and Holtz, W. G. (1957). "Research on determining the density of sands by spoon penetration testing." *Proc. 4th Int. Conf. Soil Mech. & Found. Eng.*, London, 1, 35-39.

Gifford, D. G., Wheeler, J. R., Kraemer, S. R., and McKown, A. F. (1987). "Spread footings for highway bridges." *Final Report, FHWA/RD-86/185*, Federal Highway Administration, Washington, DC.

Giraudel, J. L., and Lek, S. (2001). "A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination." *Ecological Modelling*, 146, 329-339.

Goble, G. G., Likins, G. E., and Rausche, F. (1975). "Bearing capacity of piles from dynamic measurements." *Final Report*, Dept. of Civil Eng., Case Western University.

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

Goh, A. T. C. (1994b). "Seismic liquefaction potential assessed by neural network." *J. Geotech. & Geoenv. Eng.*, 120(9), 1467-1480.

Goh, A. T. C. (1995a). "Back-propagation neural networks for modeling complex systems." *Artificial Intelligence in Engineering*, 9, 143-151.

Goh, A. T. C. (1995b). "Empirical design in geotechnics using neural networks." *Geotechnique*, 45(4), 709-714.

Goh, A. T. C. (1995c). "Modeling soil correlations using neural networks." *J. Computing in Civil Eng.*, 9(4), 275-278.

Goh, A. T. C. (1996a). "Neural-network modeling of CPT seismic liquefaction data." *J. Geotech. Eng.*, 122(1), 70-73.

Goh, A. T. C. (1996b). "Pile driving records reanalyzed using neural networks." *J. Geotech. Eng.*, 122(6), 492-495.

Goh, A. T. C., Wong, K. S., and Broms, B. B. (1995). "Estimation of lateral wall movements in braced excavation using neural networks." *Canadian Geotech. J.*, 32, 1059-1064.

Gordon, F. A., and Fletcher, F. (1965). "Standard penetration test: Its uses and abuses." *J. Soil Mechanics and Found. Div.*, 91(SM4), 67-75.

Greco, V. R., and Cherubini, C. (1993). "Probabilistic prediction of differential settlements in sand." *Probabilistic Methods in Geotechnical Engineering*, Li and Lo, eds., Balkema, Rotterdam, 303-308.

Gribb, M. M., and Gribb, G. W. (1994). "Use of neural networks for hydraulic conductivity determination in unsaturated soil." *Proc. 2nd Int. Conf. Ground Water Ecology*, Bethesda, 155-163.

Gupta, M. (1993). "Settlement of shallow footings on sand – An elastic stress-strain approach," MS thesis, North Carolina State University, Raleigh, NC.

Hammersley, J. M., and Handscomb, D. C. (1964). *Monte Carlo methods*, Wiley Press, New York.

Hammerstrom, D. (1993). "Working with neural networks." *IEEE Spectrum*, 30(7), 46-53.

Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*, MIT Press, Cambridge.

Hecht-Nielsen, R.. (1989). "Theory of the back-propagation neural network." *Proc. Int. Joint Conf. on Neural Networks*, Washington, DC, I, 593-605.

Hecht-Nielsen, R. (1987). "Kolmogrov's mapping neural network existence theorem." *Proc. First IEEE Int. Joint Conf. Neural Networks*, San Diego, CA, June 21-24, III-11-III-14.

Hecht-Nielsen, R. (1990). *Neurocomputing*, Addison-Wesely Publishing Company, Reading, MA.

Hegazy, T., Fazio, P., and Moselhi, O. (1994). "Developing practical neural network applications using back-propagation." *Microcomputers in Civil Engrg*, 9, 145-159.

Hertz, J. A., Krogh, A., and Palmer, R. G. (1991). *Introduction to the theory of neural computation*, Addison-Wesely Publishing Company, Red Wood City, California.

Hiley, A. (1922). "The efficiency of the hammer blow, and its effects with reference to piling." *Engineering*, 2, 673.

Holzlohner, U. (1984). "Settlement of shallow foundations on sand." *Jpn. Soc. Soil Mechan. Found. Eng.*, 24(4), 58-70.

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

Hubick, K. T. (1992). *Artificial neural networks in Australia*, Department of Industry, Technology and Commerce, Commonwealth of Australia, Canberra.

Hush, D. R., and Horne, B. G. (1993). "Progress in supervised neural networks." *IEEE ASSP Magazine*, January, 8-37.

Jacobs, R. A. (1988). "Incremental rates of convergence through learning rate adaptation." *Neural Networks*, 1, 295-307.

Jaksa, M. B. (1995). "The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay," PhD. thesis, The University of Adelaide, Adelaide.

Janbu, N. (1953). "Une analyse energetique du battage des pieux a l'aide de parametres sans dimension." *Norwegian Geotech. Inst.*, Oslo, Norway, 3, 63-64.

Jeyapalan, J. K., and Boehm, R. (1986). "Procedures for predicting settlements in sands." *Proc. Settlement of Shallow Foundations on Cohesionless Soils: Design and Performance*, Geotechnical Special Publication No. 5, ASCE, Seattle, Washington, 1-22.

Jorden, E. E. (1977). "Settlement in sand – Methods of calculating and factors affecting." *Ground Engineering*, 10(1), 30-37.

Journel, A. G., and Huijbregts, C. J. (1978). *Mining geostatistics*, Academic Press, London.

Juang, C. H., and Chen, C. J. (1999). "CPT-based liquefaction evaluation using artificial neural networks." *Computer-Aided Civil and Infrastructure Engineering*, 14(3), 221-229.

Karnin, E. D. (1990). "A simple procedure for pruning back-propagation trained neural networks." *IEEE Transactions on Neural Networks*, 1(2), 239-242.

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

Kavli, T. (1993). "ASMOD – an algorithm for adaptive spline modelling of observation data." *Int. J. Control*, 58(4), 947-967.

Kocjancic, R., and Zupan, J. (2000). "Modelling of the river flow rate: The influence of the training set selection." *Chemometric and Intelligent Laboratory Systems*, 54, 21-34.

Kohonen, T. (1982). "Self-organized formation of topologically correct feature maps." *Biological Cybernetics*, 43, 59-69.

Kohonen, T. (1997). *Self-organizing maps*, Springer-Verlag, Berlin Heildelberg.

Kotzais, P. C., Stamatopoulos, A. C., and Kountouris, P. J. (1990). "Exploratory graphics and geotechnical data: Some introductory remarks." *J. Geotech. Eng.*, 21(2), 127-143.

Krizek, R. J., Corotis, R. B., and El-Moursi, H. H. (1977). "Probabilistic analysis of predicted and measured settlements." *Canadian Geotech. J.*, 14(17), 17-33.

Kudrycki, T. P. (1988). "Neural network implementation of a medical diagnosis expert system," MS thesis, College of Engineering, University of Cincinnati.

Lachtermacher, G., and Fuller, J. D. (1994). "Back-propagation in hydrological time series forecasting." *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, K. W. Hipel, A. I. McLeod, U. S. Panu, and V. P. Singh, eds., Kluwer Academic, Dordrecht.

Lapedes, A., and Farber, R. (1988). "How neural networks work." *Neural Information Processing Systems*, American Institute of Physics, 442-456.

Lee, C., and Sterling, R. (1992). "Identifying probable failure modes for underground openings using a neural network." *Int. J. Rock Mechanics and Mining Science & Geomechanics Abstracts*, 29(1), 49-67.

Lee, I. K., White, W., and Ingles, O. G. (1983). *Geotechnical engineering*, Pitman Publishing Inc.

Lee, I. M., and Lee, J. H. (1996). "Prediction of pile bearing capacity using artificial neural networks." *Computers and Geotechnics*, 18(3), 189-200.

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

Maier, H. R. (1995). "Use of artificial neural networks for modelling multivariate water quality time series," PhD. thesis, University of Adelaide, Adelaide.

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, 193-209.

Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications." *Environmental Modelling & Software*, 15(2000), 101-124.

Maier, H. R., Sayed, T., and Lence, B. J. (2001). "Forecasting cyanobacterium Anabaena spp. in River Murray, South Australia, using B-spline neurofuzzy models." *Ecological Modelling*, 146, 85-96.

Mana, I. A. (1977). "Finite element analyses of deep excavation behavior in soft clay," PhD. thesis, Stanford University, Stanford, Calif.

Mansur, C. I., and Kaufman, R. I. (1958). "Pile tests, low sill structure, old river, La." *J. Soil Mech. and Found. Div.*, 82(4), 1-33.

Maren, A., Harston, C., and Pap, R. (1990). *Handbook of neural computing applications*, Academic Press, San Diego, California.

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

Maugeri, M., Castelli, F., Massimino, M. R., and Verona, G. (1998). "Observed and computed settlements of two shallow foundations on sand." *J. Geotech. & Geoenv. Eng.*, 124(7), 595-605.

McClelland, J. L., and Rumelhart, D. E. (1988). *Explorations in parallel distributed processing*, MIT Press, Cambridge, Mass.

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

Melchers, R. E. (1987). *Structural reliability: Analysis and prediction*, Ellis Horwood Limited, Chichester.

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

Meyerhof, G. G. (1976). "Bearing capacity and settlement of pile foundations." *J. Geotech. Eng.*, 102(3), 196-228.

Miller, G. F., Todd, P. M., and Hedge, S. U. (1989). "Designing neural networks using genetic algorithms." *Proc. 3rd Int. Conf. on Genetic Algorithms*, Arlington, 379-384.

Minns, A. W., and Hall, M. J. (1996). "Artificial neural networks as rainfall-runoff models." *Hydrological Sciences Journal*, 41(3), 399-417.

Moon, H. K., Na, S. M., and Lee, C. W. (1995). "Artificial neural-network integrated with expert-system for preliminary design of tunnels and slopes." *Proc. 8th Int. Congress on Rock Mechanics*, Rotterdam: Balkema, 1 & 2, 901-905.

Moorhouse, D. C. (1972). "Shallow foundations." *Proc. Conf. Performance of Earth and Earth-Supported Structures*, Purdue University, West Lafayette, Ind., II, 71-109.

Moselhi, O., Hegazy, T., and Fazio, P. (1992). "Potential applications of neural networks in construction." *Can. J. Civil Eng.*, 19, 521-529.

Mukherjee, A., and Deshpande, J. M. (1997). "Closure on: Modeling initial design process using artificial neural networks." *J. Computing in Civil Eng.*, 11(2), 145-146.

Najjar, Y. M., and Ali, H. E. (1998). "CPT-based liquefaction potential assessment: A neuronet approach." *ASCE Geotechnical Special Publication*, 1, 542-553.

Najjar, Y. M., and Ali, H. E. (1999). "Simulating the stress-strain behaviour of Nevada sand by ANN." *Proc. 5th U.S. National Congress on Computational Mechanics (USACM)*, Boulder, Colorado, August 4-6.

Najjar, Y. M., Ali, H. E., and Basheer, I. A. (1999). "On the use of neuronets for simulating the stress-strain behavior of soils." *Proc. 7th Int. Symposium on Numerical Models in Geomechanics*, Graz, Austria, VII, 657-662.

Najjar, Y. M., and Basheer, I. A. (1996a). "Neural network approach for site characterization and uncertainty prediction." *ASCE Geotechnical Special Publication*, 58(1), 134-148.

Najjar, Y. M., and Basheer, I. A. (1996b). "Utilizing computational neural networks for evaluating the permeability of compacted clay liners." *J. Geotech. and Geological Eng.*, 14, 193-212.

Najjar, Y. M., Basheer, I. A., and Naouss, W. A. (1996). "On the identification of compaction characteristics by neuronets." *J. Computers and Geotechnics*, 18(3), 167-187.

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://geotech.civeng.okstate.edu/ejge/ppr9909/index.html.

NeuralWare. (1997). *NeuralWorks Predict Release 2.1*, NeuralWare Inc., Pittsburgh.

Neusciences. (2000). *Neuframe Version 4.0*, Neusciences Corp., Southampton, Hampshire.

NGI. (1962). "Measurements at a strutted excavation, Oslo subway, Vaterland 3, kn 1450." *Technical Report 8*, Norwegian Geotechnical Institute, Oslo, Norway.

Ni, S. H., Lu, P. C., and Juang, C. H. (1996). "A fuzzy neural network approach to evaluation of slope failure potential." *J. Microcomputers in Civil Eng.*, 11, 59-66.

Orchant, C. J., Kulhawy, F. H., and Trautmann, C. H. (1987). "Critical evaluation of in-situ test methods and their variability." *Research Project 1493-4*, Cornell University, Ithaca, New York.

Padilla, J. D., and Vanmarcke, E. H. (1974). "Settlement of structures on shallow foundations: A probabilistic approach." *Research Report R74-9*, Massachusetts Institute of Technology.

Palisade. (2000). *@Risk: Risk analysis and simulation version 4.0*, Palisade Corp., New York.

Papadopoulos, B. P. (1992). "Settlements of shallow foundations on cohesionless soils." *J. Geotech. Eng.*, 118(3), 377-393.

Parry, R. H. G. (1971). "A direct method of estimating settlements in sand from SPT values." *Proc. Symposium on the Interaction of Structure and Foundations*, Birmingham, 29-37.

Peck, R. B., Hanson, W. E., and Thornburn, T. H. (1974). *Foundation engineering*, John Wiley & Sons, New York.

Peir, J. C., and Cornell, C. A. (1973). "Spatial and temporal variability of live loads." *J. Struct. Div.*, 99(ST5), 903-922.

Penumadu, D., and Jean-Lou, C. (1997). "Geomaterial modeling using artificial neural networks." *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications*, ASCE, 160-184.

Penumadu, D., Jin-Nan, L., Chameau, J.-L., and Arumugam, S. (1994). "Rate dependent behavior of clays using neural networks." *Proc. 13th Conf. Int. Soc. Soil Mech. & Found. Eng.*, New Delhi, 4, 1445-1448.

Penumadu, D., and Zhao, R. (1999). "Triaxial compression behavior of sand and gravel using artificial neural networks (ANN)." *J. Computers and Geotechnics*, 24, 207-230.

Philox, K. T. (1962). "Some recent developments in the design of high buildings in Hong Kong." *The Structural Engineer*, 40(10), 303-323.

Picornell, M., and Del Monte, E. (1988). "Prediction of settlement of cohesive granular soils." *Proc. Measured Performance of Shallow Found.*, Nashville, Tennessee, Geotech. Special Publication No. 15, 55-72.

Picton, P. D. (1994). *Introduction to neural networks*, The MacMillan Press, London.

Poulos, H. G. (1975). "Settlement of isolated foundations." *Proc. General Session of the Symposium Held at The University of New South Wales*, Kensington, NSW, July 14-18, 181-212.

Poulos, H. G. (1999). "Common procedures for foundation settlement analysis – Are they adequate?" *Proc. 8th Australia New Zealand Conf. on Geomechanics*, Hobart, 3-25.

Prekopa, A. (1995). *Stochastic programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Raghavarao, D. (1988). *Exploring statistics*, Marcel Dekker, Inc., New York.

Randolph, M. F. (1985). "Capacity of piles driven into dense sand." *Rep. Soils TR 171*, Engrg. Dept., Cambridge University, Cambridge.

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

Rao, S. S. (1992). *Reliability - based design*, McGraw-Hill, New York.

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

Ray, C., and Klindworth, K. K. (2000). "Neural networks for agrichemical vulnerability assessment of rural private wells." *J. Hydrologic Eng.*, 5(2), 162-171.

Ripley, B. D. (1993). "Statistical aspects of neural networks." *Networks and Chaos: Statistical and Probabilistic Aspects*, O. E. Barndorff-Nielsen, J. L. Jensen, and W. S. Kendall, eds., Chapman & Hall, London.

Ripley, B. D. (1996). *Pattern recognition and neural networks*, Cambridge University Press, Cambridge.

Rizzo, D. M., and Dougherty, D. E. (1994). "Application of artificial neural networks for site characterization using hard and soft information." *Proc. 10th Int. Conf.*

Computational Methods in Water Resources, Dordrecht: Kluwer Academic, 12, 793-799.

Rizzo, D. M., Lillys, T. P., and Dougherty, D. E. (1996). "Comparisons of site characterization methods using mixed data." *ASCE Geotechnical Special Publication*, 58(1), 157-179.

Rogers, L. L., and Dowla, F. U. (1994). "Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling." *Water Resources Research*, 30(2), 457-481.

Rojas, R. (1996). *Neural networks: A systematic introduction*, Springer-Verlag, Berlin.

Romero, S., and Pamukcu, S. (1996). "Characterization of granular material by low strain dynamic excitation and ANN." *ASCE Geotechnical Special Publication*, ASTM-58(2), 1134-1148.

Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo method*, Wiley Press, New York.

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.

Salchenberger, L. M., Cinar, E. M., and Lash, N. A. (1992). "Neural networks: A new tool for predicting thrift failures." *Decision Science*, 23, 899-916.

Sarle, W. S. (1994a). "Neural network implementation in SAS software." *Proc. Nineteenth Annual SAS Users Group Int. Conf.*, Cary, NC: SAS Institute, 1-28.

Sarle, W. S. (1994b). "Neural networks and statistical models." *Proc. 19th Annual SAS Users Group International Conference*, Cary, NC: SAS Institute, 1538-1550.

Sayed, T., and Razavi, A. (2000). "Comparison of neural and conventional approaches to mode choice analysis." *J. Computing in Civil Eng.*, 14(1), 23-30.

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.

Schultze, E., and Menzenbach, K. J. (1961). "Standard penetration test and compressibility of soils." *Proc. 5th Int. Conf. Soil Mech. Found. Eng.*, Paris, 1, 527-532.

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.

Schwarz, G. (1978). "Estimating the dimension of a model." *The Ann. Stat.*, 6(2), 461-464.

Seber, G. A. F. (1974). *Elementary statistics*, John Wiley & Sons, Sydney.

Seed, H. B., Tokimatsu, H., Harder, L. F., and Chung, R. M. (1985). "Influence of SPT procedure in seismic liquefaction resistance evaluations." *J. Geotech. Eng.*, 111(12), 1425-1445.

Semple, R. M., and Rigden, W. J. (1986). "Shaft capacity of driven pipe piles in clay." *Ground Engineering*, 19(1), 11-17.

Shaopei, L., and Boru, D. (1998). "Modeling of fuzzy machine learning and fuzzy neural network in structural design." *Uncertainty Modeling and Analysis in Civil Engineering*, B. M. Ayyub, ed., CRC Press, New York.

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

Shi, J. J. (2000). "Reducing prediction error by transforming input data for neural networks." *J. Computing in Civil Eng.*, 14(2), 109-116.

Shi, J. J. (2002). "Clustering technique for evaluating and validating neural network performance." *J. Comp. Civil Eng.*, 16(2), 152-155.

Shibata, T., and Teparaksa, W. (1988). "Evaluation of liquefaction potentials of soils using cone penetration tests." *Soils and Foundations*, 28(2), 49-60.

Sidarta, D. E., and Ghaboussi, J. (1998). "Constitutive modeling of geomaterials from non-uniform material tests." *J. Computers & Geomechanics*, 22(10), 53-71.

Sietsma, J., and Dow, R. J. F. (1988). "Neural net pruning – Why and how." *Proc. IEEE Int. Conf. Neural Networks*, San Diego, CA, 325-333.

Sivakugan, N., Eckersley, J. D., and Li, H. (1998). "Settlement predictions using neural networks." *Australian Civil Engineering Transactions*, CE40, 49-52.

Sivakugan, N., and Johnson, K. (2002). "Probabilistic design chart for settlements of shallow foundations in granular soils." *Australian Civil Engineering Transactions*, CE43, 19-24.

Smith, G. N. (1986). *Probability and statistics in civil engineering: An introduction*, Collins, London.

Smith, M. (1993). *Neural networks for statistical modelling*, Van Nostrand Reinhold, New York.

Sowers, G. F. (1970). *Introductory soil mechanics and foundations: Geotechnical engineering*, Macmillan Publishing Company, New York.

Standards Australia (1993). "Soil strength and consolidation tests–Determination of the penetration resistance of a soil–Standard penetration test (SPT)." *Methods of Testing Soils for Engineering Purposes*, AS 1289.6.3.1, Sydney.

Sterling, R. L., and Lee, C. A. (1992). "A neural network – Expert system hybird approach for tunnel design." *Proc. 33rd United-States Symp. Rock Mech.*, Brookfield VT: Balkema, 501-510.

Stone, M. (1974). "Cross-validatory choice and assessment of statistical predictions." *J. Royal Statistical Society*, B 36, 111-147.

Sutherland, H. B. (1963). "The use of in-situ tests for estimate the allowable bearing pressure of cohesionless soils." *The Structural Engineer*, 41(3), 85-92.

Tan, C. K., and Duncan, J. M. (1991). "Settlement of footings on sands – Accuracy and reliability." *Proc. Geotechnical Eng. Congress*, Boulde, Colorado, Geotechnical Special publication No. 27, 446-455.

Tang, W. H. (1993). "Recent development in geotechnical reliability." *Probabilistic Methods in Geotech. Eng.*, Li and Lo, eds., Balkema, Rotterdam, 3-27.

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

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

Terzaghi, K., and Peck, R. B. (1948). *Soil mechanics in engineering practice*, John Wiley, New York.

Thorburn, S. (1963). "Tentative correction chart for the standard penetration test in non-cohesive soils." *Civil Engineering & Public Works Review*, June, 752-753.

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

Touretzky, D. S., and Pomerleau, D. A. (1989). "What's hidden in the hidden layers?" *Byte*, August, 227-233.

Tutumluer, E., and Seyhan, U. (1998). "Neural network modeling of anisotropic aggregate behavior from repeated load triaxial tests." *Transportation Research Record 1615*, National Research Council, Washington, DC.

Twomey, J. M., and Smith, A. E. (1997). "Validation and verification." *Artificial neural networks for civil engineers: Fundamentals and applications*, N. Kartam, I. Flood, and J. H. Garrett, eds., ASCE, New York, 44-64.

Ural, D. N., and Saka, H. (1998). "Liquefaction assessment by neural networks." *Elect. J. Geotech. Eng.*, http://geotech.civen.okstate.edu/ejge/ppr9803/index.html.

Vitela, J. E., and Reifman, J. (1997). "Premature saturation in back-propagation networks – Mechanism and necessary conditions." *Neural Networks*, 10(4), 721-735.

Wahls, H. E. (1997). "Settlement analysis for shallow foundations on sand." *Proc. 3rd Int. Geotech. Eng. Conf.*, Cairo, Egypt, 7-28.

Wang, Z. Q., and Lu, W. X. (1982). "On the standardization of SPT and cone penetration test." *Proc. 2nd European Symposium on Penetration Testing*, Amsterdam, 1, 175-182.

Wasserman, P. D. (1989). *Neural computing theory and practice*, Van Nostrand Reinhold, New York.

Weiss, S., and Kulikowski, C. A. (1991). *Computer systems that learn: Classification and prediction methods from statistics, neural nets, machine learning, and expert systems*, Morgan Kaufmann Publishers, San Francisco, CA.

Wellington, A. M. (1892). "The iron wharf at Fort Monroe, Va." *ASCE Transactions*, 27, 129-137.

White, H. (1989). "Learning in artificial neural networks: A statistical perspective." *Neural Computation*, 1, 425-464.

Whitman, R. V. (2000). "Organizing and evaluating uncertainty in geotechnical engineering." *J. Geotech. and Geoenv. Eng.*, 126(7), 583-593.

Wroth, C. P. (1984). "The interpretation of in situ soil tests." *Geotechnique*, 34(4), 449-489.

Yeh, Y. C., Kuo, Y. H., and Hsu, D. S. (1993). "Building KBES for diagnosing PC pile with artificial neural network." *J. Comp. Civil Eng.*, 7(1), 71-93.

Yu, X.-H. (1992). "Can back-propagation error surface not have local minima." *IEEE Transaction on Neural Networks*, 3, 1019-1021.

Zadeh, L. A. (1965). "Fuzzy sets." *Information and Control*, 8, 338-353.

Zhang, M. M. (1997). "Neural network material models determined from structural tests," PhD. thesis, University of Illinois at Urbana-Champain, Urbana, Illinois.

Zhou, Y., and Wu, X. (1994). "Use of neural networks in the analysis and interpretation of site investigation data." *Computer and Geotechnics*, 16, 105-122.

Zhu, J. H., and Zaman, M. M. (1997). "Neural network modeling for a cohesionless soil." *76th Meeting of the Transportation Research Board*, Washington, DC.

Zhu, J. H., Zaman, M. M., and Anderson, S. A. (1998a). "Modeling of soil behavior with a recurrent neural network." *Canadian Geotech. J.*, 35(5), 858-872.

Zhu, J. H., Zaman, M. M., and Anderson, S. A. (1998b). "Modelling of shearing behaviour of a residual soil with recurrent neural network." *Int. J. Numerical and Analytical Methods in Geomechanics*, 22(8), 671-687.

Zolkov, E., and Wiseman, G. (1965). "Engineering properties of dune and beach sands and the little influence of stress history." *Proc. 6th Int. Conf. Soil Mech. & Found. Eng.*, Montreal, 1, 134-138.

Zurada, J. M. (1992). *Introduction to artificial neural systems*, West Publishing Company, St. Paul.

Appendix A. Database Used for ANN Models

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
1	36.6	-	193	28	0	1.5	>12.2	fine to coarse sand	dense	Chemical storage tank	18	Burland and Burbidge (1985)
2	12.2	-	130	17	1.2	7.3	>7.8	fine sand	medium dense	Dextrose storage tank	22	Burland and Burbidge (1985)
3	3.3	14	52	8	1.8	1.6	>25	fine sand	very loose	Bridge with embankment	20	Burland and Burbidge (1985)
4	3.3	14	52	8	1.8	1.6	>25	fine sand	very loose	Bridge with embankment	35	Burland and Burbidge (1985)
5	6	16	162	30	2.8	-1.5	>15	fine sand	dense	RC bridge	10.5	Burland and Burbidge (1985)
6	6	16	162	30	3.6	-2.3	>15	fine sand	dense	RC bridge	11	Burland and Burbidge (1985)
7	5.5	16	93	35	2.85	-1.6	>15.4	fine sand	dense	RC bridge	6.5	Burland and Burbidge (1985)
8	3	14.25	140	38	2.85	-1.6	>15.4	fine sand	dense	RC bridge	3	Burland and Burbidge (1985)
9	4.4	24	93	10	2.5	0.5	>18	silty sand	loose	RC bridge	8	Burland and Burbidge (1985)
10	2.6	22	147	10	2	0	>18	silty sand	loose	RC bridge	12	Burland and Burbidge (1985)
11	2.5	9.5	284	60	3	-	>13.6	fine sand	very dense	RC bridge	1	Burland and Burbidge (1985)
12	2.5	9.5	284	60	3	-	>13.6	fine sand	very dense	RC bridge	3	Burland and Burbidge (1985)
13	5.3	52.5	121	17	2.6	-0.5	>14.8	silty sand	medium dense	RC bridge	12	Burland and Burbidge (1985)
14	19	19	80	15	0	1	-	silty fine sand	loose	Embankment	52	Burland and Burbidge (1985)
15	0.8	-	78	15	0	1	-	silty fine sand	loose	Load test	7	Burland and Burbidge (1985)
16	14.5	64	74	6	1	1.7	>23	sand	very loose to loose	Residential building	74	Burland and Burbidge (1985)
17	14.5	64	74	6	1	1.7	>23	sand	very loose to loose	Residential building	75	Burland and Burbidge (1985)
18	22.4	84	64	6	1	2	>23	sand	very loose to loose	Residential building	70	Burland and Burbidge (1985)
19	22.4	84	75	6	1	1.6	>23	sand	very loose to loose	Residential building	92	Burland and Burbidge (1985)
20	25	25	70	6	1	0.6	>23	sand	very loose to loose	Residential building	121	Burland and Burbidge (1985)
21	25	25	86	6	2.6	0.8	>23	sand	very loose to loose	Residential building	120	Burland and Burbidge (1985)
22	25	25	63	6	1.9	0.8	>23	sand	very loose to loose	Residential building	84	Burbidge (1982)
23	25	25	75	6	2.3	-0.8	>23	sand	very loose to loose	Residential building	87	Burbidge (1982)
24	25	25	76	6	2.1	0.9	>23	sand	very loose to loose	Residential building	85	Burbidge (1982)
25	25	25	75	6	2.8	-0.6	>23	sand	very loose to loose	Residential building	87	Burbidge (1982)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
26	60	-	385	47	5.2	-3.7	60	slightly gravely sand	very dense	Nuclear reactor	40	Burland and Burbidge (1985)
27	3	4.8	231	20	1.5	4	8	fine/medium sand	dense to medium	Steel mill	8.1	Burbidge (1982)
28	3.4	5.4	247	20	1.7	4	8	fine/medium sand	dense to medium	Steel mill	12.2	Burbidge (1982)
29	3.7	5.9	139	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	7.4	Burbidge (1982)
30	3.7	5.9	215	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	15	Burbidge (1982)
31	3.7	5.9	215	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	6.4	Burbidge (1982)
32	3.7	5.9	225	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	7.4	Burbidge (1982)
33	3.7	5.9	252	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	16.5	Burbidge (1982)
34	3.7	5.9	279	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	8.6	Burbidge (1982)
35	3.7	5.9	290	20	1.8	4	8	fine/medium sand	dense to medium	Steel mill	11.2	Burbidge (1982)
36	4	6.4	97	20	2	4	8	fine/medium sand	dense to medium	Steel mill	6.1	Burland and Burbidge (1985)
37	4	6.4	145	20	2	4	8	fine/medium sand	dense to medium	Steel mill	7.4	Burland and Burbidge (1985)
38	4	6.4	225	20	2	4	8	fine/medium sand	dense to medium	Steel mill	9.1	Burland and Burbidge (1985)
39	4.3	6.9	102	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	7.1	Burland and Burbidge (1985)
40	4.3	6.9	134	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	10.2	Burland and Burbidge (1985)
41	4.3	6.9	139	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	7.1	Burland and Burbidge (1985)
42	4.3	6.9	145	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	11	Burland and Burbidge (1985)
43	4.3	6.9	150	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	6.8	Burland and Burbidge (1985)
44	4.3	6.9	161	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	5	Burland and Burbidge (1985)
45	4.3	6.9	177	20	2.1	4	8	fine/medium sand	dense to medium	Steel mill	8.1	Burland and Burbidge (1985)
46	4.6	7.4	113	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	5.1	Burland and Burbidge (1985)
47	4.6	7.4	166	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	8.1	Burland and Burbidge (1985)
48	4.9	7.8	97	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	4.3	Burland and Burbidge (1985)
49	4.9	7.8	102	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	6.9	Burland and Burbidge (1985)
50	4.9	7.8	107	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	3.6	Burland and Burbidge (1985)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
51	4.9	7.8	113	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	8.9	Burland and Burbidge (1985)
52	4.9	7.8	123	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	6.6	Burland and Burbidge (1985)
53	4.9	7.8	182	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	13.8	Burland and Burbidge (1985)
54	4.9	7.8	188	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	15	Burland and Burbidge (1985)
55	4.9	7.8	199	20	2.3	4	8	fine/medium sand	dense to medium	Steel mill	11.7	Burland and Burbidge (1985)
56	5.5	8.8	139	20	2.6	4	8	fine/medium sand	dense to medium	Steel mill	9.4	Burland and Burbidge (1985)
57	6.1	9.8	161	20	3	4	8	fine/medium sand	dense to medium	Steel mill	10.2	Burland and Burbidge (1985)
58	6.4	10.2	150	20	3.2	4	8	fine/medium sand	dense to medium	Steel mill	14.5	Burland and Burbidge (1985)
59	6.7	10.7	113	21	3.4	4	8	fine/medium sand	dense to medium	Steel mill	5	Burland and Burbidge (1985)
60	7	11.2	177	22	3.5	4	8	fine/medium sand	dense to medium	Steel mill	8.3	Burland and Burbidge (1985)
61	42.7	-	166	21	0	5.6	27.4	fine/medium sand	medium dense	Steel storage tank	80	Burland and Burbidge (1985)
62	33.5	-	156	19	0	5.6	27.4	fine/medium sand	loose	Steel storage tank	90	Burland and Burbidge (1985)
63	27.4	-	154	17	0	5.6	27.4	fine/medium sand	loose	Steel storage tank	100	Burland and Burbidge (1985)
64	55	101	233.6	60	9.7	-7.2	21.3	fine sand	very dense	Reactor building	25	Burland and Burbidge (1985)
65	3.5	3.5	25	12	1.5	0	16.2	medium sand	loose	Building	3	Burland and Burbidge (1985)
66	22	75	82	21	5	10	>50	medium sand	medium dense	Coke oven	7.7	Burland and Burbidge (1985)
67	22	-	79	21	5	10	>50	medium sand	medium dense	Chimney	10.5	Burland and Burbidge (1985)
68	2.5	-	245	16	0	10	>50	medium sand	loose	Plate load tests	11	Burland and Burbidge (1985)
69	1	-	247.5	16	0	10	>50	medium sand	loose	Plate load tests	9.9	Burland and Burbidge (1985)
70	16	20.5	70	12	1.5	small	26.7	fine sand	loose	Multi storey building	90	Burland and Burbidge (1985)
71	11	33.4	120	24	5	3.2	12.5	medium sand	medium dense	Building	19.6	Burbidge (1982)
72	5.2	5.2	134	22	5	3.2	12.5	medium sand	medium dense	Building	14.7	Burbidge (1982)
73	4.3	4.3	134	20	5	3.2	12.5	medium sand	medium dense	Building	15.4	Burbidge (1982)
74	4.1	4.1	125	20	5	3.2	12.5	medium sand	medium dense	Building	17.8	Burbidge (1982)
75	3.7	3.7	135	20	5	3.2	12.5	medium sand	medium dense	Building	10.1	Burbidge (1982)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
76	3.4	3.4	129	20	5	3.2	12.5	medium sand	medium dense	Building	11.5	Burbidge (1982)
77	1.5	1.5	150	35	0.6	>3	>3	medium sand	dense	Test footings	2.1	Burland and Burbidge (1985)
78	1.5	1.5	150	50	0.6	>3	>3	medium sand	very dense	Test footings	1	Burland and Burbidge (1985)
79	1.2	1.2	150	28	0.6	>3	>3	medium sand	dense	Test footings	1.3	Burland and Burbidge (1985)
80	1.2	1.2	150	45	0.6	>3	>3	medium sand	very dense	Test footings	0.6	Burland and Burbidge (1985)
81	13	31.5	193	18	2.1	0	19.2	fine to coarse sand	medium dense	Multi storey office block	22	Burland and Burbidge (1985)
82	13	27.4	193	18	2.1	0	19.2	fine to coarse sand	medium dense	Multi storey office block	23.5	Burland and Burbidge (1985)
83	13	22.5	193.8	18	2.1	0	19.2	fine to coarse sand	medium dense	Multi storey office block	18.8	Burland and Burbidge (1985)
84	17.2	43.3	34	17	4.6	0	16.7	sand/gravel	medium dense	Student centre	3.6	Burbidge (1982)
85	1.2	1.2	215	29	2.6	>25	>20	sand with gravel	dense	Plate load tests	2.5	Burland and Burbidge (1985)
86	1.2	1.2	215	26	2.6	>2.5	>20	sand with gravel	dense	Plate load tests	1.5	Burland and Burbidge (1985)
87	1.2	1.2	215	18	2.6	>2.5	>20	sand with gravel	medium dense	Plate load tests	8.6	Burland and Burbidge (1985)
88	34	57	270	30	7.9	-0.9	>27	medium sand	dense	32 Storey building	22	Burland and Burbidge (1985)
89	18.3	-	41	20	0.3	1.8	>8.9	silty fine sand	medium dense	Tank	4.8	Burland and Burbidge (1985)
90	15.2	-	33	20	0.3	1.8	>8.9	silty fine sand	medium dense	Tank	2.8	Burland and Burbidge (1985)
91	4	7	512	37	5	5.6	7	sandy gravel	dense	12 Storey tower block	12.8	Burland and Burbidge (1985)
92	1.2	1.2	300	50	0.5	dry	4.1	sand/gravel	very dense	Plate load tests	4.5	Burland and Burbidge (1985)
93	1.4	1.4	300	50	3.7	1.5	1.5	sand/gravel	very dense	Plate load tests	1.5	Burland and Burbidge (1985)
94	0.9	0.9	300	30	1.2	3.7	6.1	sand/gravel	dense	Plate load tests	4	Burland and Burbidge (1985)
95	0.9	0.9	300	20	3.1	0.9	6.1	sand/gravel	medium dense	Plate load tests	6.7	Burland and Burbidge (1985)
96	0.9	0.9	300	20	1.2	1.8	3.4	sand/gravel	medium dense	Plate load tests	2.7	Burland and Burbidge (1985)
97	4.5	30.5	91	12	2.7	-1.1	7.1	silty sand	loose	Bridge	11	Burland and Burbidge (1985)
98	1.1	1.1	78	13	1.2	1.5	-	sandy gravel	loose	Factory building	2	Burland and Burbidge (1985)
99	1.5	1.5	77	13	1.2	1.5	-	sandy gravel	medium dense	Factory building	2.1	Burland and Burbidge (1985)
100	1.5	1.5	77	13	1.2	1.5	-	sandy gravel	medium dense	Factory building	1.3	Burland and Burbidge (1985)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
101	23.6	26.9	167	35	3	0-3	>25	fine to medium sand	dense	Various structures	15.4	Burland and Burbidge (1985)
102	1.8	1.8	230	25	3	0-3	>25	fine to medium sand	medium dense	Various structures	3.4	Burland and Burbidge (1985)
103	1.4	1.4	230	25	3	0-3	>25	fine to medium sand	medium dense	Various structures	3.9	Burland and Burbidge (1985)
104	1	2.2	284	25	3	0-3	>25	fine to medium sand	medium dense	Various structures	10.5	Burland and Burbidge (1985)
105	4.5	5.7	195	35	3	0-3	>25	fine to medium sand	dense	Various structures	3.9	Burland and Burbidge (1985)
106	15	72.9	81	35	3	0-3	>25	fine to medium sand	dense	Various structures	5.4	Burland and Burbidge (1985)
107	1.6	12.6	250	25	0.4	2.6	>25	fine to medium sand	dense	Various structures	9.3	Burland and Burbidge (1985)
108	1.2	12.7	250	25	0.3	2.7	>25	fine to medium sand	medium dense	Various structures	10	Burland and Burbidge (1985)
109	1	1	294	40	0	3	>25	fine to medium sand	medium dense	Various structures	5	Burland and Burbidge (1985)
110	3.3	5.7	304	40	3	0-3	>25	fine to medium sand	medium dense	Various structures	11.6	Burland and Burbidge (1985)
111	3.6	6.3	304	40	3	0-3	>25	fine to medium sand	medium dense	Various structures	13.3	Burland and Burbidge (1985)
112	4.5	6.8	304	40	3	0-3	>25	fine to medium sand	medium dense	Various structures	18.3	Burland and Burbidge (1985)
113	22.9	32.6	165	30	3	0-3	>25	fine to medium sand	dense	Building	20.4	Burland and Burbidge (1985)
114	21.7	22.2	148	30	3	0-3	>25	fine to medium sand	dense	Building	19.8	Burland and Burbidge (1985)
115	1	1	196	25	3	0-3	>25	fine to medium sand	medium dense	Building	6	Burland and Burbidge (1985)
116	1	1	220	34	0	>2	-	fine sand	dense	Test footings	3.6	Burland and Burbidge (1985)
117	1	1	564	45	0.5	>2	-	fine sand	very dense	Test footings	4.4	Burland and Burbidge (1985)
118	1	1	339	45	0.5	0	-	compacted moist sand	medium dense	Test footings	6	Burland and Burbidge (1985)
119	1	1	284	45	0.5	0	-	compacted moist sand	medium dense	Test footings	4.7	Burland and Burbidge (1985)
120	1.2	1.2	320	25	0	-	-	sand/gravel	dense	Plate load tests	2.8	Burland and Burbidge (1985)
121	12.2	12.2	181	53	3	1	>32	coarse sand with gravel	very dense	Bridge	9.6	Burbidge (1982)
122	22.5	65	221	20	10	-2.5	>30	fine coarse sand	medium dense	25 Storey building	21	Burland and Burbidge (1985)
123	10	-	240	60	1.5	10	12	medium sand	very dense	Storage tank	7	Burland and Burbidge (1985)
124	20	20	85	5	3	-1	32	silty fine sand	very loose	Building	116	Burland and Burbidge (1985)
125	20	20	85	5	3	-1	45	silty fine sand	very loose	Building	81	Burland and Burbidge (1985)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
126	41.2	41.2	104	36	10	-5.5	13	sandy gravel	dense	Multi storey building	10	Burland and Burbidge (1985)
127	0.9	0.9	133	5	0.3	deep	-	fine sand	very loose	Plate load tests	7.6	Burland and Burbidge (1985)
128	0.9	0.9	113	6	0.9	deep	-	fine sand	very loose	Plate load tests	6.4	Burland and Burbidge (1985)
129	1.2	1.2	199	7	0.2	deep	-	fine sand	very loose	Plate load tests	13	Burland and Burbidge (1985)
130	1.2	1.2	268	8	0.9	deep	-	fine sand	very loose	Plate load tests	12.7	Burland and Burbidge (1985)
131	17.6	84	218	20	10.7	-2.2	>37	sand/gravel	medium dense	30 Storey building	26	Burland and Burbidge (1985)
132	16	43	209	14	7.3	-1.8	>23	sand/gravel	medium dense	20 Storey building	18.6	Burland and Burbidge (1985)
133	14.5	14.5	253.5	26	3.5	7.5	21.6	sand/gravel	medium dense	Boiler house	18	Burland and Burbidge (1985)
134	33	-	191	34	5.3	-2.5	8.2	sand/gravel	dense	Reactor building	43.8	Burland and Burbidge (1985)
135	15	26	136	55	6	1.5	-	sand/gravel	very dense	22 Storey building	16.2	Burland and Burbidge (1985)
136	2.6	10.7	293	37	1	-	5.1	sand/gravel	dense	Building	10.9	Burbidge (1982)
137	24.4	-	120	27	0	-	-	sand	dense	Steel oil tank	14.3	Burland and Burbidge (1985)
138	2.1	2.4	584	50	2.4	-	-	sand	very dense	Machine hall	4.4	Burland and Burbidge (1985)
139	2.1	2.1	697	50	1.5	-	-	sand	very dense	Machine hall	2.3	Burland and Burbidge (1985)
140	1.8	2.8	575	50	1.5	-	-	sand	very dense	Machine hall	2.7	Burland and Burbidge (1985)
141	2.1	2.4	584	50	3	-	-	sand	very dense	Machine hall	4.6	Burland and Burbidge (1985)
142	2.1	4.1	347	50	3	-	-	sand	very dense	Machine hall	1.8	Burland and Burbidge (1985)
143	30.2	30.8	386	18	2.7	6.5	22.3	fine coarse sand	medium dense	5 Storey building	91.6	Burland and Burbidge (1985)
144	6	6	190	7	0	0.9	18	fine sand	very loose	Test footings	74	Burland and Burbidge (1985)
145	20	20	145	7	0	0.9	18	fine sand	loose	Embankment	120	Burland and Burbidge (1985)
146	2.8	14	142	4	1	-	10	fine sand	very loose	Bridge	97	Burland and Burbidge (1985)
147	3.3	14.5	99	4	1	-	10	fine sand	very loose	Bridge	37	Burland and Burbidge (1985)
148	13.1	23.9	47.6	25	3	-0.5	-	gravely sand	-	Residential building	3.6	Maugeri et al. (1998)
149	14	22.6	18.32	15	2.5	0	-	gravely sand	-	Residential building	4.2	Maugeri et al. (1998)
150	1.5	1.5	666	18	0.762	4.14	9.94	silty fine sand	medium dense	Test footings	25	Briaud and Gibbens (1999)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
151	2.5	2.5	576	18	0.762	4.14	9.94	silty fine sand	medium dense	Test footings	25	Briaud and Gibbens (1999)
152	3	3	500	18	0.762	4.14	9.94	silty fine sand	medium dense	Test footings	25	Briaud and Gibbens (1999)
153	3	3	500	18	0.889	4	9.8	silty fine sand	medium dense	Test footings	25	Briaud and Gibbens (1999)
154	15	20	148	20	0	1.1	>30	silty gravel/silty sand	loose to medium	Load test	40	Picornell and del Monte (1988)
155	5.2	19.4	153.2	44	0	3.7	-	sand	-	Bridge	8.9	Wahls (1997)
156	5.2	19.4	127.8	58	0	3.7	-	sand	-	Bridge	17	Wahls (1997)
157	4.6	16	111.1	43	0	3.5	-	sand	-	Bridge	23.9	Wahls (1997)
158	5.1	16	116.8	19	1.2	4	-	sand	-	Bridge	19.3	Wahls (1997)
159	3.8	12.5	90	12	1.5	2.7	-	sand	-	Bridge	15.5	Wahls (1997)
160	3.4	22.7	81.4	34	0	9.4	-	sand	-	Bridge	10.7	Wahls (1997)
161	5.6	24.1	112	22	1.5	5.2	-	sand	-	Bridge	15.5	Wahls (1997)
162	6.4	6.4	100.5	18	1.5	3.4	-	sand	-	Bridge	7.1	Wahls (1997)
163	6.4	9.3	71.8	18	1.5	3.4	-	sand	-	Bridge	6.6	Wahls (1997)
164	4.9	8.2	112	20	1.5	5.2	-	sand	-	Bridge	7.4	Wahls (1997)
165	4.9	5.6	118.7	22	1.5	5.2	-	sand	-	Bridge	6.4	Wahls (1997)
166	2.5	13.1	158	21	0	8.5	-	sand	-	Bridge	11.7	Wahls (1997)
167	5.1	23.4	114.9	42	1.8	4.9	-	sand	-	Bridge	5.8	Wahls (1997)
168	4.6	23.2	112	24	2	4.4	-	sand	-	Bridge	11.2	Wahls (1997)
169	4.6	20.5	85.7	39	2.7	3.4	-	sand	-	Bridge	21.1	Wahls (1997)
170	8.5	8.5	102.5	24	0	13.4	-	sand	-	Bridge	16.3	Wahls (1997)
171	6.1	30.7	144.1	23	6.7	6.7	-	sand	-	Bridge	11.7	Wahls (1997)
172	6.1	30.7	155.6	38	1.5	1.8	-	sand	-	Bridge	16.8	Wahls (1997)
173	6.6	13.5	168.1	39	0	5.2	-	sand	-	Bridge	15.5	Wahls (1997)
174	4.9	13.6	161.4	49	0	0.9	-	sand	-	Bridge	7.1	Wahls (1997)
175	5	8.5	181.9	24	2.5	5	-	sand	-	Bridge	11.9	Wahls (1997)

Case No.	Footing width (m)	Footing length (m)	Footing net applied pressure (kPa)	Average SPT blow count	Footing embedment depth (m)	Depth of water table (m)	Thick. of soil layer (m)	Soil type	Soil density	Structure	Measured settlement (mm)	Reference
176	3.3	14.5	98.6	7	2	2	-	sand	-	Bridge	37.1	Wahls (1997)
177	3	10	230.8	50	3	4.5	-	sand	-	Bridge	21.1	Wahls (1997)
178	5.8	24	72.8	17	2.5	2	-	sand	-	Bridge	11.9	Wahls (1997)
179	2.6	21	196.3	9	2	2	-	sand	-	Bridge	33	Wahls (1997)
180	4	7	507.5	32	5	11.6	-	sand	-	Bridge	11.9	Wahls (1997)
181	6	16	158	42	2.8	1.7	-	sand	-	Bridge	7.9	Wahls (1997)
182	6	16	214.5	42	3.6	2.5	-	sand	-	Bridge	4.1	Wahls (1997)
183	7	36	131.2	42	2.3	1.2	-	sand	-	Bridge	11.9	Wahls (1997)
184	5.2	28	95.8	42	2.3	1.2	-	sand	-	Bridge	9.9	Wahls (1997)
185	9	72	115	11	4.5	0	>15	fine/medium clayey sand	medium dense	18 Storey building	25	Bazaraa (1967)
186	2.4	3.9	190	22	4.5	2.5	15.5	clayey sand	medium dense	11 Storey building	8.5	Bazaraa (1967)
187	16.2	25.2	154	16	4.8	0.6	12	fine to medium sand	-	21 Storey building	15	Bazaraa (1967)
188	2.25	2.4	400	8	2.3	2.5	1.84	sand	-	Boiler and shop	43	Bazaraa (1967)
189	25.5	25.5	175	21	2.55	0	8.7	fine grey sand	-	Regenerator	25	Bazaraa (1967)

Appendix B

Database Used for Synthetic Clean and Noisy Data

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
1	0.6	6.8	5.1	1.1	1.3
2	26.0	141.5	50.5	5.4	3.6
3	5.4	156.2	20.4	13.7	12.4
4	21.1	709.8	24.5	56.2	51.4
5	21.1	42.8	42.2	1.9	2.2
6	28.1	639.4	7.0	177.0	152.9
7	10.6	319.1	37.1	16.2	17.0
8	5.0	731.0	45.4	28.6	21.3
9	18.2	295.6	25.5	22.4	37.5
10	3.3	547.7	21.9	42.1	40.1
11	7.1	103.9	11.3	16.8	28.2
12	43.6	699.6	53.2	25.9	23.5
13	39.6	118.1	32.1	7.2	8.1
14	3.3	59.0	9.2	10.7	12.3
15	20.2	604.2	10.9	107.1	106.7
16	38.9	669.0	54.1	24.3	15.7
17	23.1	79.7	45.1	3.4	3.3
18	21.5	442.0	42.0	20.4	23.2
19	32.6	610.7	53.0	22.6	22.3
20	46.6	778.5	5.7	267.9	284.0
21	6.6	657.4	33.4	35.9	41.0
22	35.0	515.8	37.3	27.1	22.4
23	12.5	131.7	9.1	27.3	35.0
24	46.5	290.7	48.3	11.8	11.2
25	1.4	729.2	49.1	20.5	20.4
26	21.0	76.7	47.0	3.1	4.2
27	19.5	549.4	59.1	18.0	19.1
28	5.2	86.0	10.9	14.1	15.9
29	17.6	765.3	36.1	40.9	35.9
30	10.7	362.9	43.1	15.9	13.3
31	49.5	143.0	24.6	11.4	12.3
32	23.6	608.0	7.6	154.5	187.7
33	14.2	106.2	58.7	3.4	3.4
34	21.0	788.7	35.5	43.1	40.7
35	46.3	12.0	10.0	2.3	1.9
36	12.5	134.6	9.5	26.9	42.9
37	17.7	381.3	37.9	19.4	20.0
38	28.2	319.2	17.6	35.3	37.3
39	38.9	50.3	23.5	4.2	3.5
40	39.1	269.1	53.5	9.8	3.1

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
41	32.5	480.0	35.0	26.8	36.3
42	19.1	331.6	29.4	21.8	15.0
43	40.4	797.2	15.2	103.2	116.8
44	17.5	87.7	52.5	3.2	2.4
45	35.8	692.9	5.8	233.0	198.1
46	38.6	639.4	50.4	24.9	13.5
47	13.7	359.1	39.3	17.4	12.7
48	40.5	307.1	57.3	10.5	19.2
49	27.2	642.9	8.5	147.3	191.9
50	20.7	630.0	14.0	87.4	56.7
51	35.2	636.7	53.8	23.2	12.0
52	18.3	206.0	12.9	30.6	49.6
53	29.9	672.2	9.5	137.4	164.2
54	14.7	372.5	40.0	17.8	12.0
55	21.1	124.9	53.6	4.5	3.7
56	36.3	639.9	52.9	23.7	20.9
57	14.1	49.4	50.9	1.8	2.0
58	0.55	131.0	22.3	5.1	6.3
59	38.2	242.6	51.0	9.3	5.4
60	47.3	655.3	42.9	30.1	29.8
61	34.3	73.6	31.8	4.5	5.3
62	2.6	295.9	42.8	11.1	4.3
63	19.1	73.8	48.8	2.9	1.2
64	13.3	247.1	24.2	19.5	17.4
65	7.4	670.4	34.3	36.1	12.0
66	44.0	167.1	34.0	9.6	3.5
67	4.1	116.9	16.3	12.4	5.8
68	4.3	739.2	47.3	27.3	36.8
69	17.4	144.5	5.5	50.5	40.8
70	5.8	782.6	51.6	27.4	32.6
71	25.4	258.3	12.3	40.9	51.3
72	31.9	229.1	56.0	8.0	7.4
73	28.3	201.4	56.2	7.0	5.1
74	23.5	686.3	18.6	71.7	87.2
75	12.6	642.7	24.7	49.6	56.2
76	15.6	219.4	17.8	23.6	31.3
77	41.3	460.7	22.6	40.0	37.0
78	48.6	163.8	28.4	11.3	12.6
79	25.9	196.7	58.2	6.6	6.4
80	7.3	208.1	25.5	15.0	5.3
81	20.8	475.1	47.4	19.4	30.7
82	49.3	453.8	12.8	69.8	12.2
83	5.6	682.5	38.0	32.3	36.2
84	31.8	336.6	16.0	41.2	41.1
85	33.0	631.6	55.5	22.3	27.7
86	23.4	745.0	26.8	54.0	76.8
87	37.4	702.6	5.5	250.5	30.3
88	15.5	790.0	41.9	36.2	50.2
89	21.3	73.9	46.3	3.1	2.9
90	9.9	136.0	12.6	20.2	8.2

Case No.	Input variables			Output variable	
	B (m)	*q* (kPa)	*N*	S_m (mm) (clean)	S_p (mm) (noisy)
91	26.7	253.1	10.1	48.9	38.0
92	1.6	559.9	25.4	31.6	35.0
93	6.2	514.3	14.1	66.2	66.0
94	14.1	212.6	18.5	22.0	21.7
95	27.8	247.8	8.1	59.4	49.3
96	37.0	164.6	41.4	7.8	7.7
97	37.5	556.1	40.1	27.2	23.2
98	34.9	315.7	9.7	63.5	43.2
99	21.5	117.8	52.2	4.3	6.0
100	26.4	601.6	58.7	20.0	23.9
101	41.9	527.4	31.2	33.2	37.5
102	7.1	233.2	29.1	14.7	12.8
103	41.7	170.9	37.1	9.0	11.1
104	28.1	741.8	21.1	68.5	106.7
105	32.1	61.5	32.6	3.6	5.1
106	48.2	5.0	6.9	1.4	1.5
107	38.9	684.9	56.3	23.9	28.3
108	1.6	545.2	23.4	33.2	40.1
109	28.5	473.9	38.6	24.0	26.0
110	29.0	444.4	34.1	25.5	16.0
111	33.2	444.1	29.3	29.7	43.6
112	23.2	384.5	32.2	23.2	30.1
113	12.2	532.1	9.8	102.5	114.4
114	45.1	424.5	13.4	62.1	40.0
115	44.1	208.7	39.7	10.3	9.4
116	31.6	673.8	7.9	167.3	196.3
117	1.3	53.5	10.7	6.6	13.3
118	18.5	97.9	52.8	3.5	6.6
119	3.9	309.8	43.3	12.3	12.6
120	20.2	39.4	42.7	1.7	1.0
121	25.5	721.4	21.2	66.3	38.5
122	4.1	534.4	19.2	48.2	52.6
123	33.3	246.5	56.9	8.5	9.8
124	33.9	206.8	50.8	7.9	9.5
125	42.6	269.8	49.8	10.6	11.5
126	26.1	198.5	58.2	6.6	2.1
127	46.6	606.8	36.9	32.3	7.85
128	46.3	733.5	54.8	26.4	34.1
129	14.4	433.4	48.8	17.0	14.9
130	28.1	217.6	58.7	7.2	7.6
131	7.5	84.0	8.1	19.1	18.4
132	17.4	507.6	55.7	17.6	33.8
133	40.3	166.0	38.0	8.6	6.9
134	49.0	576.5	30.0	37.8	35.6
135	49.7	630.6	36.8	33.7	45.5
136	6.0	273.0	35.8	13.8	10.2
137	22.4	90.9	47.5	3.7	3.9
138	21.2	780.1	34.1	44.4	43.4
139	0.9	5.5	59.5	0.1	0.1
140	46.0	43.6	14.6	5.8	6.9

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
141	46.0	20.6	11.5	3.5	2.7
142	44.9	635.3	42.7	29.3	42.9
143	0.9	658.0	39.9	18.4	25.5
144	6.5	774.5	49.7	28.4	33.3
145	11.4	403.5	48.0	15.9	20.0
146	40.1	382.3	13.1	57.3	85.6
147	17.1	688.1	26.0	51.0	61.2
148	34.4	177.4	46.1	7.5	3.6
149	45.3	667.2	46.7	28.1	27.9
150	26.7	597.4	57.7	20.2	22.4
151	0.9	596.6	31.4	21.3	14.3
152	20.4	792.0	36.7	41.8	38.0
153	31.0	544.2	45.6	23.3	21.3
154	32.7	303.0	10.3	57.2	64.7
155	20.5	310.7	24.9	24.1	20.8
156	20.3	326.1	27.3	23.1	24.2
157	31.9	529.2	42.5	24.3	18.1
158	39.5	66.0	25.0	5.1	8.7
159	35.3	322.7	10.2	61.9	59.0
160	25.4	662.0	13.2	97.8	163.8
161	47.9	401.2	7.0	111.9	101.6
162	4.0	570.1	24.2	40.7	45.8
163	33.6	507.1	37.6	26.4	30.4
164	22.6	771.2	31.3	47.8	47.7
165	37.4	702.7	5.4	252.8	163.5
166	20.0	264.5	19.2	26.6	25.8
167	49.9	119.7	20.9	11.2	12.8
168	25.9	394.5	30.6	25.1	24.9
169	28.9	543.8	47.9	22.2	23.5
170	14.6	63.0	52.3	2.3	2.6
171	41.7	493.7	26.8	36.2	29.9
172	29.1	401.0	27.9	28.0	35.9
173	19.9	600.3	10.8	107.7	101.6
174	33.2	412.0	24.9	32.4	32.3
175	31.9	420.3	27.5	29.9	40.0
176	7.0	446.5	58.7	13.9	14.9
177	29.7	697.3	13.2	103.0	116.5
178	12.4	291.6	31.3	17.7	15.5
179	3.2	41.1	6.9	9.9	8.3
180	15.2	515.3	59.2	16.7	18.0
181	5.2	52.9	6.3	14.8	18.0
182	45.7	161.5	31.3	10.1	10.1
183	48.2	57.7	14.2	7.9	7.5
184	18.2	606.9	13.6	86.3	72.2
185	43.3	703.0	54.0	25.6	40.8
186	41.7	133.1	31.9	8.2	8.4
187	14.8	160.9	10.6	29.0	30.6
188	43.1	380.6	9.5	78.5	66.1
189	37.2	717.1	7.6	183.4	58.0
190	25.9	149.1	51.6	5.6	7.6

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
191	42.0	182.0	38.3	9.3	6.4
192	40.1	173.8	39.2	8.7	9.8
193	1.8	781.8	55.9	20.6	15.7
194	37.2	605.0	47.1	25.2	21.4
195	36.8	711.1	7.2	191.7	104.4
196	45.7	438.9	14.7	58.7	43.0
197	34.8	425.6	25.0	33.4	60.9
198	26.4	656.1	11.2	114.3	148.9
199	32.1	325.6	14.2	44.8	29.1
200	31.9	218.2	54.5	7.8	4.0
201	33.3	307.6	10.3	58.2	94.1
202	46.1	594.5	35.8	32.7	39.0
203	21.1	485.8	48.5	19.4	13.0
204	36.8	447.0	25.8	34.0	28.5
205	1.6	358.3	52.6	9.7	8.6
206	49.7	767.0	55.7	27.2	30.5
207	20.5	132.5	55.3	4.6	5.6
208	8.1	779.6	48.7	29.7	17.3
209	30.8	655.3	6.2	206.9	204.8
210	34.3	241.2	55.0	8.6	10.1
211	45.2	585.4	35.5	32.4	12.7
212	6.1	498.2	11.9	75.6	32.6
213	44.0	387.0	9.4	80.7	72.2
214	43.5	409.8	13.1	61.4	20.5
215	6.3	392.9	52.1	13.7	5.0
216	48.3	149.1	26.8	10.9	5.2
217	20.1	197.4	9.7	39.2	52.8
218	21.6	45.1	42.0	2.0	1.6
219	6.7	729.9	43.3	30.8	36.8
220	43.6	156.4	33.0	9.3	11.7
221	48.9	84.7	17.2	9.7	9.0
222	26.5	508.6	45.6	21.7	16.0
223	5.0	201.4	27.1	13.2	16.1
224	40.7	213.4	44.1	9.5	10.7
225	38.3	531.2	35.7	29.2	38.7
226	49.2	472.9	15.5	60.0	55.4
227	17.4	442.2	46.6	18.3	20.4
228	48.5	470.8	16.0	57.8	56.7
229	10.9	629.2	24.7	48.1	17.0
230	46.7	451.9	15.5	57.5	90.9
231	24.6	575.3	57.1	19.6	3.4
232	0.8	371.7	55.3	7.4	8.3
233	18.8	83.7	50.4	3.2	3.2
234	29.8	469.6	36.6	25.1	31.2
235	11.2	330.6	38.0	16.4	23.4
236	14.2	128.8	6.8	36.1	4.3
237	19.6	446.3	44.8	19.3	26.7
238	21.9	34.4	40.1	1.6	1.5
239	48.0	221.0	37.0	11.7	4.8
240	25.4	693.4	17.5	77.2	60.0

Case No.	Input variables			Output variable	
	B (m)	**q (kPa)**	**N**	**S_m (mm) (clean)**	**S_p (mm) (noisy)**
241	29.2	537.6	46.7	22.5	25.0
242	25.0	444.0	38.5	22.5	22.4
243	2.0	290.9	42.8	10.3	10.2
244	47.1	55.0	15.0	7.2	5.9
245	4.9	491.9	12.3	70.6	70.4
246	8.4	303.9	37.5	15.0	12.8
247	22.1	146.6	55.5	5.1	3.4
248	29.4	8.7	28.3	0.6	0.8
249	19.8	635.5	15.7	78.5	93.8
250	8.2	183.6	21.0	16.2	18.2
251	1.0	192.5	30.3	7.5	6.5
252	20.3	557.2	59.3	18.2	22.4
253	34.2	94.0	34.8	5.3	8.2
254	3.7	136.9	19.6	11.9	16.6
255	49.8	422.2	7.9	105.4	114.6
256	9.5	60.4	57.6	1.9	2.3
257	24.3	419.0	35.8	22.8	27.5
258	6.4	448.9	59.8	13.7	14.9
259	39.2	664.3	53.1	24.6	15.4
260	0.5	559.5	26.6	17.7	26.0
261	19.9	49.0	44.4	2.1	2.7
262	29.4	464.2	36.3	24.9	27.8
263	45.0	220.7	40.3	10.7	6.9
264	40.8	612.9	44.2	27.2	24.8
265	19.8	797.0	38.0	40.6	47.6
266	23.9	292.7	18.7	30.4	61.6
267	48.3	799.4	6.8	232.1	430.3
268	13.4	511.9	5.7	170.2	174.2
269	29.3	251.9	7.1	69.2	39.8
270	46.0	249.9	43.2	11.3	6.6
271	36.6	342.1	11.4	58.6	63.9
272	8.2	622.7	26.8	43.2	50.2
273	41.1	33.2	18.8	3.4	4.1
274	47.4	442.1	13.2	65.6	71.0
275	43.3	730.0	57.7	24.9	8.1
276	37.5	232.3	50.3	9.0	2.2
277	4.8	606.1	28.3	37.9	49.0
278	47.6	121.8	23.7	10.1	8.8
279	33.4	273.2	5.5	97.2	102.9
280	11.8	701.4	33.7	39.5	38.1
281	3.2	668.9	38.8	28.8	55.4
282	32.4	22.3	26.9	1.6	1.3
283	16.0	326.1	32.1	19.5	18.4
284	29.0	425.0	31.3	26.5	35.7
285	36.0	443.3	26.1	33.3	24.7
286	38.7	138.2	35.8	7.5	7.9
287	43.5	115.5	27.5	8.2	8.1
288	11.1	509.9	8.0	120.0	128.2
289	26.1	643.4	9.8	128.1	151.3
290	19.0	137.2	57.6	4.6	3.5

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
291	31.4	378.2	22.2	33.3	48.9
292	27.5	720.7	18.9	74.5	103.0
293	43.3	426.7	15.7	53.4	62.6
294	42.5	436.8	18.0	47.7	60.0
295	1.0	228.6	35.2	7.9	11.9
296	45.0	625.2	41.2	29.8	35.8
297	22.5	415.8	37.3	21.7	10.5
298	32.1	791.9	23.6	65.6	65.2
299	33.9	700.5	9.0	152.7	169.8
300	44.5	599.6	38.3	30.8	20.7
301	12.7	578.7	15.8	69.8	63.4
302	40.1	791.3	14.7	105.8	96.8
303	45.9	176.9	33.3	10.4	11.8
304	37.4	132.9	36.7	7.1	6.1
305	34.4	644.6	55.7	22.7	23.8
306	11.6	696.8	33.3	39.7	29.5
307	6.2	717.4	42.2	30.9	51.8
308	32.5	14.9	25.7	1.1	1.0
309	38.2	794.3	17.2	90.5	151.5
310	41.1	663.5	50.9	25.6	23.3
311	16.7	161.9	8.6	36.2	40.7
312	28.6	283.4	12.2	45.3	52.0
313	31.5	762.7	20.3	73.6	73.3
314	21.4	749.3	29.7	49.0	31.7
315	34.8	150.2	41.9	7.0	6.8
316	34.3	514.5	37.8	26.7	30.3
317	24.6	656.8	13.3	96.2	95.2
318	35.3	548.4	41.5	25.9	27.5
319	40.3	167.4	38.1	8.6	9.8
320	43.4	692.8	52.4	26.0	21.5
321	12.7	36.1	50.6	1.3	1.7
322	44.1	61.4	19.2	6.2	5.9
323	36.1	597.3	47.3	24.8	24.7
324	15.6	498.2	56.3	17.0	22.7
325	39.4	149.7	36.7	8.0	8.5
326	25.4	635.6	9.4	130.9	148.0
327	37.1	795.6	18.6	84.0	73.7
328	38.2	17.5	19.7	1.7	1.4
329	14.9	312.2	31.4	19.0	20.5
330	4.3	435.0	5.2	144.9	176.1
331	32.5	59.7	32.0	3.6	3.6
332	26.3	513.2	46.5	21.5	20.3
333	27.4	466.7	38.9	23.4	19.6
334	41.6	762.6	9.0	165.3	262.9
335	25.7	226.4	7.5	58.3	60.3
336	28.0	305.1	15.9	37.4	39.4
337	28.6	122.8	45.0	5.3	4.4
338	20.8	203.9	9.9	39.8	12.5
339	38.8	285.1	56.1	10.0	13.5
340	45.9	300.5	50.4	11.7	8.1

Case No.	Input variables			Output variable	
	B (m)	*q* (kPa)	*N*	S_m (mm) (clean)	S_p (mm) (noisy)
341	27.1	152.2	50.7	5.8	6.6
342	37.5	70.4	27.9	4.9	3.7
343	36.9	21.7	21.8	1.9	1.6
344	46.4	711.3	51.6	27.2	14.8
345	19.3	15.9	40.6	0.7	0.5
346	46.5	496.9	21.9	44.7	81.5
347	4.7	696.9	41.0	30.0	39.1
348	8.4	313.5	38.8	15.0	9.7
349	46.2	84.9	20.2	8.2	4.2
350	15.9	335.2	33.5	19.2	31.1
351	34.3	693.8	7.6	177.7	212.2
352	38.0	459.9	26.2	34.5	23.2
353	4.1	37.4	5.4	11.9	10.0
354	34.2	74.0	32.1	4.5	4.0
355	10.3	372.4	44.8	15.6	17.6
356	7.0	553.5	18.5	54.8	66.7
357	42.2	611.9	42.5	28.3	16.5
358	36.9	257.9	54.5	9.3	9.2
359	39.3	553.2	37.6	28.9	34.2
360	27.9	709.4	16.9	82.0	32.2
361	35.8	491.8	33.1	29.2	12.5
362	34.0	199.0	49.6	7.8	7.0
363	22.9	120.2	51.0	4.5	1.5
364	34.1	697.0	8.3	163.3	60.1
365	8.9	371.8	46.3	15.0	7.1
366	8.9	585.0	20.8	52.6	70.8
367	14.2	221.3	19.6	21.5	17.4
368	26.8	555.6	51.9	20.9	24.9
369	44.2	417.4	13.4	61.2	76.7
370	25.0	468.0	41.7	21.8	20.4
371	6.1	682.4	37.4	33.1	24.4
372	49.5	719.7	49.4	28.7	34.9
373	18.8	781.6	37.0	40.8	46.3
374	6.3	185.8	23.4	14.4	19.1
375	48.7	86.5	17.6	9.6	8.9
376	15.0	687.3	28.2	46.7	52.2
377	30.8	135.8	44.3	6.0	5.8
378	25.6	701.5	18.3	74.5	26.4
379	35.5	240.5	53.5	8.8	13.9
380	30.3	427.9	30.3	27.6	4.8
381	24.6	289.0	17.4	32.2	36.1
382	23.4	577.9	58.8	19.1	19.0
383	8.6	245.2	29.1	15.7	19.5
384	8.2	383.1	48.6	14.6	20.8
385	21.8	721.4	25.4	55.1	6.6
386	35.9	152.1	40.9	7.3	10.1
387	24.0	766.4	29.1	51.2	48.5
388	2.2	281.4	41.2	10.6	4.3
389	13.3	648.3	24.7	50.0	38.9
390	7.0	613.0	26.8	42.0	46.6

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
391	31.4	542.7	45.0	23.6	23.5
392	4.9	570.5	23.3	43.4	42.9
393	26.7	388.9	28.9	26.2	21.8
394	12.1	389.7	45.3	16.3	16.3
395	47.7	162.6	29.2	10.9	9.3
396	2.1	328.2	47.8	10.5	7.2
397	18.1	575.9	9.3	118.6	163.3
398	19.7	406.0	39.1	20.1	24.0
399	38.6	712.4	5.5	253.0	285.3
400	14.0	226.1	20.5	21.0	18.4
401	4.7	521.7	16.8	54.8	67.5
402	19.4	733.1	29.6	47.9	74.5
403	34.6	401.1	21.8	36.0	50.4
404	22.9	431.0	39.0	21.5	23.3
405	37.3	635.5	51.3	24.3	28.8
406	36.5	194.8	46.2	8.2	10.0
407	45.9	615.0	38.9	31.1	33.9
408	15.4	428.8	47.0	17.5	11.0
409	33.3	668.6	5.3	245.7	361.0
410	20.7	538.8	56.3	18.5	24.1
411	3.5	657.7	36.8	30.3	33.8
412	38.3	196.3	44.3	8.7	5.6
413	8.9	271.7	32.4	15.6	14.2
414	21.5	492.5	49.0	19.5	22.9
415	34.5	577.5	46.3	24.4	49.6
416	49.6	164.3	27.4	11.8	21.9
417	39.3	795.6	16.2	96.2	98.5
418	42.1	396.8	12.8	60.7	34.9
419	12.6	84.8	57.5	2.8	1.6
420	9.5	194.5	21.1	17.2	18.8
421	39.2	509.7	31.8	31.5	36.6
422	34.6	351.7	14.9	46.1	55.1
423	26.2	240.2	8.8	52.8	57.1
424	41.6	576.1	38.2	29.6	9.6
425	29.7	456.3	34.9	25.5	6.2
426	23.0	642.7	13.1	95.1	123.1
427	14.9	684.6	27.9	47.0	41.3
428	21.1	497.0	50.1	19.2	20.3
429	42.9	392.2	11.4	67.8	65.3
430	22.9	100.2	48.2	4.0	7.7
431	16.2	684.9	26.6	49.6	40.2
432	2.2	173.9	26.4	10.2	9.6
433	32.8	447.5	30.3	28.9	39.0
434	43.7	367.6	7.0	102.2	76.0
435	17.1	740.2	33.2	42.9	45.0
436	3.0	506.2	16.5	50.7	49.6
437	33.9	672.7	5.2	252.1	269.4
438	38.7	599.1	44.6	26.4	31.2
439	11.1	300.8	34.1	16.7	12.9
440	12.7	79.1	56.6	2.6	3.9

Case No.	Input variables			Output variable	
	B (m)	q (kPa)	N	S_m (mm) (clean)	S_p (mm) (noisy)
441	44.6	588.7	36.7	31.6	43.7
442	39.6	103.8	30.2	6.7	7.9
443	46.3	272.4	46.1	11.6	14.6
444	27.8	75.5	39.3	3.7	5.6
445	25.2	451.8	39.3	22.4	26.9
446	43.8	611.8	40.7	29.5	14.3
447	24.1	290.8	18.3	30.9	30.7
448	42.5	226.1	43.9	10.1	11.2
449	31.7	781.0	22.6	67.7	45.6
450	40.8	238.2	47.4	9.8	8.9
451	24.4	235.3	10.2	44.6	40.8
452	44.8	333.2	56.1	11.7	13.2
453	34.9	273.2	58.8	9.1	7.8
454	40.7	236.5	47.3	9.8	10.3
455	26.0	736.8	22.9	62.8	46.7
456	45.3	37.1	14.6	5.0	8.4
457	38.8	575.3	41.3	27.4	26.1
458	47.4	717.3	51.3	27.5	46.1
459	44.3	221.5	41.2	10.5	9.6
460	40.7	231.7	46.6	9.7	11.0
461	48.3	550.0	27.2	39.8	45.7
462	37.4	63.7	27.0	4.6	4.6
463	2.1	739.4	49.7	23.0	14.8
464	36.7	387.0	17.5	43.3	42.0
465	5.4	430.3	58.3	13.2	15.0
466	34.2	453.8	29.6	30.1	29.7
467	8.8	241.1	28.4	15.8	16.8
468	34.4	749.7	15.3	95.8	109.5
469	35.7	88.6	32.3	5.3	4.4
470	43.1	37.1	17.0	4.2	5.5
471	15.3	247.4	22.0	21.5	20.3
472	4.6	501.5	14.1	62.4	62.3
473	34.9	684.5	5.7	234.4	313.3
474	35.7	410.5	21.9	36.7	39.1
475	18.3	216.1	14.3	29.1	32.9
476	2.0	50.0	9.4	8.1	7.1
477	26.7	278.8	13.7	39.7	33.4
478	44.8	400.4	10.3	76.0	82.0
479	26.8	395.5	29.7	26.0	31.6
480	48.3	121.6	22.9	10.4	10.4
481	25.3	627.5	8.4	144.3	136.3
482	15.2	26.1	46.5	1.0	0.9
483	22.6	368.0	30.6	23.3	37.1
484	32.7	266.1	5.2	98.8	102.2
485	40.8	637.0	47.6	26.3	27.7
486	15.7	756.8	37.0	39.3	33.1
487	21.8	763.2	31.1	47.6	15.0
488	5.0	574.8	23.7	43.1	58.3
489	49.6	56.0	12.4	8.8	6.1
490	11.8	497.3	5.5	171.7	194.3

Case No.	Input variables			Output variable	
	B (m)	**q (kPa)**	**N**	**S_m (mm) (clean)**	**S_p (mm) (noisy)**
491	22.1	115.8	51.3	4.3	3.3
492	18.6	473.4	49.5	18.4	15.7
493	25.3	14.3	33.6	0.8	0.4
494	3.2	419.3	59.3	11.8	8.6
495	40.7	706.0	57.2	24.2	44.2
496	31.5	55.0	32.4	3.3	4.3
497	48.6	134.1	24.4	10.8	7.0
498	21.3	608.4	10.2	115.0	59.4
499	29.4	252.7	7.1	69.3	112.1
500	20.8	494.0	50.0	19.1	22.9

Training = 1 to 300, testing = 301 to 400 and validation = 401 to 500

Appendix C

Membership Values of Fuzzy Clustering

NUMBER OF CLUSTERS 16
****

FUZZY CLUSTERING
****

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
001	0.1208	0.0387	0.0594	0.1207	0.0364	0.0825	0.0526	0.0655	0.0438	0.0511	0.0651	0.0525	0.0713	0.0514	0.0441	0.0441
002	0.0111	0.3835	0.0046	0.0110	0.0301	0.0146	0.0419	0.0240	0.1028	0.0457	0.0244	0.0423	0.0177	0.0457	0.1004	0.1003
003	0.0669	0.0423	0.2320	0.0668	0.0406	0.0586	0.0489	0.0536	0.0450	0.0483	0.0536	0.0489	0.0555	0.0485	0.0452	0.0452
004	0.0266	0.1440	0.0128	0.0266	0.0752	0.0328	0.0643	0.0463	0.0957	0.0672	0.0469	0.0647	0.0376	0.0674	0.0959	0.0959
005	0.2135	0.0228	0.0154	0.2149	0.0192	0.0993	0.0375	0.0564	0.0276	0.0358	0.0557	0.0374	0.0729	0.0360	0.0278	0.0279
006	0.0328	0.0591	0.0125	0.0328	0.0843	0.0444	0.0801	0.0691	0.0731	0.0802	0.0699	0.0804	0.0519	0.0807	0.0742	0.0744
007	0.0689	0.0176	0.0085	0.0684	0.0145	0.3700	0.0338	0.0658	0.0224	0.0317	0.0641	0.0337	0.1234	0.0319	0.0227	0.0227
008	0.2882	0.0170	0.0128	0.2850	0.0147	0.0810	0.0271	0.0407	0.0203	0.0259	0.0402	0.0271	0.0529	0.0261	0.0205	0.0205
009	0.0287	0.0699	0.0134	0.0286	0.2262	0.0354	0.0610	0.0491	0.0702	0.0623	0.0497	0.0613	0.0397	0.0626	0.0709	0.0710
010	0.0177	0.0381	0.0056	0.0176	0.0206	0.0269	0.1320	0.0631	0.0648	0.1260	0.0647	0.1297	0.0378	0.1242	0.0656	0.0656
011	0.0279	0.0247	0.0066	0.0278	0.0174	0.0562	0.0737	0.1721	0.0356	0.0653	0.1682	0.0727	0.1137	0.0656	0.0362	0.0362
012	0.0127	0.1461	0.0050	0.0127	0.0296	0.0170	0.0581	0.0297	0.1541	0.0643	0.0303	0.0587	0.0211	0.0641	0.1485	0.1479
013	0.0156	0.0287	0.0047	0.0155	0.0169	0.0258	0.1464	0.0787	0.0494	0.1295	0.0811	0.1414	0.0392	0.1266	0.0502	0.0502
014	0.0113	0.2044	0.0045	0.0113	0.0301	0.0152	0.0502	0.0264	0.1508	0.0559	0.0269	0.0508	0.0188	0.0557	0.1442	0.1436
015	0.0148	0.1943	0.0061	0.0148	0.0377	0.0195	0.0574	0.0324	0.1295	0.0624	0.0329	0.0579	0.0237	0.0624	0.1273	0.1271
016	0.0212	0.0222	0.0054	0.0211	0.0160	0.0423	0.0758	0.2054	0.0332	0.0662	0.2020	0.0746	0.0807	0.0664	0.0338	0.0338
017	0.0113	0.0304	0.0037	0.0113	0.0156	0.0176	0.1497	0.0446	0.0612	0.1618	0.0459	0.1459	0.0248	0.1525	0.0619	0.0619
018	0.0140	0.0277	0.0043	0.0140	0.0165	0.0230	0.1544	0.0692	0.0493	0.1383	0.0715	0.1489	0.0343	0.1343	0.0502	0.0502
019	0.0117	0.3780	0.0049	0.0117	0.0303	0.0154	0.0432	0.0251	0.1008	0.0470	0.0255	0.0436	0.0187	0.0470	0.0986	0.0984

020	0.0406	0.0420	0.0121	0.0405	0.0424	0.0628	0.0805	0.0997	0.0557	0.0770	0.0999	0.0805	0.0754	0.0775	0.0566	0.0567
021	0.2680	0.0172	0.0115	0.2701	0.0146	0.0935	0.0287	0.0448	0.0209	0.0274	0.0442	0.0287	0.0607	0.0275	0.0211	0.0211
022	0.3507	0.0115	0.0091	0.3677	0.0101	0.0498	0.0183	0.0270	0.0138	0.0175	0.0267	0.0183	0.0343	0.0176	0.0139	0.0139
023	0.0119	0.0462	0.0042	0.0119	0.0234	0.0176	0.1079	0.0390	0.1058	0.1256	0.0401	0.1083	0.0234	0.1224	0.1062	0.1061
024	0.0187	0.0580	0.0066	0.0187	0.0307	0.0263	0.1025	0.0516	0.0944	0.1066	0.0527	0.1026	0.0341	0.1062	0.0952	0.0952
025	0.0115	0.2045	0.0046	0.0114	0.0327	0.0155	0.0506	0.0267	0.1483	0.0562	0.0272	0.0511	0.0190	0.0561	0.1427	0.1422
026	0.0096	0.0666	0.0036	0.0095	0.0191	0.0135	0.0637	0.0262	0.1890	0.0745	0.0268	0.0643	0.0173	0.0734	0.1725	0.1707
027	0.0358	0.0552	0.0133	0.0357	0.0756	0.0483	0.0805	0.0743	0.0690	0.0797	0.0751	0.0808	0.0561	0.0803	0.0701	0.0702
028	0.0060	0.0201	0.0027	0.0060	0.8019	0.0077	0.0156	0.0115	0.0198	0.0162	0.0116	0.0157	0.0089	0.0162	0.0201	0.0201
029	0.0292	0.0252	0.0070	0.0291	0.0188	0.0639	0.0674	0.1731	0.0355	0.0611	0.1669	0.0668	0.1224	0.0614	0.0361	0.0361
030	0.0093	0.1052	0.0036	0.0093	0.0213	0.0128	0.0506	0.0235	0.2003	0.0577	0.0240	0.0511	0.0162	0.0573	0.1799	0.1779
031	0.0147	0.0348	0.0047	0.0146	0.0212	0.0227	0.1380	0.0586	0.0641	0.1348	0.0604	0.1360	0.0318	0.1328	0.0653	0.0654
032	0.0141	0.1200	0.0056	0.0141	0.0324	0.0190	0.0651	0.0333	0.1459	0.0715	0.0340	0.0657	0.0236	0.0714	0.1425	0.1421
033	0.0397	0.0207	0.0074	0.0395	0.0158	0.1081	0.0479	0.1187	0.0278	0.0439	0.1138	0.0476	0.2686	0.0442	0.0282	0.0282
034	0.0505	0.0154	0.0070	0.0501	0.0126	0.4375	0.0302	0.0620	0.0197	0.0282	0.0602	0.0301	0.1283	0.0284	0.0199	0.0199
035	0.0100	0.0570	0.0036	0.0100	0.0190	0.0143	0.0767	0.0289	0.1650	0.0913	0.0296	0.0773	0.0186	0.0893	0.1553	0.1541
036	0.0756	0.0356	0.2718	0.0755	0.0335	0.0595	0.0441	0.0511	0.0389	0.0433	0.0509	0.0441	0.0542	0.0435	0.0391	0.0391
037	0.0120	0.3686	0.0050	0.0119	0.0307	0.0158	0.0440	0.0256	0.1019	0.0478	0.0261	0.0444	0.0190	0.0478	0.0997	0.0996
038	0.0322	0.0306	0.0082	0.0321	0.0236	0.0586	0.0812	0.1420	0.0434	0.0737	0.1412	0.0805	0.0900	0.0742	0.0442	0.0443
039	0.0178	0.0714	0.0078	0.0177	0.3555	0.0229	0.0498	0.0349	0.0681	0.0520	0.0354	0.0502	0.0266	0.0522	0.0689	0.0690
040	0.1102	0.0404	0.0602	0.1100	0.0395	0.0823	0.0543	0.0669	0.0456	0.0529	0.0666	0.0543	0.0718	0.0532	0.0460	0.0460
041	0.0122	0.3725	0.0051	0.0121	0.0342	0.0159	0.0435	0.0257	0.0995	0.0472	0.0261	0.0439	0.0191	0.0472	0.0980	0.0978
042	0.0106	0.0487	0.0038	0.0106	0.0209	0.0155	0.0953	0.0331	0.1296	0.1151	0.0339	0.0959	0.0204	0.1119	0.1276	0.1271
043	0.0140	0.0393	0.0046	0.0139	0.0191	0.0208	0.1322	0.0479	0.0756	0.1379	0.0492	0.1305	0.0287	0.1342	0.0762	0.0762
044	0.0337	0.0162	0.0060	0.0335	0.0127	0.1250	0.0364	0.0961	0.0216	0.0335	0.0916	0.0362	0.3800	0.0337	0.0219	0.0219
045	0.0117	0.0279	0.0037	0.0117	0.0159	0.0186	0.1576	0.0513	0.0540	0.1563	0.0529	0.1527	0.0267	0.1491	0.0549	0.0550
046	0.0109	0.0342	0.0037	0.0109	0.0163	0.0166	0.1362	0.0396	0.0742	0.1600	0.0407	0.1343	0.0229	0.1504	0.0746	0.0745
047	0.0145	0.0312	0.0045	0.0144	0.0190	0.0232	0.1440	0.0656	0.0564	0.1369	0.0676	0.1407	0.0336	0.1337	0.0574	0.0575
048	0.0105	0.0521	0.0038	0.0105	0.0191	0.0152	0.0886	0.0316	0.1428	0.1066	0.0324	0.0891	0.0200	0.1036	0.1374	0.1366
049	0.3097	0.0153	0.0117	0.3188	0.0132	0.0640	0.0244	0.0358	0.0183	0.0233	0.0354	0.0243	0.0455	0.0234	0.0185	0.0185
050	0.0097	0.0857	0.0037	0.0097	0.0204	0.0135	0.0567	0.0253	0.1951	0.0651	0.0258	0.0572	0.0172	0.0644	0.1761	0.1742
051	0.0434	0.0238	0.0085	0.0432	0.0194	0.1141	0.0541	0.1360	0.0318	0.0497	0.1308	0.0538	0.1766	0.0501	0.0323	0.0324
052	0.0500	0.0271	0.4752	0.0500	0.0259	0.0416	0.0326	0.0369	0.0293	0.0321	0.0368	0.0326	0.0386	0.0322	0.0294	0.0295
053	0.0107	0.0371	0.0036	0.0106	0.0171	0.0160	0.1258	0.0369	0.0864	0.1534	0.0379	0.1249	0.0217	0.1449	0.0865	0.0864
054	0.0197	0.0845	0.0088	0.0197	0.2876	0.0251	0.0538	0.0379	0.0756	0.0562	0.0384	0.0542	0.0292	0.0565	0.0764	0.0765
055	0.0202	0.0240	0.0053	0.0201	0.0167	0.0384	0.0905	0.1789	0.0368	0.0782	0.1800	0.0888	0.0687	0.0783	0.0375	0.0375
056	0.0107	0.0380	0.0037	0.0106	0.0182	0.0160	0.1225	0.0368	0.0903	0.1490	0.0378	0.1220	0.0216	0.1417	0.0906	0.0905

057	0.0827	0.0379	0.2146	0.0826	0.0359	0.0647	0.0474	0.0551	0.0415	0.0465	0.0550	0.0474	0.0586	0.0466	0.0418	0.0418
058	0.0137	0.0708	0.0051	0.0137	0.0254	0.0192	0.0841	0.0369	0.1346	0.0936	0.0377	0.0846	0.0246	0.0926	0.1320	0.1316
059	0.0103	0.1451	0.0041	0.0103	0.0239	0.0141	0.0508	0.0250	0.1792	0.0572	0.0255	0.0514	0.0175	0.0569	0.1651	0.1637
060	0.0122	0.0290	0.0039	0.0122	0.0156	0.0193	0.1554	0.0518	0.0552	0.1546	0.0534	0.1504	0.0278	0.1472	0.0560	0.0560
061	0.3437	0.0118	0.0086	0.3528	0.0101	0.0582	0.0192	0.0292	0.0142	0.0184	0.0288	0.0192	0.0384	0.0185	0.0144	0.0144
062	0.0142	0.2249	0.0059	0.0141	0.0413	0.0187	0.0542	0.0310	0.1243	0.0588	0.0315	0.0547	0.0227	0.0589	0.1225	0.1224
063	0.0105	0.0373	0.0036	0.0105	0.0177	0.0157	0.1239	0.0362	0.0893	0.1511	0.0373	0.1234	0.0212	0.1434	0.0896	0.0895
064	0.0060	0.0201	0.0027	0.0060	0.8019	0.0077	0.0156	0.0115	0.0198	0.0162	0.0116	0.0157	0.0089	0.0162	0.0201	0.0201
065	0.0345	0.0536	0.0124	0.0344	0.0686	0.0475	0.0830	0.0764	0.0687	0.0818	0.0772	0.0832	0.0561	0.0824	0.0699	0.0700
066	0.0315	0.0190	0.0063	0.0313	0.0145	0.0901	0.0459	0.1337	0.0258	0.0418	0.1266	0.0455	0.2933	0.0421	0.0262	0.0262
067	0.0190	0.0289	0.0054	0.0189	0.0190	0.0317	0.1300	0.1014	0.0472	0.1111	0.1042	0.1267	0.0493	0.1108	0.0481	0.0482
068	0.0109	0.0294	0.0036	0.0109	0.0155	0.0170	0.1516	0.0437	0.0600	0.1646	0.0450	0.1476	0.0240	0.1547	0.0608	0.0608
069	0.0164	0.0336	0.0051	0.0164	0.0193	0.0254	0.1415	0.0650	0.0583	0.1302	0.0669	0.1384	0.0365	0.1285	0.0593	0.0593
070	0.0205	0.0346	0.0061	0.0204	0.0211	0.0322	0.1277	0.0822	0.0558	0.1148	0.0842	0.1253	0.0472	0.1144	0.0567	0.0568
071	0.0127	0.0400	0.0057	0.0126	0.5962	0.0161	0.0319	0.0238	0.0398	0.0330	0.0241	0.0321	0.0185	0.0331	0.0403	0.0403
072	0.0202	0.0239	0.0053	0.0201	0.0164	0.0382	0.0919	0.1771	0.0367	0.0789	0.1784	0.0901	0.0689	0.0791	0.0374	0.0374
073	0.0106	0.0437	0.0037	0.0106	0.0186	0.0156	0.1074	0.0343	0.1122	0.1318	0.0352	0.1075	0.0209	0.1263	0.1109	0.1106
074	0.0106	0.0550	0.0038	0.0106	0.0228	0.0152	0.0838	0.0314	0.1471	0.0989	0.0322	0.0845	0.0198	0.0971	0.1440	0.1434
075	0.0277	0.0670	0.0125	0.0276	0.2315	0.0347	0.0616	0.0493	0.0697	0.0629	0.0499	0.0619	0.0393	0.0632	0.0705	0.0706
076	0.0142	0.0362	0.0046	0.0142	0.0200	0.0223	0.1339	0.0553	0.0678	0.1395	0.0568	0.1319	0.0312	0.1349	0.0686	0.0686
077	0.0781	0.0363	0.2527	0.0780	0.0342	0.0612	0.0452	0.0524	0.0397	0.0444	0.0523	0.0452	0.0557	0.0445	0.0400	0.0400
078	0.3199	0.0144	0.0113	0.3299	0.0126	0.0615	0.0228	0.0336	0.0171	0.0218	0.0333	0.0228	0.0424	0.0220	0.0173	0.0173
079	0.1249	0.0231	0.0131	0.1236	0.0196	0.2121	0.0404	0.0690	0.0285	0.0384	0.0678	0.0404	0.1028	0.0386	0.0288	0.0289
080	0.0316	0.0190	0.0063	0.0315	0.0148	0.0912	0.0466	0.1437	0.0259	0.0424	0.1357	0.0462	0.2696	0.0427	0.0264	0.0264
081	0.0206	0.1982	0.0093	0.0206	0.0587	0.0261	0.0596	0.0393	0.1038	0.0631	0.0398	0.0600	0.0306	0.0632	0.1036	0.1036
082	0.0351	0.0346	0.0094	0.0350	0.0299	0.0591	0.0838	0.1230	0.0484	0.0773	0.1232	0.0834	0.0809	0.0780	0.0493	0.0494
083	0.0097	0.0970	0.0037	0.0097	0.0245	0.0134	0.0542	0.0248	0.1902	0.0617	0.0253	0.0548	0.0168	0.0614	0.1771	0.1757
084	0.0146	0.0330	0.0046	0.0145	0.0175	0.0226	0.1461	0.0571	0.0597	0.1398	0.0587	0.1424	0.0323	0.1361	0.0604	0.0605
085	0.0542	0.0155	0.0072	0.0538	0.0127	0.4438	0.0300	0.0598	0.0197	0.0281	0.0582	0.0299	0.1188	0.0283	0.0200	0.0200
086	0.0091	0.0666	0.0034	0.0091	0.0188	0.0128	0.0600	0.0247	0.1984	0.0702	0.0253	0.0606	0.0164	0.0692	0.1788	0.1766
087	0.0307	0.0184	0.0061	0.0305	0.0140	0.0876	0.0446	0.1314	0.0250	0.0406	0.1244	0.0443	0.3107	0.0409	0.0254	0.0254
088	0.0090	0.0638	0.0034	0.0090	0.0188	0.0127	0.0613	0.0248	0.1965	0.0722	0.0254	0.0619	0.0164	0.0710	0.1780	0.1759
089	0.0325	0.0228	0.0071	0.0323	0.0179	0.0791	0.0576	0.1717	0.0315	0.0523	0.1635	0.0572	0.1577	0.0527	0.0320	0.0320
090	0.0139	0.2155	0.0057	0.0138	0.0383	0.0184	0.0546	0.0307	0.1288	0.0595	0.0313	0.0551	0.0224	0.0595	0.1264	0.1262
091	0.0337	0.0200	0.0067	0.0335	0.0149	0.0903	0.0479	0.1285	0.0271	0.0437	0.1227	0.0476	0.2845	0.0440	0.0275	0.0275
092	0.0819	0.0283	0.0128	0.0817	0.0228	0.1313	0.0548	0.0955	0.0361	0.0513	0.0939	0.0546	0.1304	0.0517	0.0365	0.0366
093	0.0422	0.0178	0.0070	0.0419	0.0139	0.1631	0.0383	0.0908	0.0234	0.0355	0.0874	0.0381	0.3173	0.0357	0.0237	0.0238

094	0.0106	0.0379	0.0036	0.0106	0.0173	0.0158	0.1232	0.0363	0.0901	0.1509	0.0373	0.1224	0.0215	0.1426	0.0901	0.0899
095	0.0093	0.1052	0.0036	0.0093	0.0212	0.0128	0.0506	0.0235	0.2007	0.0576	0.0240	0.0511	0.0161	0.0572	0.1799	0.1779
096	0.0207	0.1637	0.0093	0.0207	0.0906	0.0264	0.0604	0.0401	0.1023	0.0640	0.0406	0.0609	0.0309	0.0641	0.1026	0.1027
097	0.0412	0.0242	0.0082	0.0411	0.0182	0.0940	0.0578	0.1341	0.0328	0.0527	0.1299	0.0574	0.1890	0.0531	0.0333	0.0333
098	0.0127	0.0069	0.8666	0.0126	0.0066	0.0105	0.0083	0.0094	0.0075	0.0082	0.0093	0.0083	0.0098	0.0082	0.0075	0.0075
099	0.3018	0.0161	0.0123	0.3092	0.0138	0.0666	0.0255	0.0374	0.0192	0.0245	0.0370	0.0255	0.0476	0.0246	0.0194	0.0194
100	0.0115	0.1871	0.0046	0.0115	0.0292	0.0156	0.0520	0.0271	0.1541	0.0578	0.0276	0.0525	0.0193	0.0576	0.1466	0.1458
101	0.0651	0.0165	0.0079	0.0646	0.0135	0.4052	0.0316	0.0619	0.0209	0.0296	0.0604	0.0315	0.1190	0.0299	0.0212	0.0212
102	0.0247	0.0207	0.0057	0.0246	0.0153	0.0554	0.0603	0.2098	0.0296	0.0537	0.1985	0.0596	0.1277	0.0540	0.0301	0.0302
103	0.0200	0.0312	0.0058	0.0199	0.0209	0.0330	0.1240	0.0999	0.0504	0.1091	0.1025	0.1217	0.0498	0.1090	0.0514	0.0515
104	0.0649	0.0399	0.2670	0.0649	0.0384	0.0564	0.0465	0.0513	0.0425	0.0459	0.0512	0.0465	0.0532	0.0460	0.0427	0.0427
105	0.0196	0.2008	0.0087	0.0195	0.0686	0.0250	0.0581	0.0380	0.1038	0.0618	0.0385	0.0586	0.0294	0.0619	0.1038	0.1038
106	0.2153	0.0226	0.0152	0.2168	0.0191	0.0992	0.0372	0.0560	0.0273	0.0355	0.0553	0.0371	0.0725	0.0357	0.0276	0.0276
107	0.0189	0.1127	0.0080	0.0188	0.1287	0.0247	0.0638	0.0402	0.1039	0.0679	0.0408	0.0644	0.0295	0.0681	0.1048	0.1049
108	0.0188	0.0310	0.0056	0.0188	0.0212	0.0311	0.1275	0.0959	0.0509	0.1127	0.0985	0.1251	0.0465	0.1126	0.0519	0.0520
109	0.0106	0.0332	0.0035	0.0106	0.0161	0.0161	0.1377	0.0388	0.0731	0.1627	0.0399	0.1356	0.0222	0.1526	0.0737	0.0736
110	0.0412	0.0173	0.0068	0.0410	0.0137	0.1855	0.0370	0.0893	0.0227	0.0343	0.0859	0.0369	0.3079	0.0345	0.0230	0.0230
111	0.2620	0.0177	0.0123	0.2585	0.0153	0.1031	0.0292	0.0456	0.0214	0.0279	0.0450	0.0291	0.0615	0.0280	0.0217	0.0217
112	0.0480	0.0135	0.0063	0.0477	0.0111	0.5028	0.0264	0.0538	0.0173	0.0247	0.0524	0.0263	0.1097	0.0249	0.0175	0.0175
113	0.0286	0.0326	0.0077	0.0285	0.0225	0.0479	0.0983	0.1198	0.0480	0.0879	0.1209	0.0971	0.0741	0.0884	0.0488	0.0489
114	0.0196	0.0257	0.0053	0.0195	0.0171	0.0357	0.1064	0.1490	0.0404	0.0911	0.1516	0.1040	0.0614	0.0910	0.0411	0.0412
115	0.0092	0.0652	0.0034	0.0091	0.0206	0.0129	0.0618	0.0252	0.1913	0.0724	0.0257	0.0624	0.0165	0.0714	0.1773	0.1756
116	0.0236	0.0226	0.0058	0.0235	0.0172	0.0485	0.0694	0.2065	0.0330	0.0616	0.2002	0.0686	0.0901	0.0620	0.0336	0.0337
117	0.0155	0.1086	0.0061	0.0155	0.0565	0.0212	0.0679	0.0373	0.1304	0.0743	0.0379	0.0686	0.0260	0.0742	0.1300	0.1300
118	0.2560	0.0213	0.0192	0.2560	0.0186	0.0777	0.0324	0.0459	0.0251	0.0312	0.0455	0.0324	0.0565	0.0314	0.0253	0.0253
119	0.0175	0.0618	0.0079	0.0174	0.4281	0.0220	0.0441	0.0324	0.0577	0.0457	0.0328	0.0444	0.0253	0.0459	0.0584	0.0584
120	0.0233	0.0219	0.0057	0.0232	0.0165	0.0485	0.0673	0.2123	0.0319	0.0596	0.2046	0.0665	0.0936	0.0600	0.0325	0.0325
121	0.0091	0.0653	0.0034	0.0091	0.0194	0.0128	0.0615	0.0250	0.1941	0.0719	0.0255	0.0621	0.0165	0.0709	0.1777	0.1758
122	0.0151	0.0301	0.0047	0.0151	0.0184	0.0240	0.1490	0.0671	0.0533	0.1337	0.0691	0.1450	0.0349	0.1318	0.0543	0.0544
123	0.0149	0.1522	0.0061	0.0149	0.0425	0.0198	0.0611	0.0335	0.1352	0.0665	0.0341	0.0617	0.0241	0.0665	0.1336	0.1334
124	0.0280	0.0148	0.7107	0.0280	0.0140	0.0230	0.0179	0.0203	0.0160	0.0176	0.0203	0.0179	0.0213	0.0177	0.0161	0.0161
125	0.0122	0.3556	0.0051	0.0121	0.0365	0.0159	0.0442	0.0259	0.1032	0.0480	0.0263	0.0447	0.0192	0.0480	0.1016	0.1015
126	0.0096	0.0670	0.0036	0.0096	0.0192	0.0135	0.0636	0.0262	0.1894	0.0743	0.0268	0.0641	0.0174	0.0730	0.1724	0.1705
127	0.0647	0.0161	0.0078	0.0642	0.0134	0.4272	0.0306	0.0597	0.0204	0.0288	0.0583	0.0305	0.1080	0.0290	0.0206	0.0207
128	0.0107	0.0293	0.0035	0.0107	0.0154	0.0166	0.1520	0.0424	0.0607	0.1651	0.0437	0.1481	0.0233	0.1555	0.0615	0.0615
129	0.0109	0.0342	0.0037	0.0109	0.0163	0.0166	0.1365	0.0396	0.0741	0.1600	0.0407	0.1344	0.0229	0.1502	0.0746	0.0745
130	0.0138	0.0469	0.0048	0.0137	0.0222	0.0200	0.1153	0.0432	0.0949	0.1253	0.0443	0.1150	0.0267	0.1230	0.0954	0.0954

131	0.0169	0.0315	0.0051	0.0169	0.0186	0.0268	0.1426	0.0729	0.0537	0.1268	0.0749	0.1390	0.0394	0.1255	0.0546	0.0547
132	0.2032	0.0199	0.0125	0.2006	0.0170	0.1510	0.0339	0.0555	0.0243	0.0323	0.0547	0.0339	0.0796	0.0325	0.0246	0.0247
133	0.0170	0.0568	0.0076	0.0170	0.4480	0.0215	0.0432	0.0319	0.0551	0.0447	0.0322	0.0435	0.0248	0.0449	0.0558	0.0559
134	0.0150	0.0267	0.0045	0.0149	0.0163	0.0248	0.1531	0.0801	0.0461	0.1292	0.0826	0.1473	0.0380	0.1272	0.0470	0.0471
135	0.0125	0.0816	0.0047	0.0125	0.0257	0.0177	0.0725	0.0335	0.1532	0.0824	0.0342	0.0730	0.0226	0.0813	0.1467	0.1459
136	0.0139	0.2067	0.0057	0.0139	0.0368	0.0184	0.0552	0.0307	0.1316	0.0601	0.0312	0.0557	0.0224	0.0601	0.1290	0.1288
137	0.0101	0.1441	0.0040	0.0101	0.0237	0.0137	0.0500	0.0245	0.1821	0.0563	0.0249	0.0505	0.0171	0.0560	0.1672	0.1657
138	0.0227	0.0228	0.0056	0.0226	0.0160	0.0459	0.0744	0.1984	0.0336	0.0653	0.1938	0.0733	0.0918	0.0657	0.0341	0.0342
139	0.0406	0.0236	0.0082	0.0404	0.0194	0.1049	0.0552	0.1445	0.0319	0.0506	0.1391	0.0549	0.1704	0.0511	0.0324	0.0325
140	0.0199	0.0262	0.0054	0.0198	0.0173	0.0363	0.1057	0.1474	0.0410	0.0912	0.1496	0.1034	0.0622	0.0911	0.0417	0.0418
141	0.0827	0.0379	0.2146	0.0826	0.0359	0.0647	0.0473	0.0551	0.0415	0.0465	0.0550	0.0474	0.0586	0.0466	0.0418	0.0418
142	0.0193	0.0404	0.0061	0.0192	0.0222	0.0290	0.1256	0.0656	0.0667	0.1202	0.0671	0.1239	0.0404	0.1191	0.0675	0.0676
143	0.0716	0.0357	0.0158	0.0714	0.0307	0.0953	0.0648	0.0963	0.0450	0.0613	0.0956	0.0647	0.0990	0.0618	0.0456	0.0456
144	0.0108	0.0563	0.0039	0.0108	0.0212	0.0156	0.0839	0.0319	0.1476	0.0992	0.0326	0.0844	0.0203	0.0969	0.1427	0.1419
145	0.0258	0.0227	0.0061	0.0256	0.0162	0.0546	0.0675	0.1912	0.0327	0.0600	0.1843	0.0667	0.1195	0.0604	0.0332	0.0333
146	0.0358	0.0500	0.0122	0.0357	0.0572	0.0504	0.0843	0.0828	0.0653	0.0823	0.0836	0.0845	0.0601	0.0829	0.0664	0.0666
147	0.0173	0.0602	0.0074	0.0172	0.3954	0.0224	0.0483	0.0345	0.0618	0.0502	0.0350	0.0486	0.0261	0.0504	0.0626	0.0627
148	0.3398	0.0127	0.0096	0.3413	0.0109	0.0599	0.0203	0.0304	0.0152	0.0195	0.0301	0.0203	0.0396	0.0196	0.0154	0.0154
149	0.0378	0.0591	0.0151	0.0377	0.0928	0.0488	0.0757	0.0694	0.0697	0.0755	0.0700	0.0760	0.0550	0.0759	0.0707	0.0709
150	0.0255	0.1537	0.0121	0.0255	0.0713	0.0316	0.0636	0.0451	0.0973	0.0666	0.0456	0.0640	0.0364	0.0667	0.0974	0.0974
151	0.0126	0.0069	0.8667	0.0126	0.0066	0.0105	0.0083	0.0094	0.0075	0.0082	0.0093	0.0083	0.0098	0.0082	0.0075	0.0075
152	0.0131	0.0620	0.0047	0.0130	0.0267	0.0186	0.0893	0.0377	0.1284	0.1004	0.0385	0.0899	0.0241	0.0992	0.1273	0.1271
153	0.0343	0.0202	0.0068	0.0341	0.0151	0.0906	0.0485	0.1297	0.0273	0.0443	0.1240	0.0482	0.2770	0.0446	0.0277	0.0278
154	0.0124	0.0711	0.0046	0.0124	0.0310	0.0176	0.0768	0.0344	0.1452	0.0875	0.0351	0.0775	0.0224	0.0865	0.1429	0.1426
155	0.0144	0.0758	0.0054	0.0144	0.0327	0.0199	0.0809	0.0377	0.1325	0.0891	0.0385	0.0815	0.0253	0.0887	0.1318	0.1316
156	0.3462	0.0119	0.0094	0.3623	0.0104	0.0515	0.0189	0.0279	0.0142	0.0181	0.0276	0.0189	0.0354	0.0183	0.0144	0.0144
157	0.0076	0.0258	0.0034	0.0076	0.7468	0.0097	0.0200	0.0146	0.0255	0.0207	0.0147	0.0201	0.0112	0.0208	0.0258	0.0259
158	0.0337	0.0162	0.0060	0.0335	0.0127	0.1237	0.0363	0.0947	0.0216	0.0334	0.0906	0.0361	0.3840	0.0337	0.0219	0.0219
159	0.0241	0.0230	0.0060	0.0240	0.0177	0.0492	0.0702	0.2024	0.0336	0.0624	0.1966	0.0695	0.0900	0.0629	0.0342	0.0343
160	0.0855	0.0204	0.0101	0.0849	0.0172	0.2838	0.0387	0.0741	0.0258	0.0363	0.0725	0.0386	0.1234	0.0366	0.0261	0.0262
161	0.0185	0.0925	0.0072	0.0185	0.0413	0.0250	0.0786	0.0442	0.1172	0.0840	0.0449	0.0791	0.0310	0.0839	0.1171	0.1170
162	0.0150	0.0267	0.0045	0.0150	0.0163	0.0249	0.1529	0.0801	0.0462	0.1292	0.0825	0.1472	0.0381	0.1272	0.0471	0.0471
163	0.1321	0.0264	0.0154	0.1315	0.0230	0.1461	0.0458	0.0753	0.0325	0.0435	0.0742	0.0458	0.0987	0.0438	0.0329	0.0330
164	0.0159	0.1076	0.0063	0.0159	0.0576	0.0217	0.0687	0.0381	0.1280	0.0750	0.0388	0.0693	0.0267	0.0748	0.1278	0.1278
165	0.0290	0.0674	0.0132	0.0290	0.2139	0.0362	0.0630	0.0511	0.0703	0.0642	0.0516	0.0633	0.0408	0.0645	0.0712	0.0713
166	0.0146	0.0079	0.8478	0.0146	0.0075	0.0121	0.0095	0.0107	0.0085	0.0093	0.0107	0.0095	0.0112	0.0093	0.0085	0.0085
167	0.0182	0.0416	0.0060	0.0181	0.0313	0.0279	0.1175	0.0667	0.0710	0.1169	0.0683	0.1170	0.0379	0.1164	0.0725	0.0726

168	0.2077	0.0271	0.0290	0.2083	0.0241	0.0811	0.0398	0.0538	0.0315	0.0385	0.0534	0.0398	0.0635	0.0387	0.0318	0.0318
169	0.0239	0.1662	0.0111	0.0238	0.0679	0.0298	0.0626	0.0433	0.0995	0.0657	0.0438	0.0630	0.0345	0.0658	0.0996	0.0996
170	0.0185	0.0498	0.0063	0.0184	0.0280	0.0266	0.1122	0.0555	0.0837	0.1139	0.0568	0.1119	0.0352	0.1134	0.0848	0.0849
171	0.3418	0.0118	0.0085	0.3493	0.0102	0.0613	0.0193	0.0296	0.0143	0.0185	0.0293	0.0193	0.0392	0.0186	0.0144	0.0144
172	0.0096	0.1169	0.0037	0.0096	0.0224	0.0132	0.0505	0.0239	0.1944	0.0572	0.0244	0.0510	0.0165	0.0568	0.1759	0.1741
173	0.0167	0.0277	0.0049	0.0166	0.0172	0.0285	0.1377	0.0967	0.0463	0.1187	0.0992	0.1334	0.0448	0.1172	0.0472	0.0472
174	0.0138	0.0319	0.0044	0.0137	0.0167	0.0215	0.1486	0.0554	0.0587	0.1446	0.0569	0.1444	0.0310	0.1396	0.0594	0.0594
175	0.0147	0.0273	0.0044	0.0147	0.0165	0.0245	0.1508	0.0784	0.0475	0.1318	0.0807	0.1455	0.0373	0.1290	0.0484	0.0485
176	0.0629	0.0262	0.0106	0.0626	0.0213	0.1338	0.0548	0.1082	0.0342	0.0509	0.1059	0.0546	0.1534	0.0513	0.0347	0.0347
177	0.0117	0.3701	0.0049	0.0117	0.0301	0.0155	0.0437	0.0253	0.1028	0.0475	0.0257	0.0441	0.0187	0.0474	0.1005	0.1003
178	0.0423	0.0179	0.0070	0.0420	0.0140	0.1626	0.0385	0.0909	0.0235	0.0356	0.0876	0.0383	0.3163	0.0359	0.0238	0.0238
179	0.0142	0.0508	0.0050	0.0142	0.0234	0.0212	0.1078	0.0453	0.0993	0.1211	0.0463	0.1075	0.0283	0.1178	0.0990	0.0988
180	0.0101	0.0597	0.0037	0.0101	0.0226	0.0144	0.0739	0.0290	0.1650	0.0871	0.0296	0.0746	0.0186	0.0856	0.1584	0.1575
181	0.2838	0.0174	0.0131	0.2806	0.0150	0.0822	0.0277	0.0415	0.0208	0.0265	0.0410	0.0277	0.0541	0.0267	0.0210	0.0210
182	0.1216	0.0383	0.0675	0.1213	0.0360	0.0825	0.0516	0.0642	0.0432	0.0503	0.0639	0.0516	0.0705	0.0505	0.0435	0.0436
183	0.0312	0.0285	0.0078	0.0311	0.0230	0.0627	0.0745	0.1557	0.0404	0.0681	0.1531	0.0740	0.0991	0.0686	0.0411	0.0411
184	0.1872	0.0297	0.0341	0.1875	0.0263	0.0827	0.0429	0.0570	0.0343	0.0415	0.0566	0.0429	0.0663	0.0417	0.0346	0.0346
185	0.0180	0.0735	0.0079	0.0180	0.3410	0.0232	0.0508	0.0355	0.0699	0.0531	0.0360	0.0512	0.0271	0.0533	0.0707	0.0708
186	0.0227	0.0308	0.0064	0.0226	0.0228	0.0393	0.1048	0.1264	0.0475	0.0937	0.1283	0.1035	0.0602	0.0941	0.0485	0.0486
187	0.2520	0.0176	0.0115	0.2492	0.0150	0.1161	0.0294	0.0470	0.0213	0.0280	0.0464	0.0294	0.0656	0.0282	0.0216	0.0216
188	0.0237	0.1674	0.0110	0.0236	0.0685	0.0296	0.0624	0.0431	0.0996	0.0655	0.0436	0.0628	0.0343	0.0656	0.0997	0.0998
189	0.0222	0.1802	0.0101	0.0221	0.0651	0.0279	0.0612	0.0413	0.1016	0.0645	0.0418	0.0616	0.0325	0.0646	0.1016	0.1017

SILHOUETTES
****

FOR THE ENTIRE DATA SET, THE AVERAGE SILHOUETTE WIDTH IS 0.30

CLOSEST HARD CLUSTERING
****

FOR THIS HARD CLUSTERING, IT TURNS OUT THAT
ONLY THE FIRST 13 CLUSTERS ARE NONEMPTY.

CLUSTER NUMBER	SIZE	OBJECTS
1	8	001 008 040 111 132 181 182 187
2	22	002 004 014 015 019 025 037 041 062 081 090 096 100 105 123 125 136 150 169 177 188 189
3	11	003 036 052 057 077 098 104 124 141 151 166
4	14	005 021 022 049 061 078 099 106 118 148 156 168 171 184
5	16	006 009 028 039 054 064 071 075 107 119 133 147 149 157 165 185
6	10	007 034 079 085 092 101 112 127 160 163
7	22	010 013 018 031 045 047 060 067 069 070 084 103 108 122 131 134 142 162 167 173 174 175
8	12	011 016 029 038 089 102 116 120 138 145 159 183
9	29	012 026 030 032 035 042 048 050 058 059 074 083 086 088 095 115 117 121 126 135 137 144 152 154 155 161 164 172 180
10	18	017 023 024 043 046 053 056 063 068 073 076 094 109 128 129 130 170 179
11	8	020 055 072 082 113 114 140 186
12	3	027 065 146
13	16	033 044 051 066 080 087 091 093 097 110 139 143 153 158 176 178

Appendix D

Statistics for Different Proportions of Data Sets

10(70-30)

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Min.	Max.	Range
Footing width, B (m)					
Training set	8.8	10.8	0.8	60.0	59.2
Testing set	8.8	8.8	1.0	42.7	41.7
Validation set	7.8	9.2	1.0	34.0	33.0
Footing net applied pressure, q (kPa)					
Training set	185.5	118.3	18.3	697.0	678.6
Testing set	192.5	128.1	33.0	584.0	551.0
Validation set	182.2	145.3	64.0	666.0	602.0
Average SPT blow count, N					
Training set	25.3	14.1	4.0	60.0	56.0
Testing set	24.0	12.4	4.0	50.0	46.0
Validation set	21.2	12.5	5.0	45.0	40.0
Footing geometry, L/B					
Training set	2.1	1.8	1.0	10.5	9.5
Testing set	2.3	1.6	1.0	8.0	7.0
Validation set	2.3	1.9	1.0	7.9	6.8
Footing embedment ratio, D_f/B					
Training set	0.53	0.58	0.0	3.4	3.4
Testing set	0.54	0.63	0.0	3.0	3.0
Validation set	0.46	0.31	0.04	1.09	1.05
Measured settlement, S_m (mm)					
Training set	18.9	24.1	0.6	121.0	120.4
Testing set	21.8	29.0	1.5	120.0	118.5
Validation set	25.8	34.4	1.3	116.0	114.7

10(80-20)

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Min.	Max.	Range
Footing width, *B* (m)					
Training set	8.8	10.3	0.8	60.0	59.2
Testing set	7.7	9.0	0.9	41.2	40.3
Validation set	10.0	11.1	1.0	34.0	33.0
Footing net applied pressure, *q* (kPa)					
Training set	185.6	120.6	18.3	697.0	678.7
Testing set	179.6	123.6	52.0	584.0	532.0
Validation set	210.6	143.8	25.0	576.0	551.0
Average SPT blow count, *N*					
Training set	24.7	13.4	4.0	60.0	56.0
Testing set	24.1	12.5	6.0	50.0	44.0
Validation set	24.1	16.4	6.0	60.0	54.0
Footing geometry, *L/B*					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.3	1.6	1.0	6.7	5.6
Validation set	2.5	2.1	1.0	8.0	7.0
Footing embedment ratio, D_f/B					
Training set	0.53	0.58	0.0	3.4	3.4
Testing set	0.54	0.55	0.0	3.0	3.0
Validation set	0.49	0.60	0.0	2.6	2.6
Measured settlement, S_m (mm)					
Training set	20.3	27.1	0.6	121.0	120.4
Testing set	20.4	26.5	1.3	100.0	98.7
Validation set	21.3	23.9	1.5	92.0	90.5

10(90-10)

Model variables and data sets	Statistical parameters				
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Range**
Footing width, B (m)					
Training set	8.4	9.7	0.8	60.0	59.2
Testing set	9.1	9.3	0.9	33.0	32.1
Validation set	10.9	13.8	1.0	55.0	54.0
Footing net applied pressure, q (kPa)					
Training set	185.1	119.2	18.3	697.0	678.7
Testing set	198.3	142.7	34.0	584.0	550.0
Validation set	192.4	142.2	41.0	666.0	625.0
Average SPT blow count, N					
Training set	24.9	13.9	4.0	60.0	56.0
Testing set	22.7	12.5	6.0	50.0	44.0
Validation set	23.1	11.5	7.0	60.0	53.0
Footing geometry, L/B					
Training set	2.3	1.9	1.0	10.5	9.5
Testing set	1.6	0.9	1.0	4.1	3.1
Validation set	1.7	1.1	1.0	5.0	4.0
Footing embedment ratio, D_f/B					
Training set	0.53	0.58	0.0	3.4	3.4
Testing set	0.50	0.45	0.0	1.4	1.4
Validation set	0.52	0.66	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training set	19.6	25.6	0.6	121.0	120.4
Testing set	22.9	29.3	2.7	120.0	117.3
Validation set	24.8	31.9	1.3	100.0	98.7

20(80-20)

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Min.	Max.	Range
Footing width, B (m)					
Training set	8.4	9.6	0.8	60.0	59.2
Testing set	9.2	12.3	0.9	55.0	54.1
Validation set	9.4	10.1	0.9	41.2	40.3
Footing net applied pressure, q (kPa)					
Training set	186.5	123.5	18.3	697.0	678.7
Testing set	188.0	135.7	25.0	584.0	559.0
Validation set	187.9	114.6	33.0	575.0	542.0
Average SPT blow count, N					
Training set	24.2	13.3	4.0	60.0	56.0
Testing set	26.1	13.8	5.0	60.0	55.0
Validation set	24.3	14.1	4.0	55.0	51.0
Footing geometry, L/B					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.4	2.1	1.0	9.9	8.9
Validation set	2.1	1.8	1.0	8.0	7.0
Footing embedment ratio, D_f/B					
Training set	0.49	0.55	0.0	3.4	3.4
Testing set	0.58	0.61	0.0	3.0	3.0
Validation set	0.59	0.64	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training set	20.1	26.5	0.6	121.0	120.4
Testing set	21.4	29.2	1.0	120.0	119.0
Validation set	20.4	25.2	1.3	120.0	118.7

20(90-10)

Model variables and data sets	Statistical parameters				
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Range**
Footing width, *B* (m)					
Training set	9.0	10.8	0.8	60.0	59.2
Testing set	8.7	9.8	0.9	55.0	54.1
Validation set	7.8	7.7	1.0	33.0	32.0
Footing net applied pressure, *q* (kPa)					
Training set	184.4	124.1	18.3	697.0	678.7
Testing set	179.7	116.9	25.0	666.0	641.0
Validation set	190.8	120.3	41.0	584.0	543.0
Average SPT blow count, *N*					
Training set	25.0	13.8	4.0	60.0	56.0
Testing set	24.3	13.1	5.0	60.0	55.0
Validation set	24.2	13.3	4.0	60.0	56.0
Footing geometry, *L/B*					
Training set	2.2	1.9	1.0	10.5	9.5
Testing set	2.1	1.7	1.0	9.9	8.9
Validation set	2.2	1.5	1.0	6.7	5.7
Footing embedment ratio, D_f/B					
Training set	0.53	0.56	0.0	3.4	3.4
Testing set	0.51	0.52	0.0	3.0	3.0
Validation set	0.56	0.63	0.0	3.0	3.0
Measured settlement, S_m (mm)					
Training set	19.6	26.4	0.6	121.0	120.4
Testing set	19.3	25.0	1.0	120.0	119.0
Validation set	20.8	25.2	1.3	120.0	118.7

30(70-30)

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Min.	Max.	Range
Footing width, *B* (m)					
Training set	7.9	9.9	0.8	60.0	59.2
Testing set	9.3	10.2	0.9	42.7	41.8
Validation set	9.6	10.4	0.9	55.0	54.1
Footing net applied pressure, *q* (kPa)					
Training set	180.4	117.7	18.3	697.0	678.7
Testing set	188.3	126.3	63.0	666.0	603.0
Validation set	196.9	131.0	34.0	584.0	550.0
Average SPT blow count, *N*					
Training set	24.2	14.1	4.0	60.0	56.0
Testing set	24.9	13.4	5.0	55.0	50.0
Validation set	24.8	12.6	4.0	60.0	56.0
Footing geometry, *L/B*					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.2	2.1	1.0	9.9	8.9
Validation set	2.1	1.6	1.0	8.4	7.4
Footing embedment ratio, D_f/B					
Training set	0.57	0.65	0.0	3.4	3.4
Testing set	0.50	0.56	0.0	2.6	2.6
Validation set	0.48	0.45	0.0	2.1	2.1
Measured settlement, S_m (mm)					
Training set	20.3	28.1	0.6	121.0	120.4
Testing set	21.5	26.1	1.0	91.6	90.6
Validation set	19.7	24.7	1.3	116.0	114.7

30(80-20)

Model variables and data sets	Statistical parameters				
	Mean	Std. Dev.	Min.	Max.	Range
Footing width, *B* (m)					
Training set	8.2	10.3	0.8	60.0	59.2
Testing set	9.0	9.0	0.9	33.0	32.1
Validation set	9.6	10.4	0.9	55.0	54.1
Footing net applied pressure, *q* (kPa)					
Training set	179.9	118.8	18.3	697.0	678.7
Testing set	194.0	125.9	75.0	666.0	591.0
Validation set	196.9	131.0	34.0	584.0	550.0
Average SPT blow count, *N*					
Training set	24.7	14.0	4.0	60.0	56.0
Testing set	23.3	13.4	5.0	55.0	50.0
Validation set	24.8	12.6	4.0	60.0	56.0
Footing geometry, *L/B*					
Training set	2.1	1.7	1.0	10.5	9.5
Testing set	2.2	2.2	1.0	9.9	8.9
Validation set	2.1	1.6	1.0	8.5	7.5
Footing embedment ratio, D_f/B					
Training set	0.55	0.63	0.0	3.4	3.4
Testing set	0.53	0.59	0.0	2.6	2.6
Validation set	0.48	0.45	0.0	2.1	2.1
Measured settlement, S_m (mm)					
Training set	20.3	28.0	0.6	121.0	120.4
Testing set	22.1	25.6	1.5	91.6	90.1
Validation set	19.7	24.7	1.3	116.0	114.7

30(90-10)

Model variables and data sets	Statistical parameters				
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Range**
Footing width, B (m)					
Training set	9.2	10.7	0.8	60.0	59.2
Testing set	7.7	7.9	1.2	30.2	29.0
Validation set	8.0	9.3	0.9	42.7	41.8
Footing net applied pressure, q (kPa)					
Training set	173.1	104.6	18.3	697.0	678.7
Testing set	157.6	94.6	70.0	386.0	316.0
Validation set	222.8	155.2	41.0	666.0	625.0
Average SPT blow count, N					
Training set	25.1	14.5	4.0	60.0	56.0
Testing set	20.4	12.3	4.0	50.0	46.0
Validation set	24.2	11.3	5.0	50.0	45.0
Footing geometry, L/B					
Training set	2.3	1.9	1.0	10.5	9.5
Testing set	2.3	1.9	1.0	6.8	5.8
Validation set	1.8	1.3	1.0	7.9	6.9
Footing embedment ratio, D_f/B					
Training set	0.54	0.64	0.0	3.4	3.4
Testing set	0.47	0.40	0.1	1.2	1.1
Validation set	0.52	0.45	0.0	2.1	2.1
Measured settlement, S_m (mm)					
Training set	19.4	25.7	0.6	121.0	120.4
Testing set	25.8	30.1	2.1	91.6	89.5
Validation set	21.3	27.7	3.6	120.0	116.4

Appendix E

Null Hypothesis Tests for Different Proportions of Data Sets

10(70-30)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	0.00	−1.97	1.97	Accept	1.51	0.64	1.64	Accept
Validation	0.38	−1.97	1.97	Accept	1.37	0.53	2.26	Reject
q								
Testing	−0.34	−1.97	1.97	Accept	0.85	0.64	1.64	Accept
Validation	0.11	−1.97	1.97	Accept	0.66	0.53	2.26	Accept
N								
Testing	0.57	−1.97	1.97	Accept	1.29	0.64	1.64	Accept
Validation	1.19	−1.97	1.97	Accept	1.27	0.53	2.26	Accept
L/B								
Testing	−0.68	−1.97	1.97	Accept	1.26	0.64	1.64	Accept
Validation	−0.45	−1.97	1.97	Accept	0.89	0.53	2.26	Accept
D_f/B								
Testing	−0.10	−1.97	1.97	Accept	0.85	0.64	1.64	Accept
Validation	0.51	−1.97	1.97	Accept	3.50	0.53	2.26	Reject
S_m								
Testing	−0.67	−1.97	1.97	Accept	0.69	0.64	1.64	Accept
Validation	−1.09	−1.97	1.97	Accept	0.49	0.53	2.26	Reject

10(80-20)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	0.57	−1.97	1.97	Accept	1.31	0.61	1.81	Accept
Validation	−0.47	−1.97	1.97	Accept	0.86	0.54	2.24	Accept
q								
Testing	0.26	−1.97	1.97	Accept	0.95	0.61	1.81	Accept
Validation	−0.83	−1.97	1.97	Accept	0.70	0.54	2.24	Accept
N								
Testing	0.24	−1.97	1.97	Accept	1.15	0.61	1.81	Accept
Validation	0.18	−1.97	1.97	Accept	0.67	0.54	2.24	Accept
L/B								
Testing	−0.62	−1.97	1.97	Accept	1.13	0.61	1.81	Accept
Validation	−0.93	−1.97	1.97	Accept	0.65	0.54	2.24	Accept
D_f/B								
Testing	−0.09	−1.97	1.97	Accept	1.11	0.61	1.81	Accept
Validation	0.28	−1.97	1.97	Accept	0.93	0.54	2.24	Accept
S_m								
Testing	−0.02	−1.97	1.97	Accept	1.05	0.61	1.81	Accept
Validation	−0.15	−1.97	1.97	Accept	1.29	0.54	2.24	Accept

10(90-10)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	−0.28	−1.97	1.97	Accept	1.09	0.53	2.37	Accept
Validation	−1.0	−1.97	1.97	Accept	0.49	0.54	2.24	Reject
q								
Testing	−0.42	−1.97	1.97	Accept	0.69	0.53	2.37	Accept
Validation	−0.24	−1.97	1.97	Accept	0.70	0.54	2.24	Accept
N								
Testing	0.62	−1.97	1.97	Accept	1.24	0.53	2.37	Accept
Validation	0.54	−1.97	1.97	Accept	1.46	0.54	2.24	Accept
L/B								
Testing	1.49	−1.97	1.97	Accept	4.46	0.53	2.37	Reject
Validation	1.35	−1.97	1.97	Accept	2.98	0.54	2.24	Reject
D_f/B								
Testing	0.21	−1.97	1.97	Accept	1.66	0.53	2.37	Accept
Validation	0.07	−1.97	1.97	Accept	0.77	0.54	2.24	Accept
S_m								
Testing	−0.49	−1.97	1.97	Accept	0.76	0.53	2.37	Accept
Validation	−0.81	−1.97	1.97	Accept	0.64	0.54	2.24	Accept

20(80-20)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.39	–1.97	1.97	Accept	0.61	0.59	1.87	Accept
Validation	–0.54	–1.97	1.97	Accept	0.90	0.61	1.77	Accept
q								
Testing	–0.05	–1.97	1.97	Accept	0.83	0.59	1.87	Accept
Validation	–0.06	–1.97	1.97	Accept	1.16	0.61	1.77	Accept
N								
Testing	–0.70	–1.97	1.97	Accept	0.93	0.59	1.87	Accept
Validation	–0.04	–1.97	1.97	Accept	0.89	0.61	1.77	Accept
L/B								
Testing	–0.83	–1.97	1.97	Accept	0.65	0.59	1.87	Accept
Validation	0.0	–1.97	1.97	Accept	0.89	0.61	1.77	Accept
D_f/B								
Testing	–0.79	–1.97	1.97	Accept	0.81	0.59	1.87	Accept
Validation	–0.93	–1.97	1.97	Accept	0.74	0.61	1.77	Accept
S_m								
Testing	–0.24	–1.97	1.97	Accept	0.82	0.59	1.87	Accept
Validation	–0.06	–1.97	1.97	Accept	1.11	0.61	1.77	Accept

20(90-10)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	0.10	–1.97	1.97	Accept	1.21	0.51	2.54	Accept
Validation	0.63	–1.97	1.97	Accept	1.96	0.62	1.76	Accept
q								
Testing	0.14	–1.97	1.97	Accept	1.13	0.51	2.54	Accept
Validation	–0.28	–1.97	1.97	Accept	1.06	0.62	1.76	Accept
N								
Testing	0.19	–1.97	1.97	Accept	1.11	0.51	2.54	Accept
Validation	0.32	–1.97	1.97	Accept	1.08	0.62	1.76	Accept
L/B								
Testing	0.19	–1.97	1.97	Accept	1.25	0.51	2.54	Accept
Validation	0.0	–1.97	1.97	Accept	1.60	0.62	1.76	Accept
D_f/B								
Testing	0.13	–1.97	1.97	Accept	1.16	0.51	2.54	Accept
Validation	–0.28	–1.97	1.97	Accept	0.79	0.62	1.76	Accept
S_m								
Testing	0.04	–1.97	1.97	Accept	1.12	0.51	2.54	Accept
Validation	–0.25	–1.97	1.97	Accept	1.09	0.62	1.76	Accept

30(70-30)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.74	–1.97	1.97	Accept	0.94	0.60	1.76	Accept
Validation	–0.99	–1.97	1.97	Accept	0.91	0.63	1.63	Accept
q								
Testing	–0.35	–1.97	1.97	Accept	0.87	0.60	1.76	Accept
Validation	–0.79	–1.97	1.97	Accept	0.81	0.63	1.63	Accept
N								
Testing	–0.27	–1.97	1.97	Accept	1.11	0.60	1.76	Accept
Validation	–0.26	–1.97	1.97	Accept	1.25	0.63	1.63	Accept
L/B								
Testing	–0.29	–1.97	1.97	Accept	0.65	0.60	1.76	Accept
Validation	0.0	–1.97	1.97	Accept	1.13	0.63	1.63	Accept
D_f/B								
Testing	0.51	–1.97	1.97	Accept	1.35	0.60	1.76	Accept
Validation	0.91	–1.97	1.97	Accept	2.06	0.63	1.63	Reject
S_m								
Testing	–0.23	–1.97	1.97	Accept	1.16	0.60	1.76	Accept
Validation	0.13	–1.97	1.97	Accept	1.29	0.63	1.63	Accept

30(80-20)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	–0.37	–1.97	1.97	Accept	1.31	0.57	1.97	Accept
Validation	–0.82	–1.97	1.97	Accept	0.98	0.64	1.62	Accept
q								
Testing	–0.54	–1.97	1.97	Accept	0.89	0.57	1.97	Accept
Validation	–0.84	–1.97	1.97	Accept	0.82	0.64	1.62	Accept
N								
Testing	0.47	–1.97	1.97	Accept	1.09	0.57	1.97	Accept
Validation	–0.04	–1.97	1.97	Accept	1.23	0.64	1.62	Accept
L/B								
Testing	–0.26	–1.97	1.97	Accept	0.59	0.57	1.97	Accept
Validation	0.0	–1.97	1.97	Accept	1.13	0.64	1.62	Accept
D_f/B								
Testing	0.14	–1.97	1.97	Accept	1.14	0.57	1.97	Accept
Validation	0.74	–1.97	1.97	Accept	1.96	0.64	1.62	Reject
S_m								
Testing	–0.30	–1.97	1.97	Accept	1.19	0.57	1.97	Accept
Validation	0.14	–1.97	1.97	Accept	1.29	0.64	1.62	Accept

30(90-10)

Variable and data sets	*t*-value	Lower critical value	Upper critical value	*t*-test	*F*-value	Lower critical value	Upper critical value	*F-test*
B								
Testing	0.49	−1.97	1.97	Accept	1.83	0.49	2.79	Accept
Validation	0.73	−1.97	1.97	Accept	1.32	0.64	1.60	Accept
q								
Testing	0.51	−1.97	1.97	Accept	1.22	0.49	2.79	Accept
Validation	−2.5	−1.97	1.97	Reject	0.45	0.64	1.60	Reject
N								
Testing	1.12	−1.97	1.97	Accept	1.39	0.49	2.79	Accept
Validation	0.41	−1.97	1.97	Accept	1.65	0.64	1.60	Accept
L/B								
Testing	0.00	−1.97	1.97	Accept	1.0	0.49	2.79	Accept
Validation	1.79	−1.97	1.97	Accept	2.13	0.64	1.60	Accept
D_f/B								
Testing	0.39	−1.97	1.97	Accept	2.56	0.49	2.79	Accept
Validation	0.21	−1.97	1.97	Accept	2.0	0.64	1.60	Reject
S_m								
Testing	−0.84	−1.97	1.97	Accept	0.73	0.49	2.79	Accept
Validation	−0.45	−1.97	1.97	Accept	0.86	0.64	1.60	Accept

Appendix F

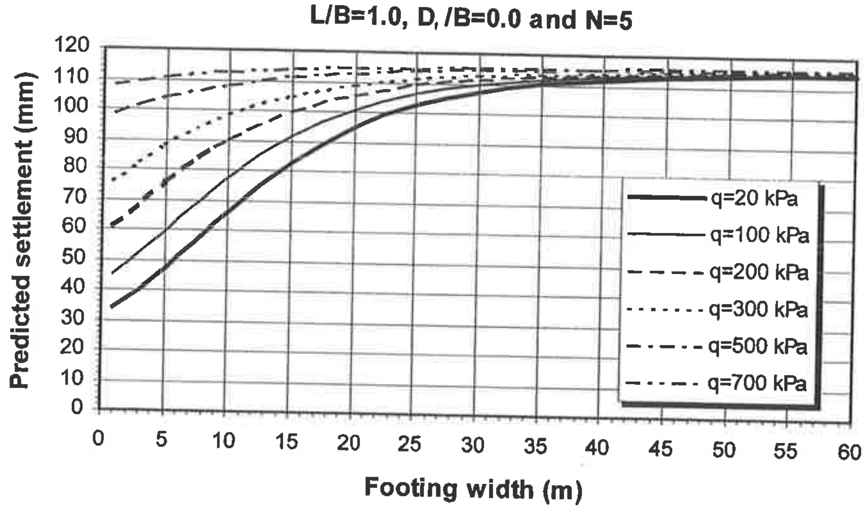
FORTRAN Code for the ANN Model

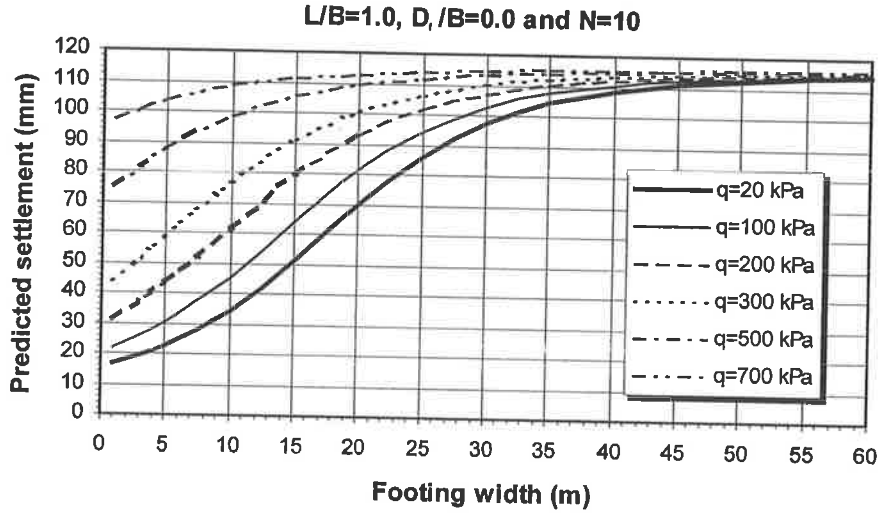
```
C      THIS PROGRAM CALCULATES SETTLEMENT OF SHALLOW FOUNDATIONS
C      ON COHESIONLESS SOILS USING ARTIFICIAL NEURAL NETWORKS
C----------------------------------------------------------------------
C      B-FOOTING WIDTH (m)
C      q-FOOTING NET APPLIED PRESSURE (kPa)
C      N-AVERAGE SPT BLOW COUNT
C      LB-FOOTING GEOMETRY (LENGTH/WIDTH)
C      Df-FOOTING EMBEDMENT RATIO (FOUNDATION DEPTH/FOOTING WIDTH)
C      S-PREDICTED SETTLEMENT (mm)
C----------------------------------------------------------------------
       REAL B,B1,q,q1,N,N1,LB,LB1,Df,Df1,SM,SM1
       REAL H1,H2,Y1,Y2,SUMM
       CHARACTER*20 NAME
       WRITE(*,*)'Enter data from screen or file (S/F)?'
       READ(*,5)NAME
 5     FORMAT(A20)
       IF ( (NAME .EQ. 'S') .OR. (NAME .EQ. 's') ) THEN
       WRITE(*,*)'Footing width (m)='
       READ(*,*)B
       WRITE(*,*)'Footing net applied pressure (kPa)='
       READ(*,*)q
       WRITE(*,*)'Average SPT blow count ='
       READ(*,*)N
       WRITE(*,*)'Footing geometry ='
       READ(*,*)LB
       WRITE(*,*)'Footing embedment ratio ='
       READ(*,*)Df
       B1=(B-0.8)/59.2
       q1=(q-18.32)/678.68
       N1=(N-4)/56
       LB1=(LB-1.0)/9.5833
       Df1=(Df-0)/3.4444
       H1=(0.12418158)+(0.22735420*B1)+(0.48116106*q1)+(0.22959290*N1)
      *-(0.01703092*LB1)+(0.06734087*Df1)
       H2=(0.18811224)-(2.44251320*B1)-(1.11489062*q1)+(4.23963997*N1)
      *-(0.49885305*LB1)+(2.50030115*Df1)
       Y1=TANH(H1)
       Y2=TANH(H2)
       SUMM=(-0.31274397)+(0.72535181*Y1)-(2.98416472*Y2)
       SM1=1/(1+EXP(-SUMM))
       SM=(SM1*120.4)+0.6
       WRITE(*,400)SM
400    FORMAT(' Settlement =',F10.2,' mm')
       ELSE
       CALL TRY
       END IF
       END
C----------------------------------------------------------------
       SUBROUTINE TRY
       DIMENSION SUM(500),S1(500),S(500),Y1(500),Y2(500)
```

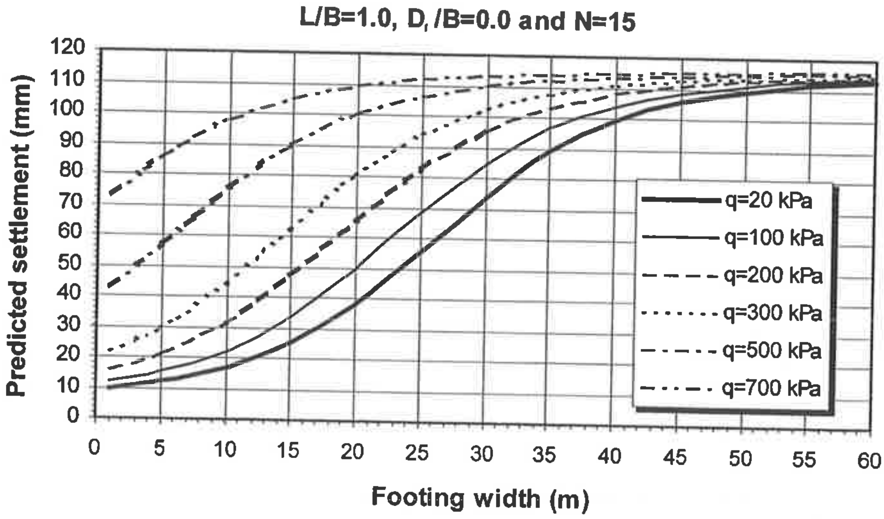
```
      DIMENSION H1(500),H2(500),R(500,5),RN(500,5)
      CHARACTER*20 FILIN,FILOUT
      WRITE(*,*)'Enter the number of case records ='
      READ(*,*)N
      WRITE(*,*)'Enter the name of input file ='
      READ(*,10)FILIN
      WRITE(*,*)'Enter the name of output file ='
      READ(*,10)FILOUT
 10   FORMAT(A20)
      OPEN(UNIT=1,FILE=FILIN,STATUS='OLD')
      OPEN(UNIT=2,FILE=FILOUT,STATUS='NEW')
      DO 20 I=1,N
      READ(1,*)(R(I,J),J=1,5)
 20   CONTINUE
      DO 50 I=1,N
      RN(I,1)=(R(I,1)-0.8)/59.2
      RN(I,2)=(R(I,2)-18.32)/678.68
      RN(I,3)=(R(I,3)-4)/56
      RN(I,4)=(R(I,4)-1.0)/9.5833
      RN(I,5)=(R(I,5)-0.0)/3.4444
 50   CONTINUE
      DO 60 I=1,N
      H1(I)=(0.12418158)+(0.22735420*RN(I,1))+(0.48116106*RN(I,2))
     *+(0.22959290*RN(I,3))-(0.01703092*RN(I,4))+(0.06734087*RN(I,5))
      H2(I)=(0.18811224)-(2.44251320*RN(I,1))-(1.11489062*RN(I,2))
     *+(4.23963997*RN(I,3))-(0.49885305*RN(I,4))+(2.50030115*RN(I,5))
 60   CONTINUE
      DO 70 I=1,N
      Y1(I)=TANH(H1(I))
      Y2(I)=TANH(H2(I))
 70   CONTINUE
      DO 80 I=1,N
      SUM(I)=(-0.31274397)+(0.72535181*Y1(I))-(2.98416472*Y2(I))
      S1(I)=1/(1+EXP(-SUM(I)))
      S(I)=(S1(I)*120.4)+0.6
 80   CONTINUE
      WRITE(2,15)
 15   FORMAT(6X,'FOOT.',3X,'APP.',7X,'SPT-N',6X,'GEOMETRY',2X,'EMBED.'
     *,3x,'SETTLEMENT')
      WRITE(2,25)
 25   FORMAT(6X,'WIDTH',3X,'PRESS.',16X,'RATIO',5X,'RATIO')
      WRITE(2,35)
 35   FORMAT(6X,'(m)',5X,'(kPa)',38X,'(mm)')
      WRITE(2,*)'-----------------------------------------------------
     *--------------------'
      DO 90 I=1,N
      WRITE(2,200)(R(I,J),J=1,5),S(I)
 200  FORMAT(6F10.2)
 90   CONTINUE
      RETURN
      END
```

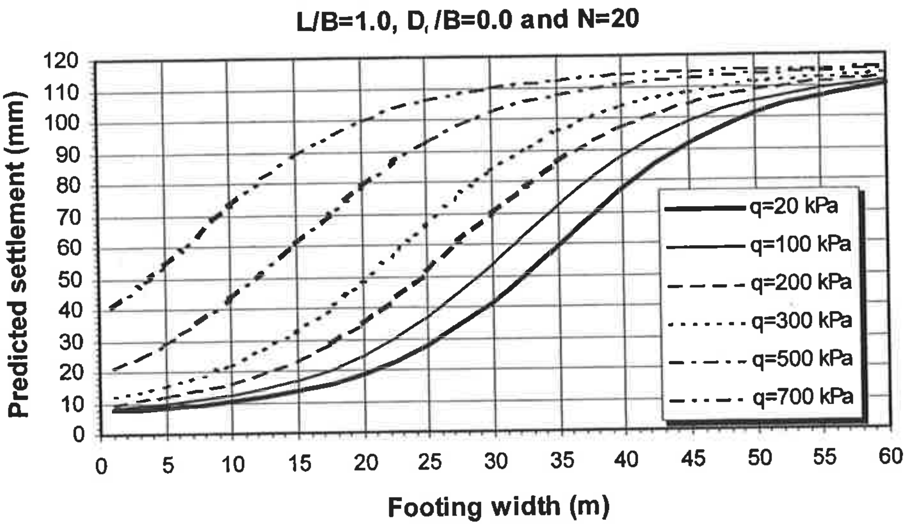
Appendix G. ANN-Based Design Charts

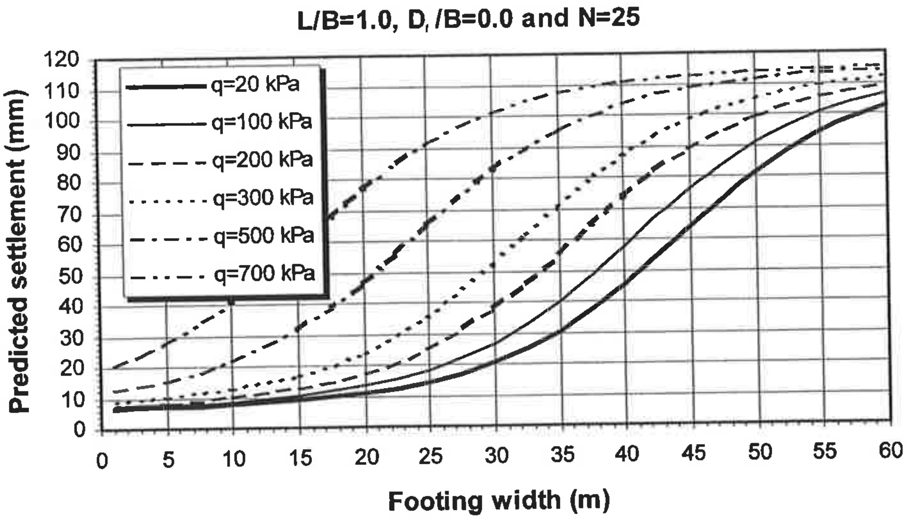
- $L/B = 1.0, D_f/B = 0.0$

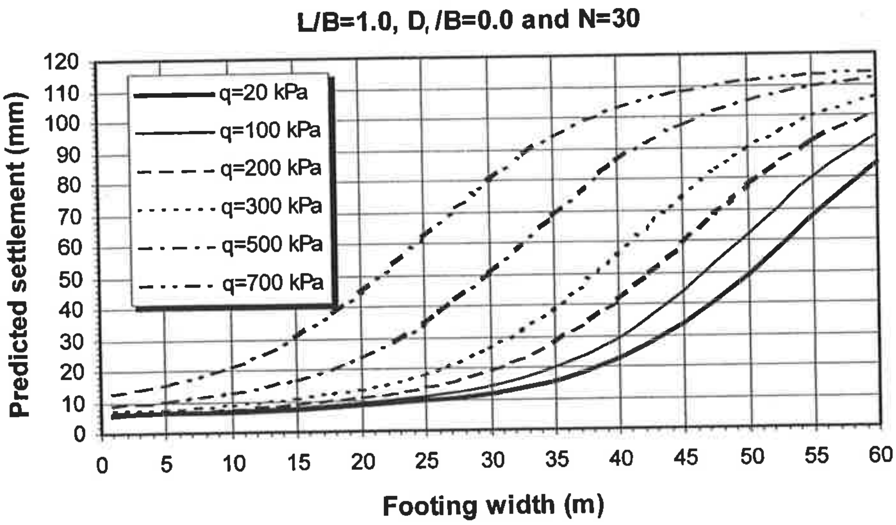


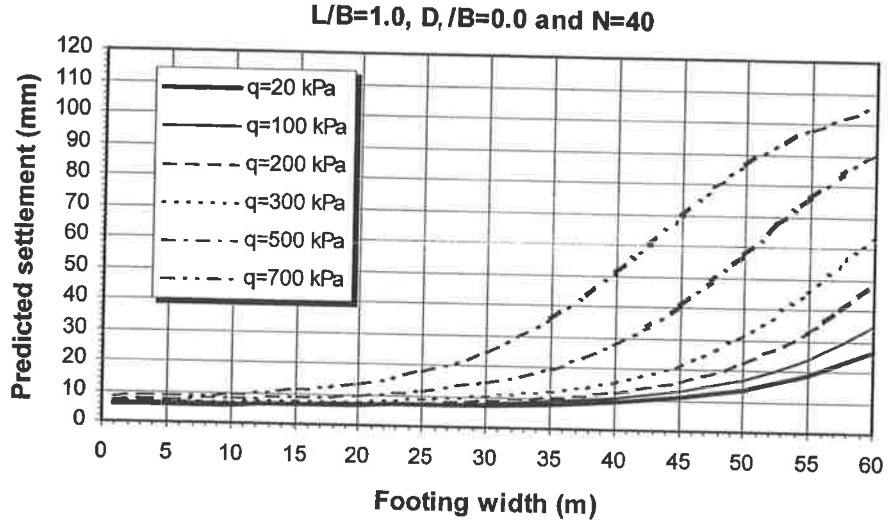


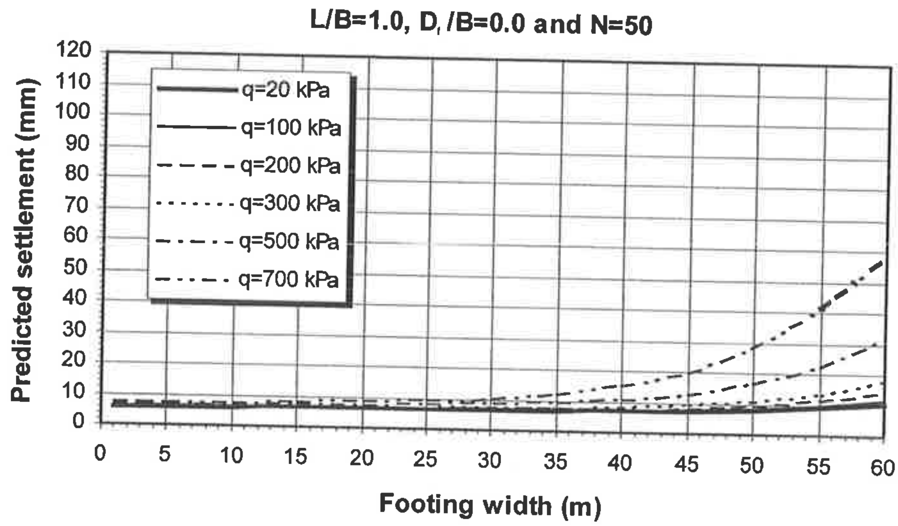


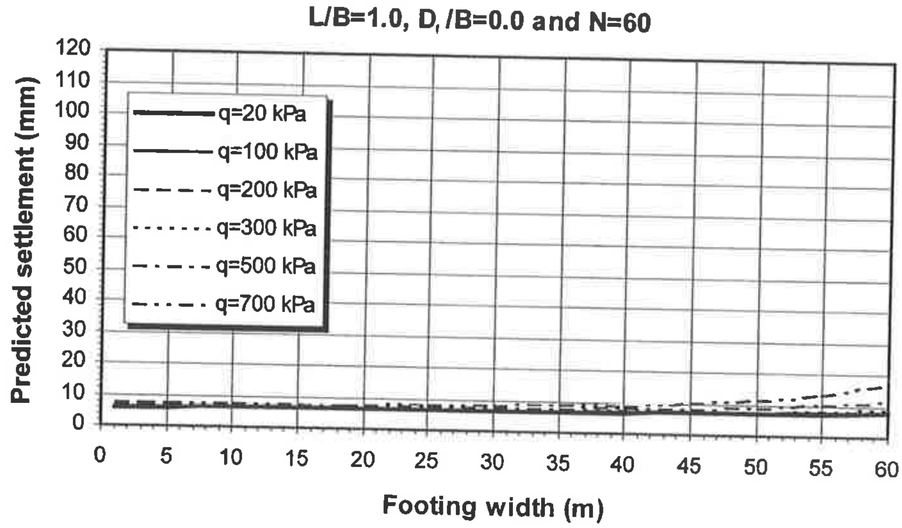

L/B=1.0, D_f/B=0.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=1.0, D_f/B=0.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

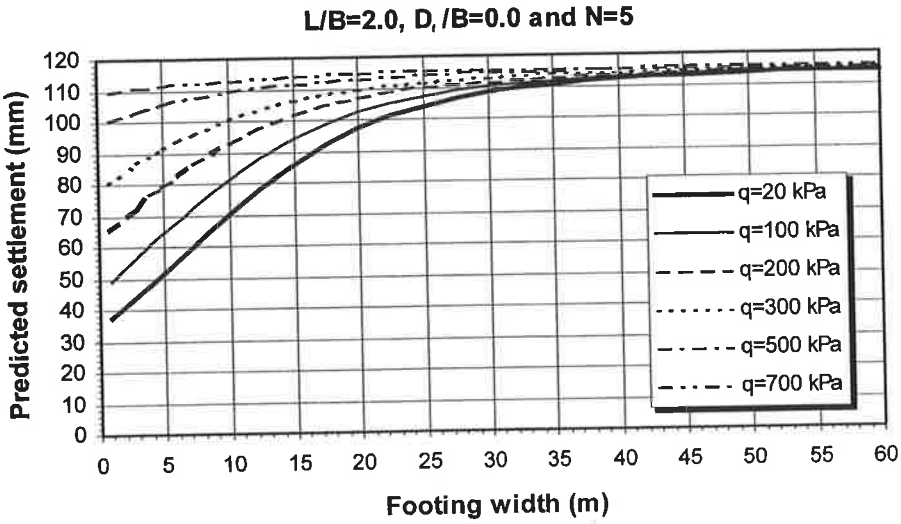

L/B=1.0, D_f/B=0.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

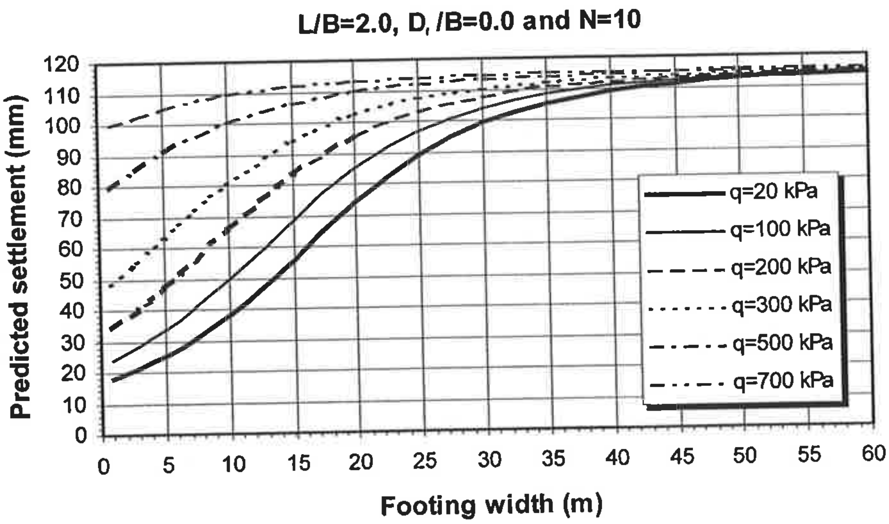

L/B=1.0, D_f /B=0.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

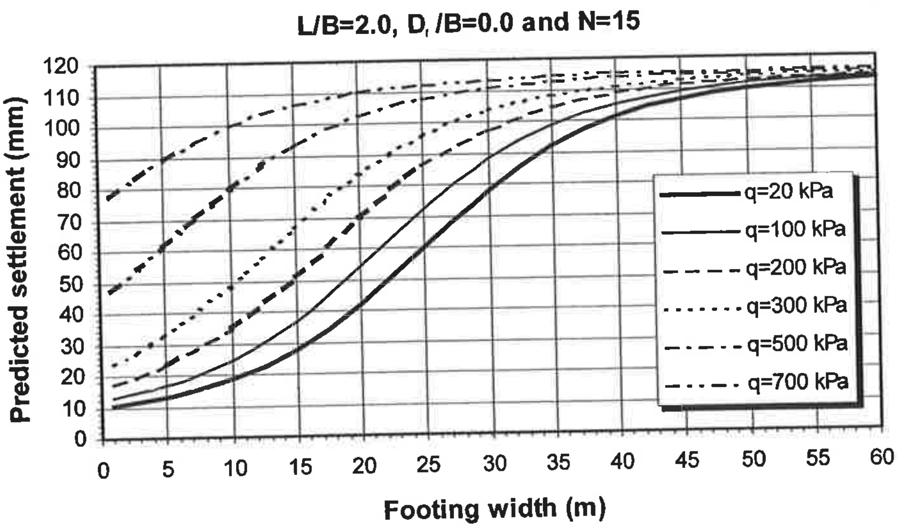

L/B=1.0, D_f /B=0.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

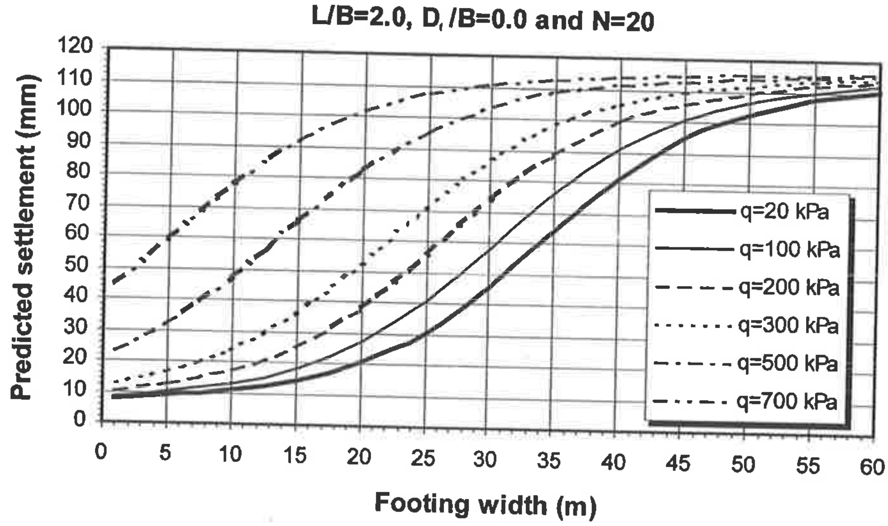

L/B=1.0, D_f /B=0.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

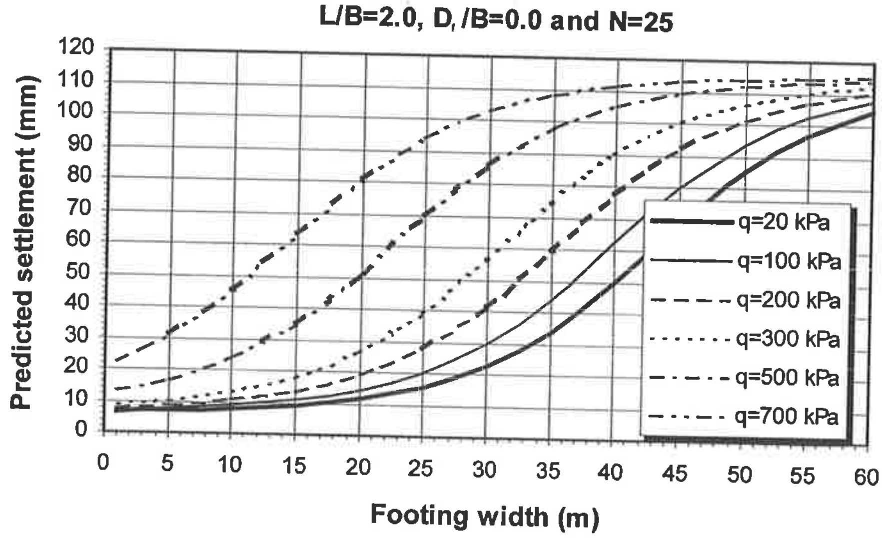
- $L/B = 2.0, D_f/B = 0.0$

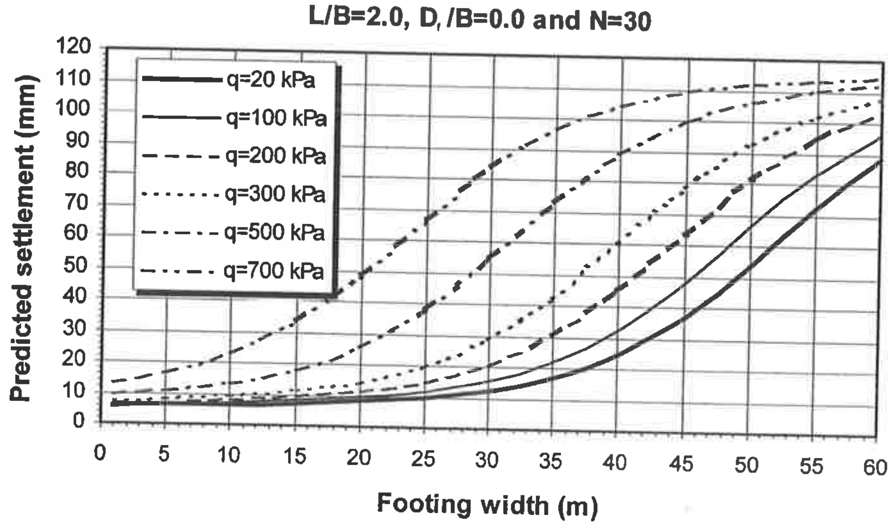

L/B=2.0, Df/B=0.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, Df/B=0.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

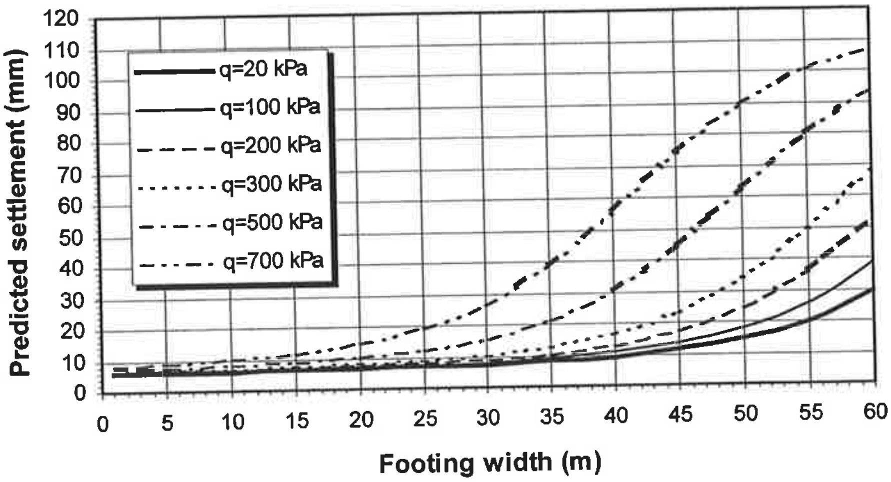

L/B=2.0, Df/B=0.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D_f/B=0.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

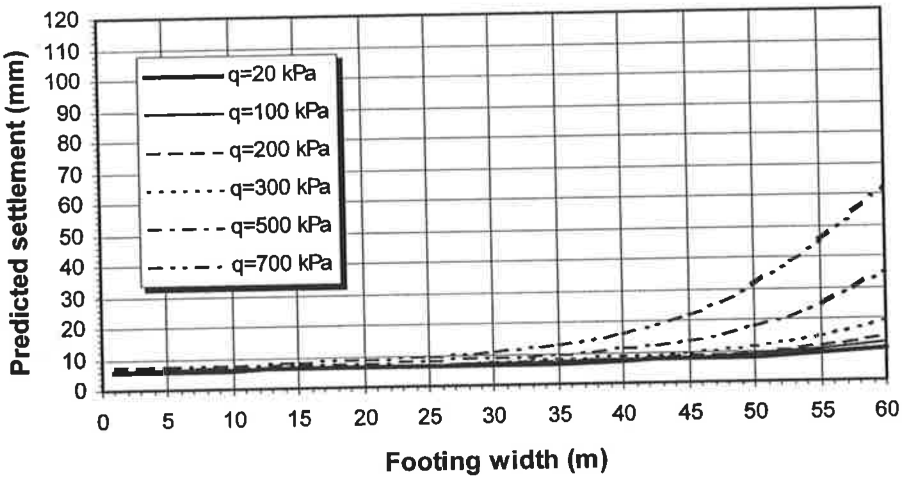

L/B=2.0, D_f/B=0.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D_f/B=0.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

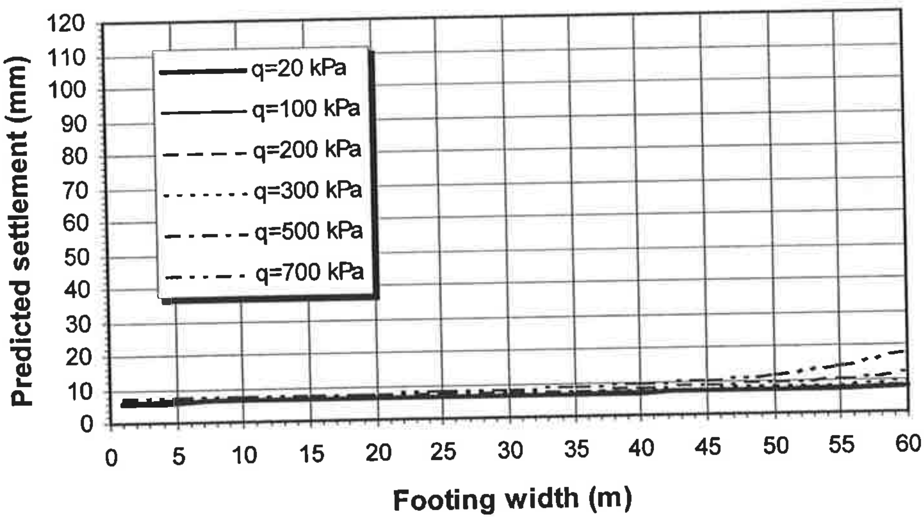
L/B=2.0, D_f/B=0.0 and N=40



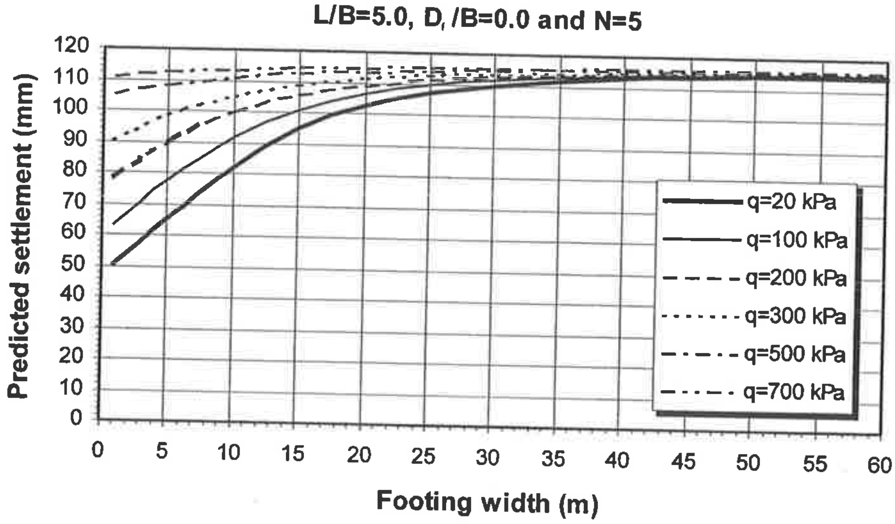
L/B=2.0, D_f/B=0.0 and N=50

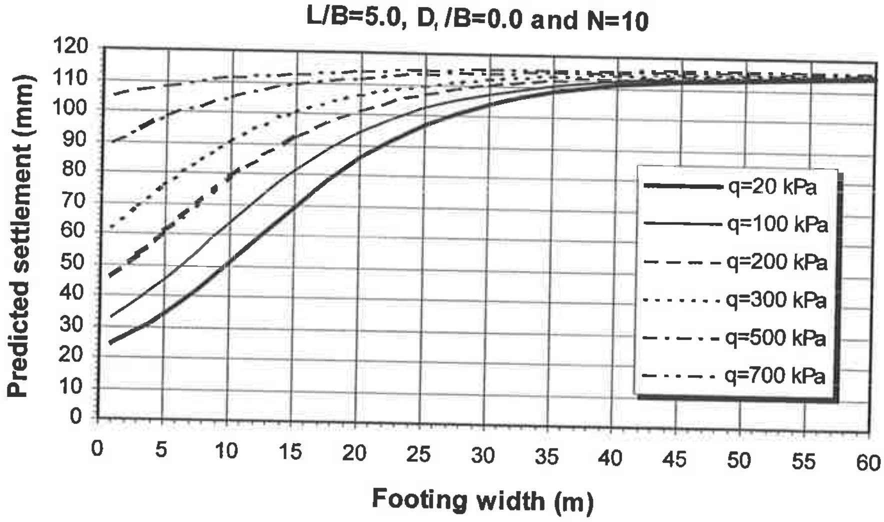


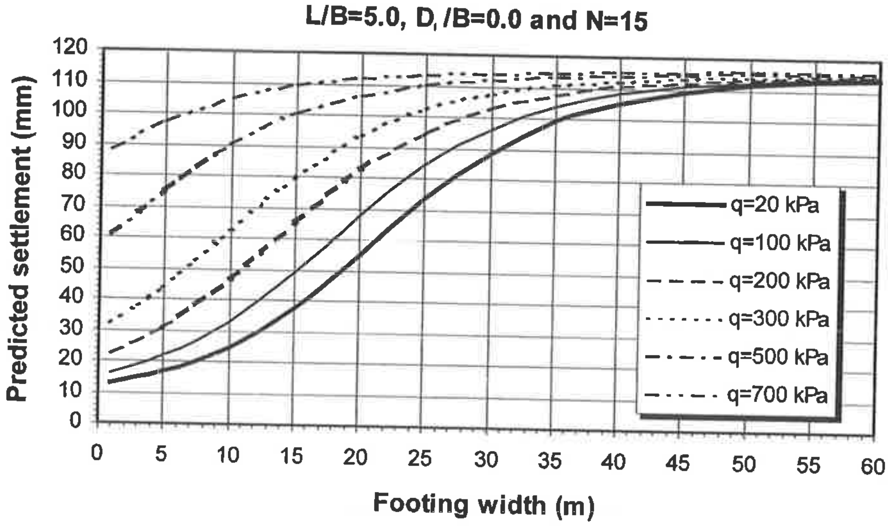
L/B=2.0, D_f/B=0.0 and N=60

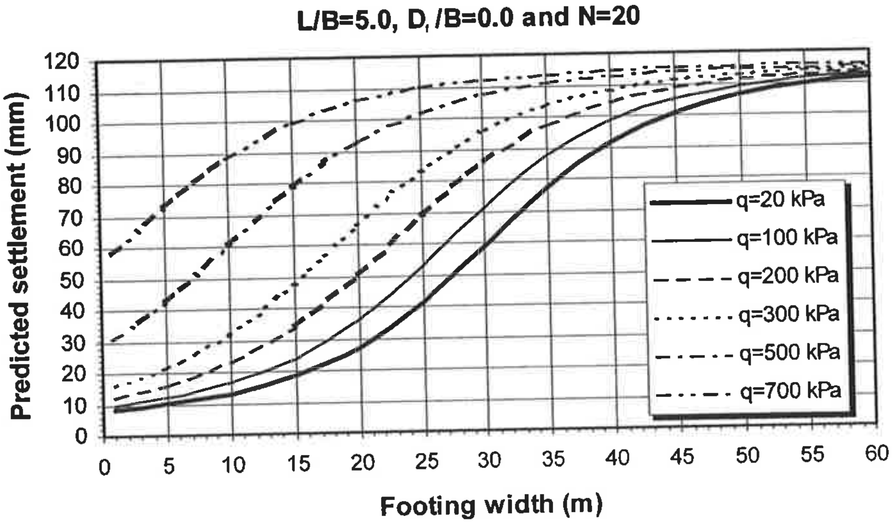


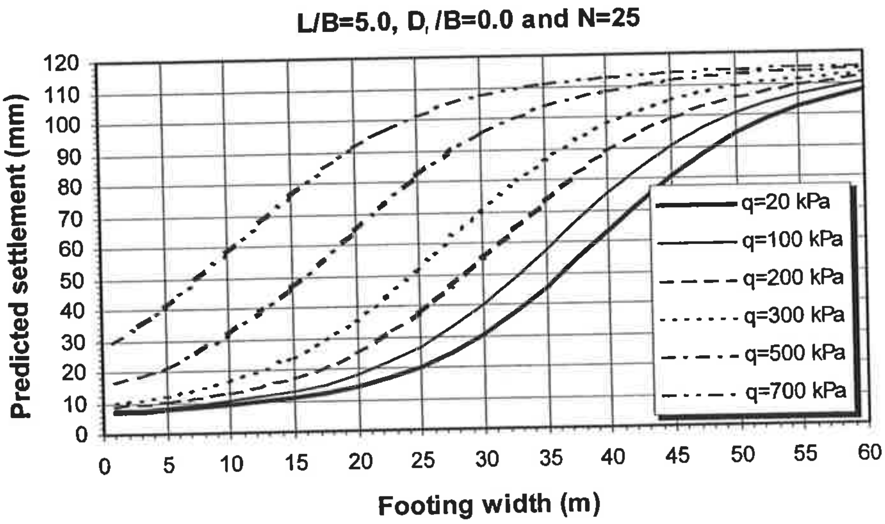
- *L/B* = 5.0, D_f/B = 0.0

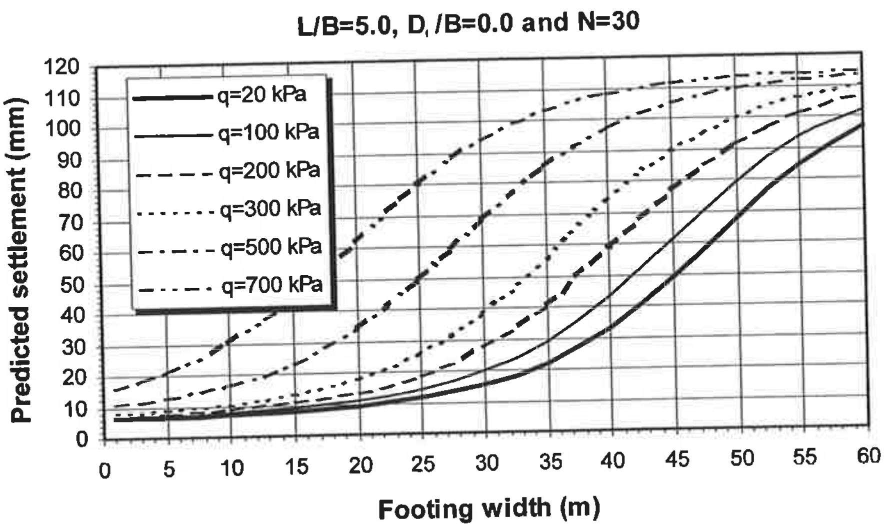

L/B=5.0, Df /B=0.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

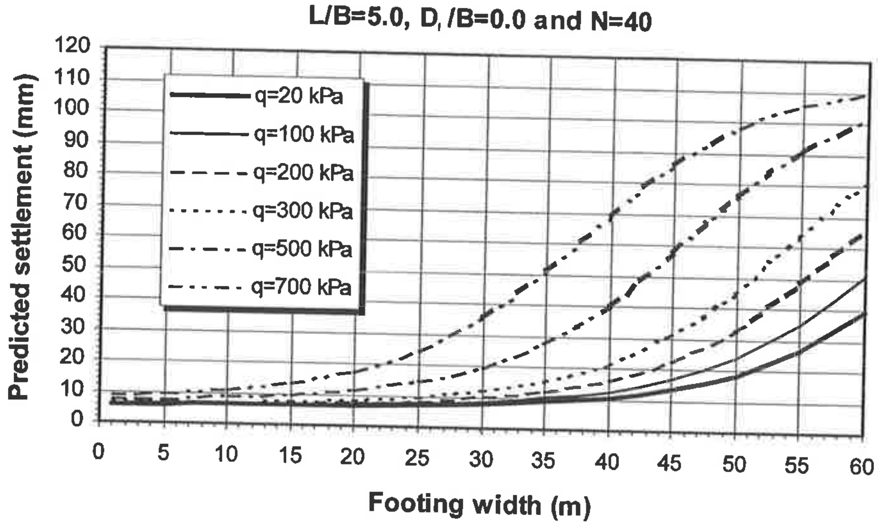

L/B=5.0, Df /B=0.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

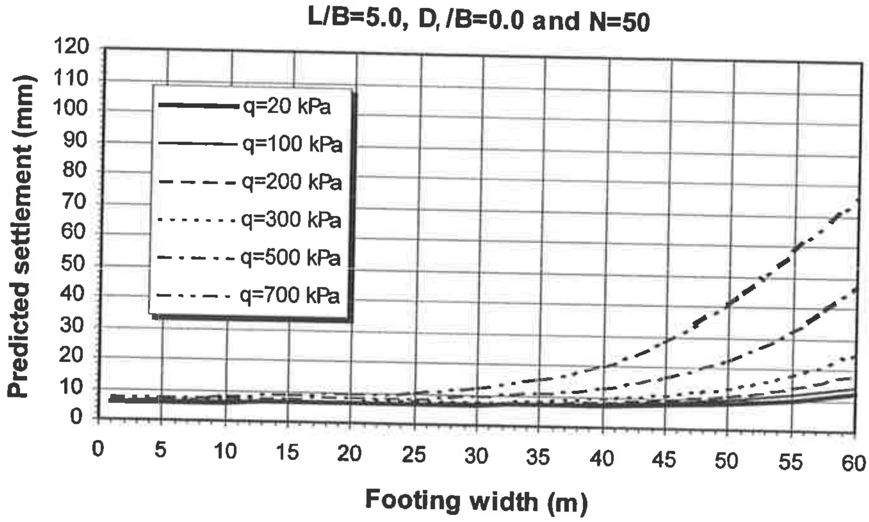

L/B=5.0, Df /B=0.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

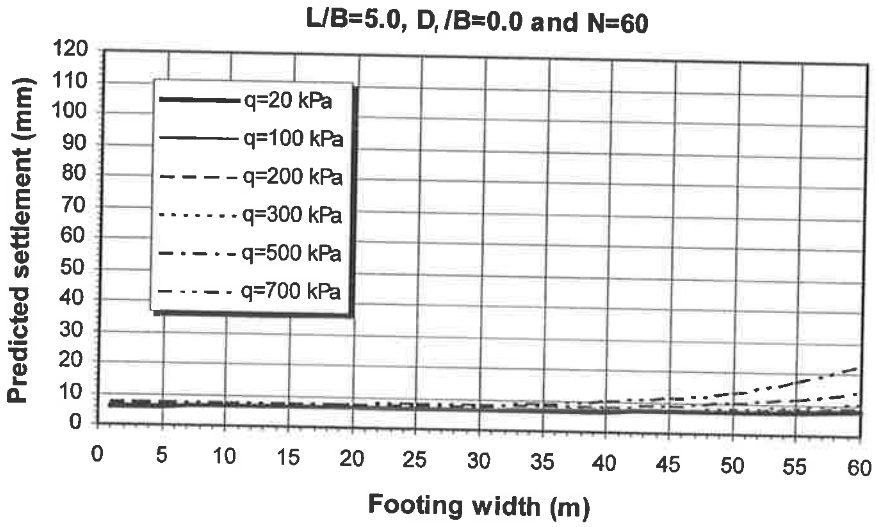

L/B=5.0, D_f/B=0.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=5.0, D_f/B=0.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

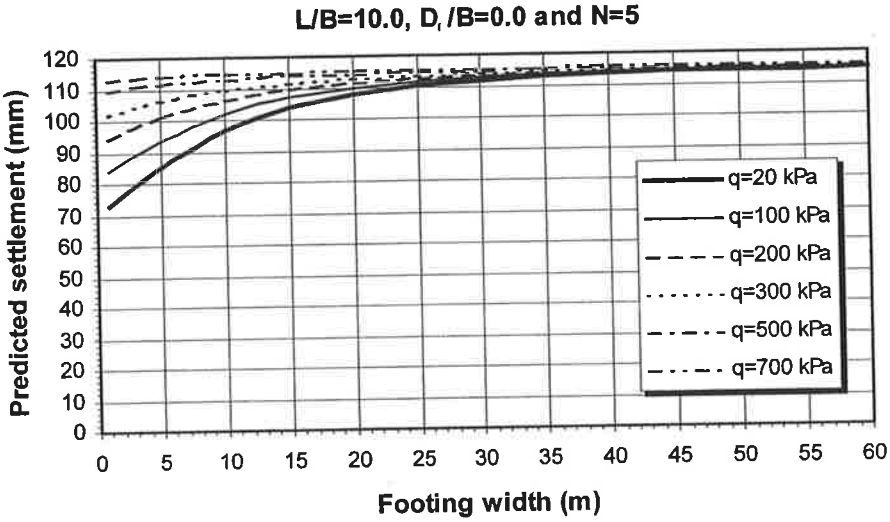

L/B=5.0, D_f/B=0.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

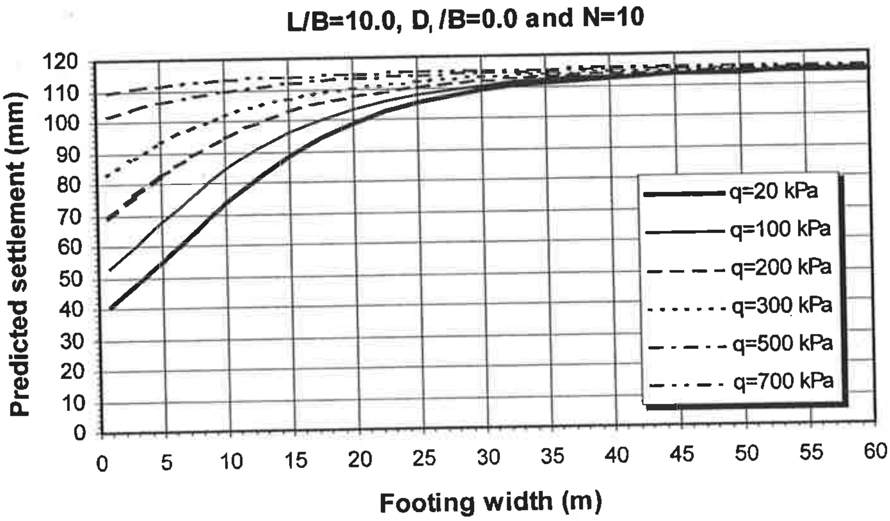

L/B=5.0, D_f/B=0.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

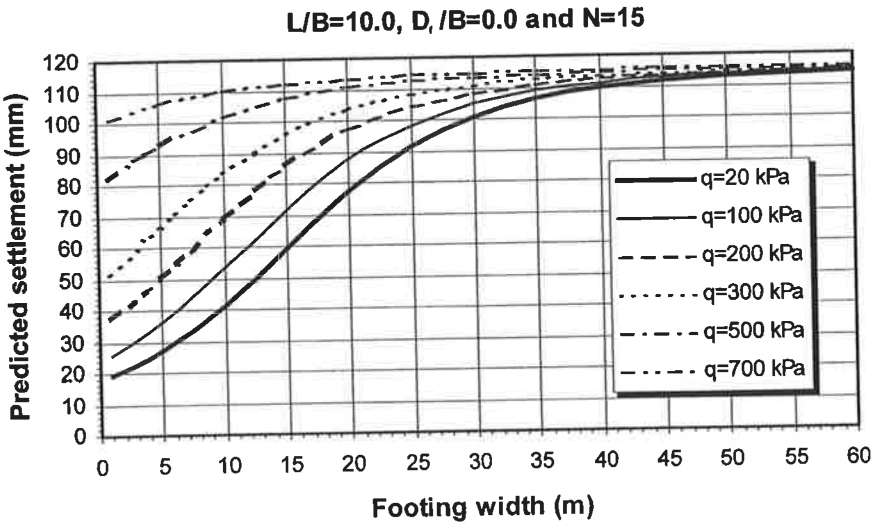

L/B=5.0, D_f/B=0.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

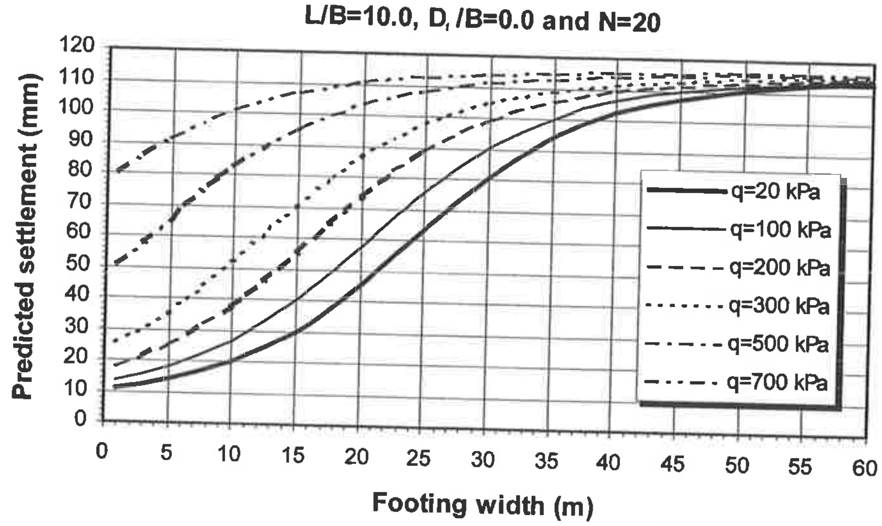

L/B=5.0, D_f/B=0.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

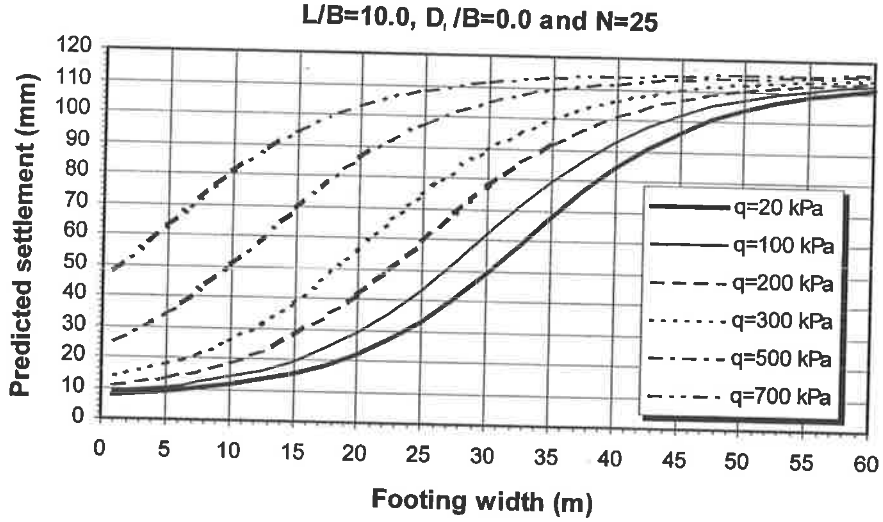
- $L/B = 10.0, D_f/B = 0.0$

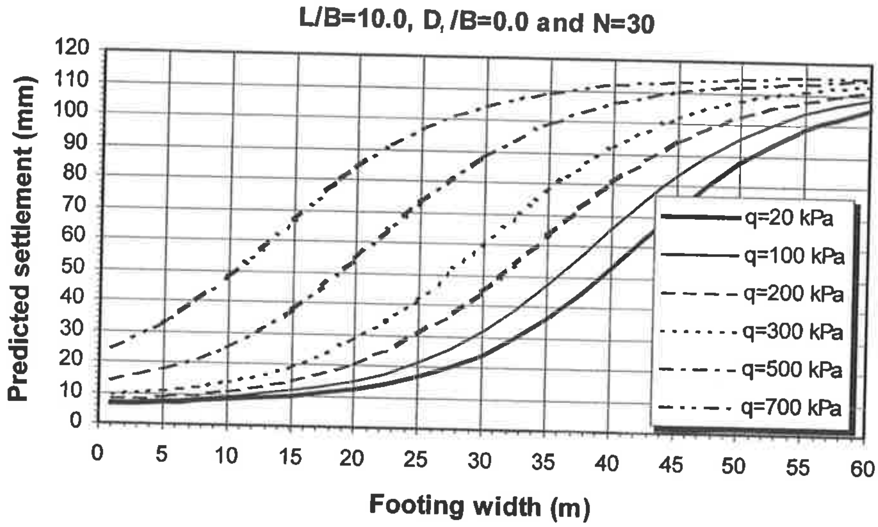

L/B=10.0, Df /B=0.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

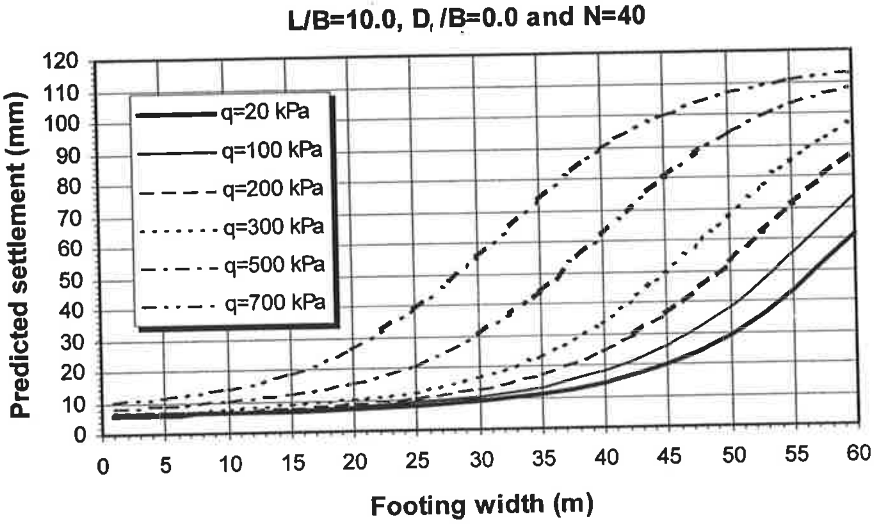

L/B=10.0, Df /B=0.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

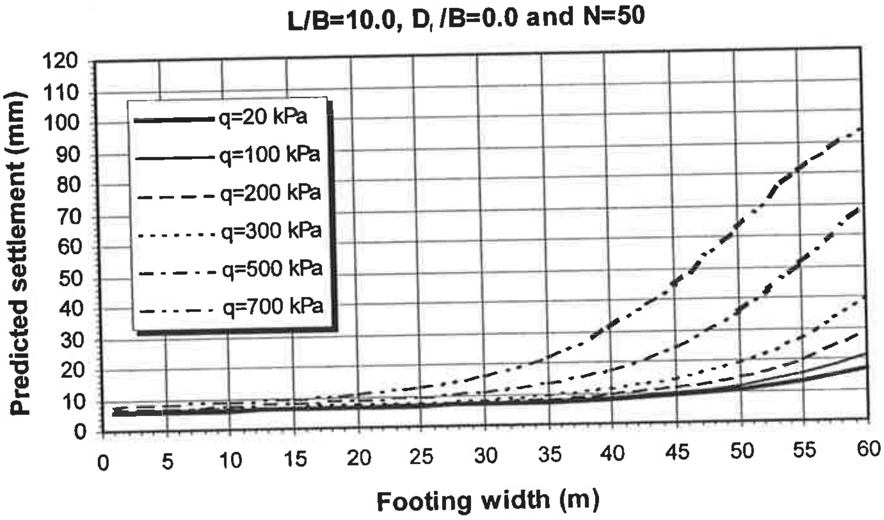

L/B=10.0, Df /B=0.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

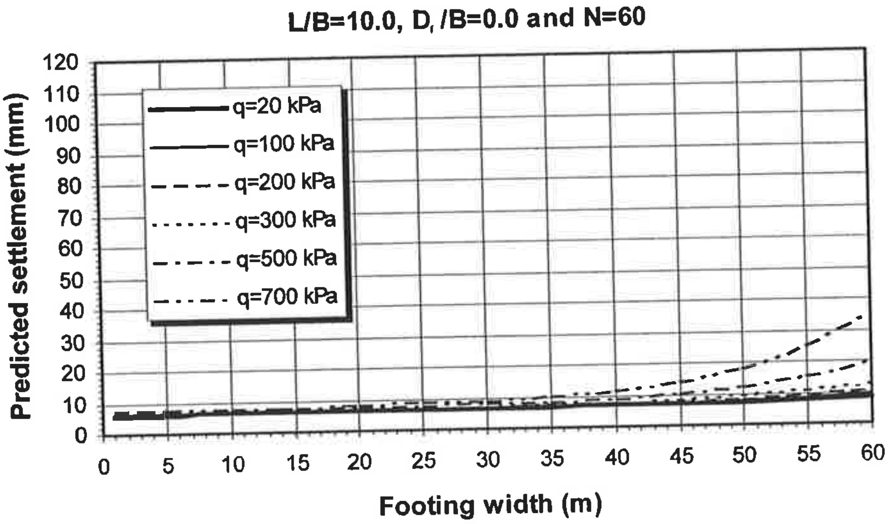

L/B=10.0, Df /B=0.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=10.0, Df /B=0.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

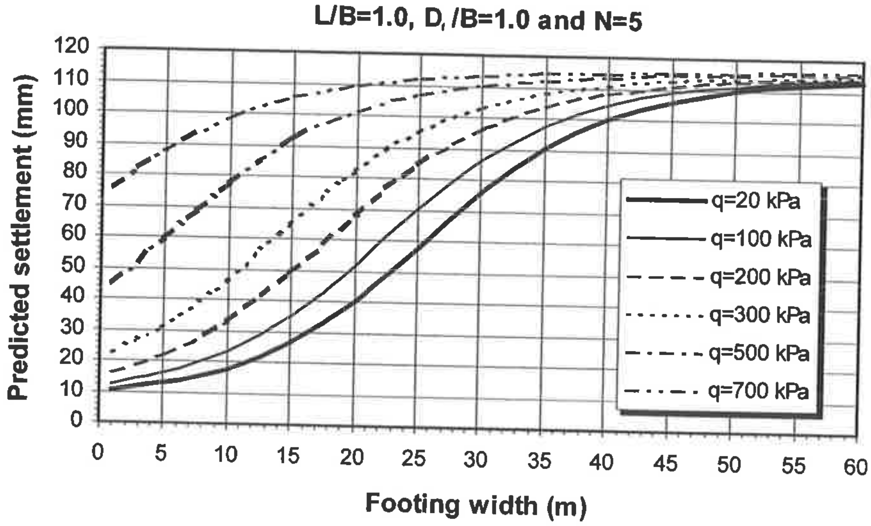

L/B=10.0, Df /B=0.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

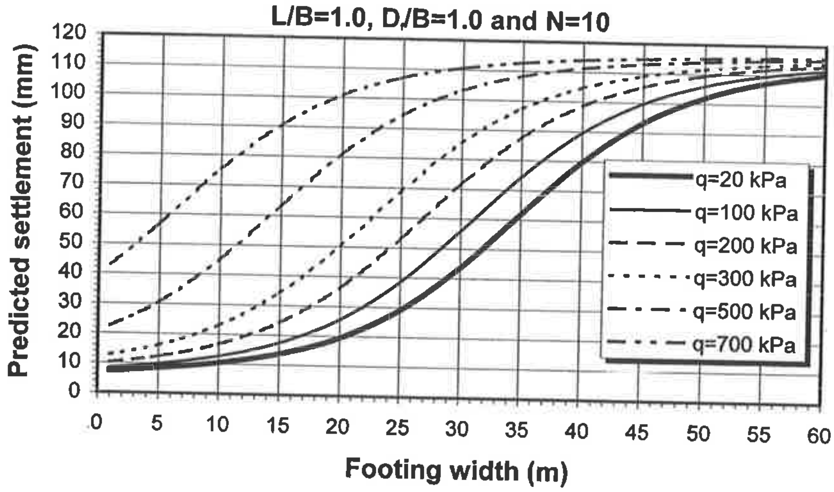

L/B=10.0, D_f/B=0.0 and N=40
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

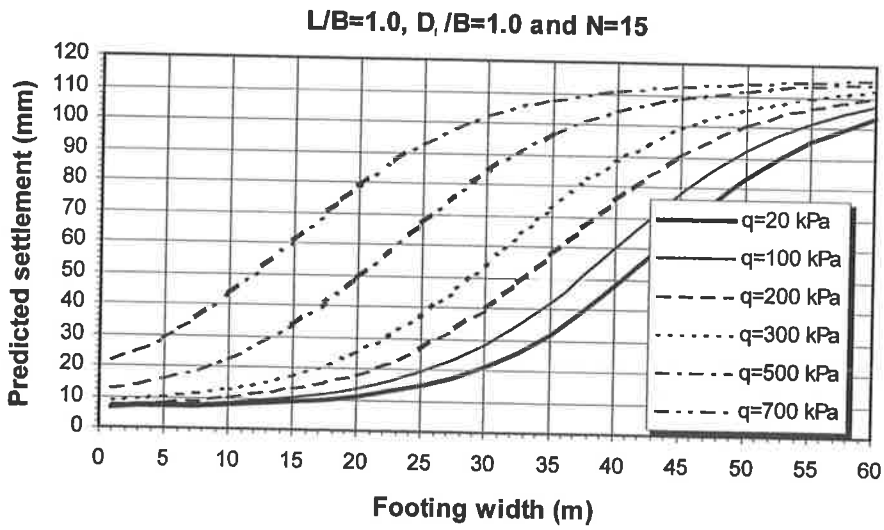

L/B=10.0, D_f/B=0.0 and N=50
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)


L/B=10.0, D_f/B=0.0 and N=60
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

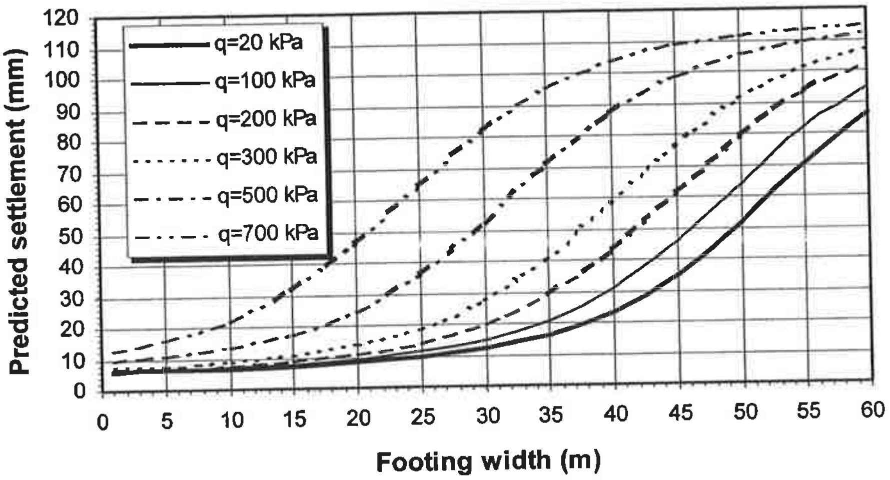
- ***L/B* = 1.0, *D$_f$/B* = 1.0**


L/B=1.0, D$_f$ /B=1.0 and N=5
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

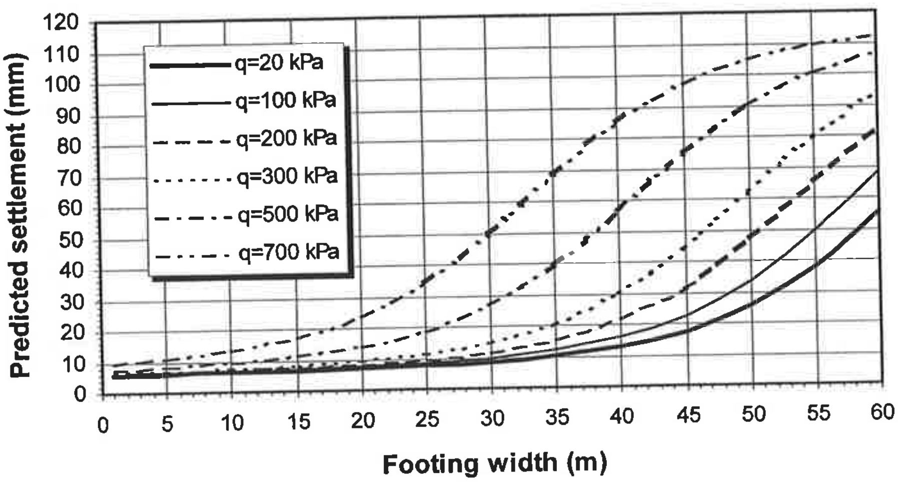

L/B=1.0, D$_f$/B=1.0 and N=10
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=1.0, D$_f$ /B=1.0 and N=15
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

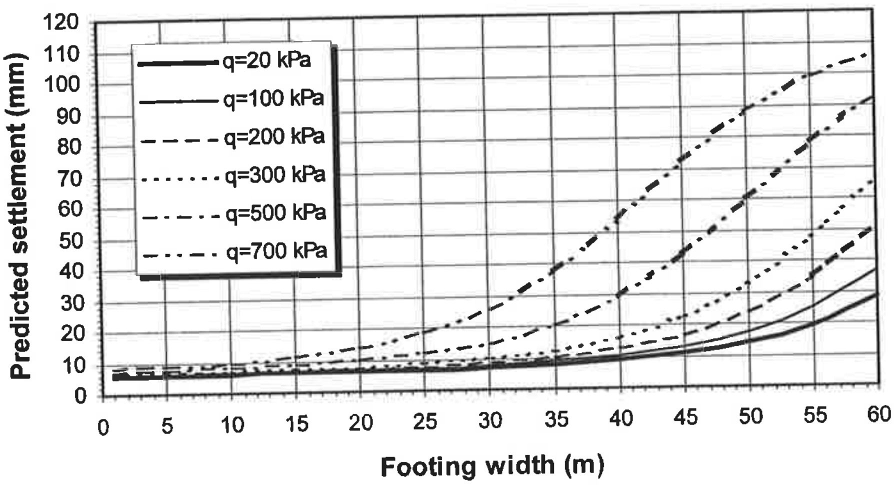
L/B=1.0, D_f/B=1.0 and N=20

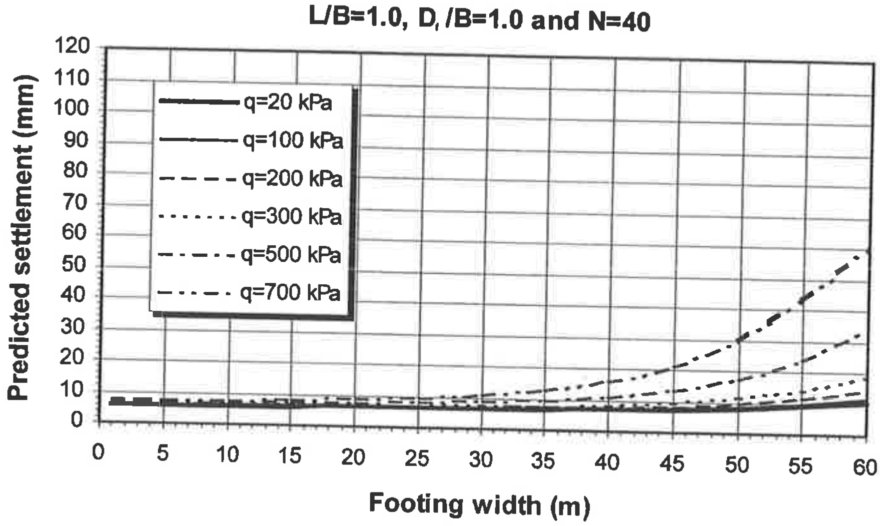


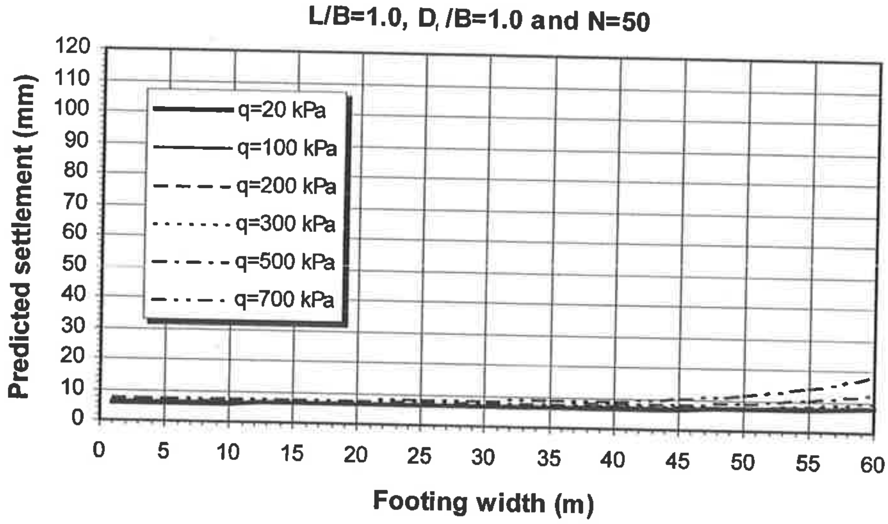
L/B=1.0, D_f/B=1.0 and N=25

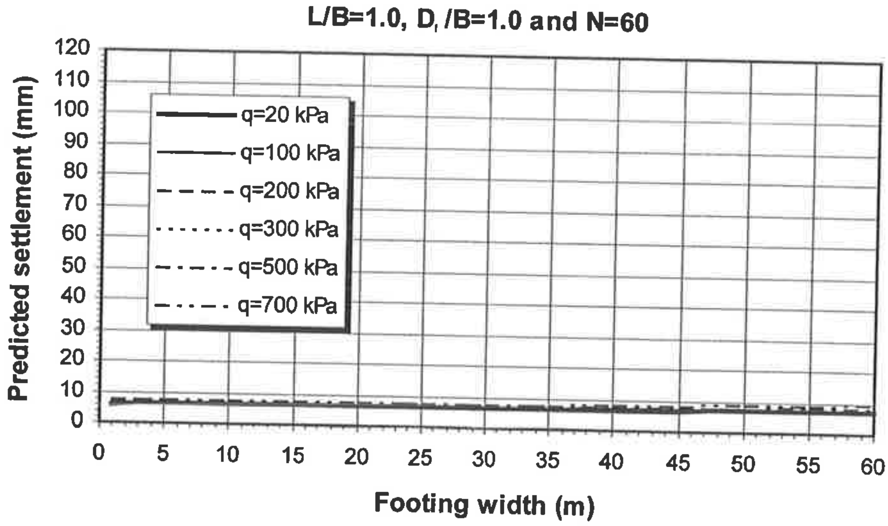


L/B=1.0, D_f/B=1.0 and N=30

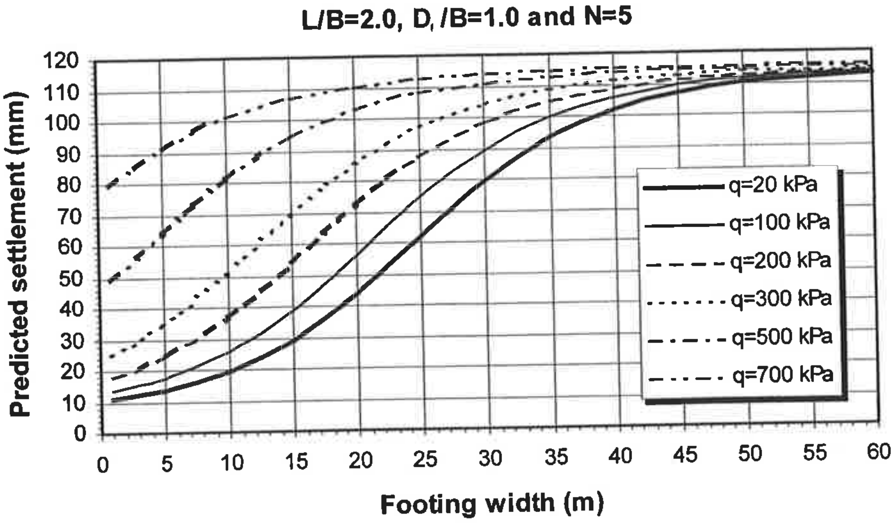


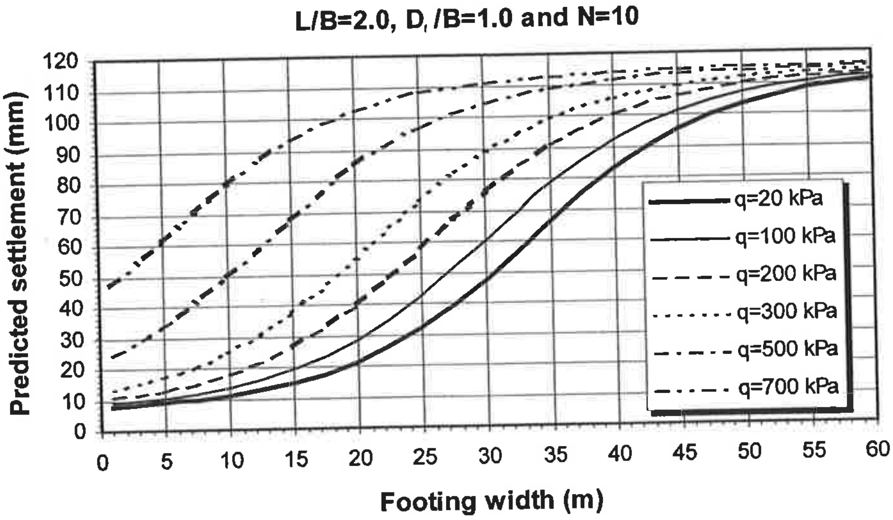

L/B=1.0, D_f/B=1.0 and N=40
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

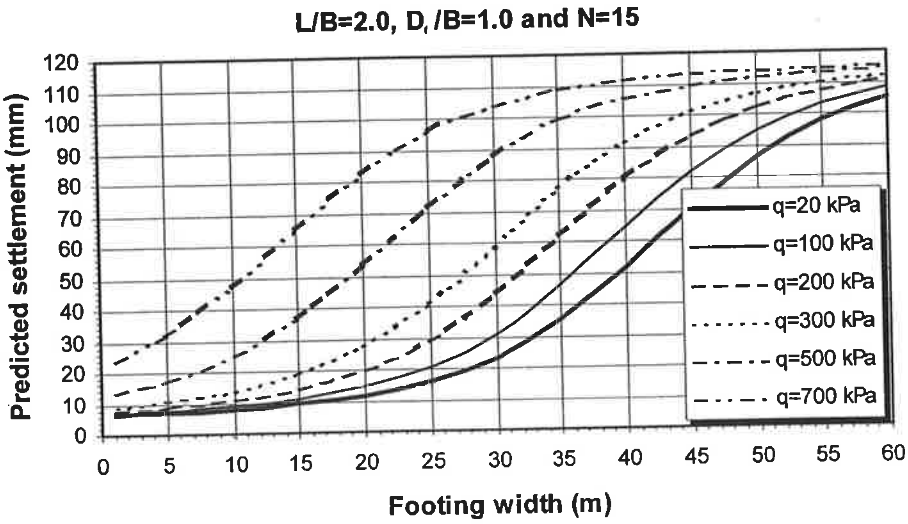

L/B=1.0, D_f/B=1.0 and N=50
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

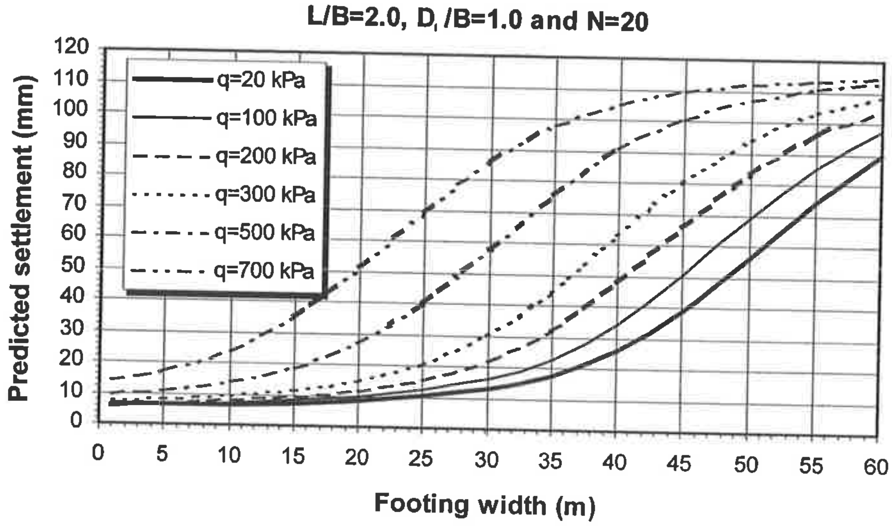

L/B=1.0, D_f/B=1.0 and N=60
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

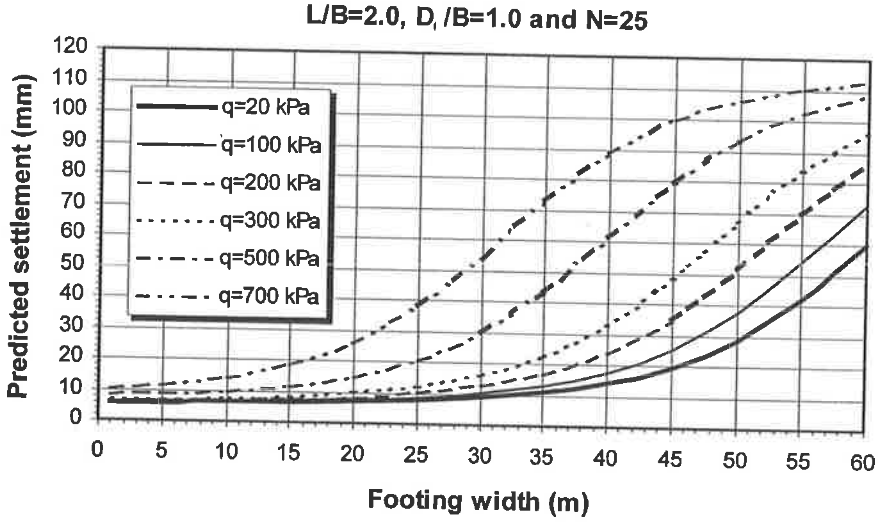
- $L/B = 2.0, D_f/B = 1.0$

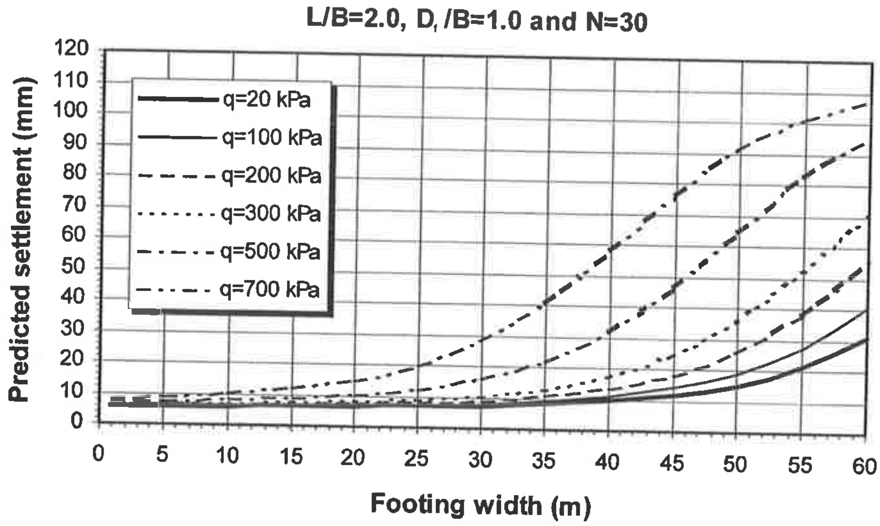

L/B=2.0, D_f /B=1.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

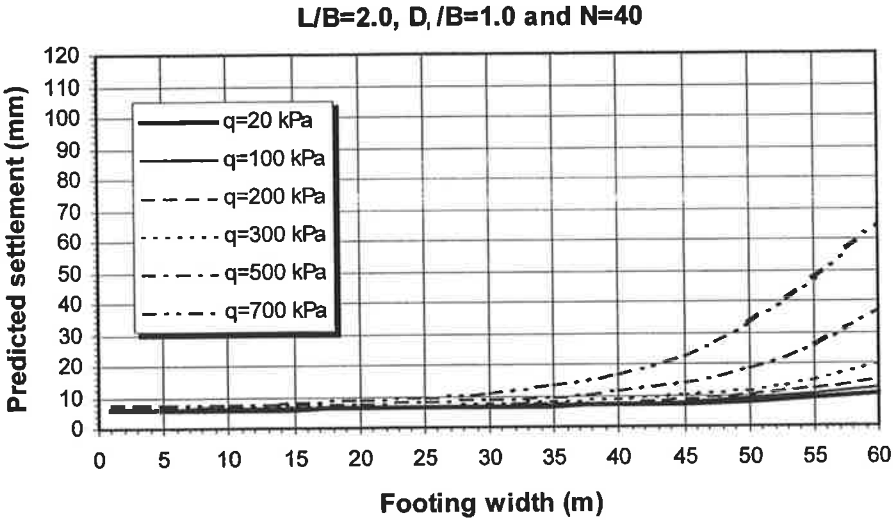

L/B=2.0, D_f /B=1.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

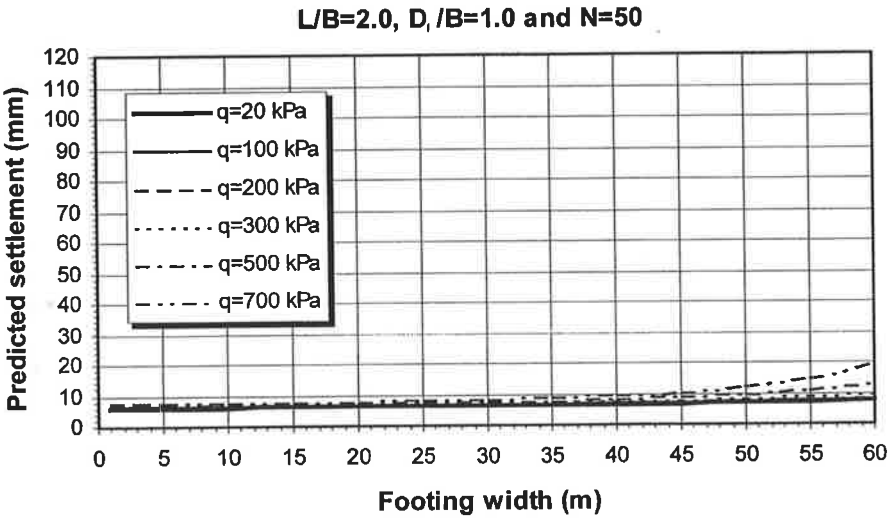

L/B=2.0, D_f /B=1.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

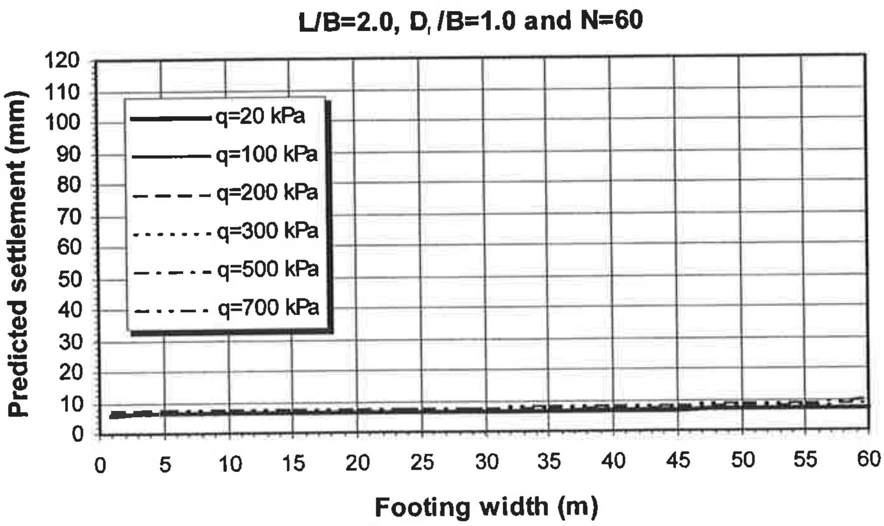

L/B=2.0, D_f/B=1.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D_f/B=1.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

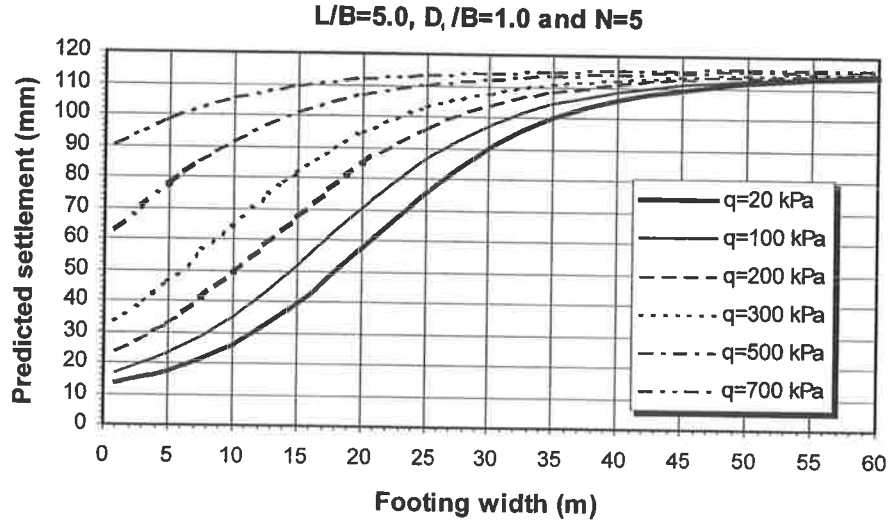

L/B=2.0, D_f/B=1.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

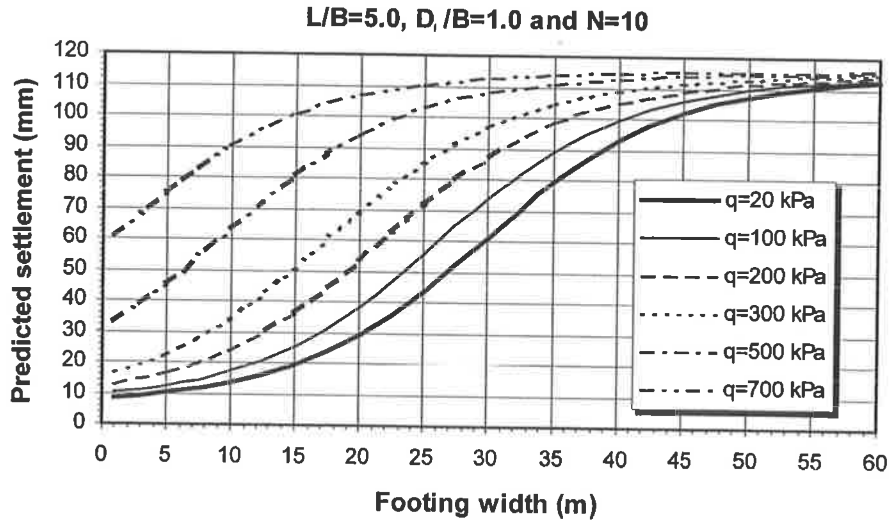

L/B=2.0, D_f/B=1.0 and N=40
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

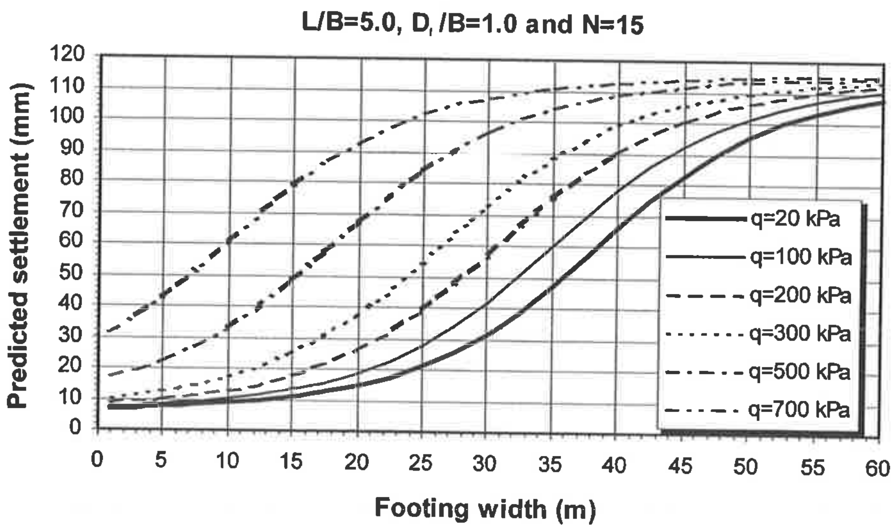

L/B=2.0, D_f/B=1.0 and N=50
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

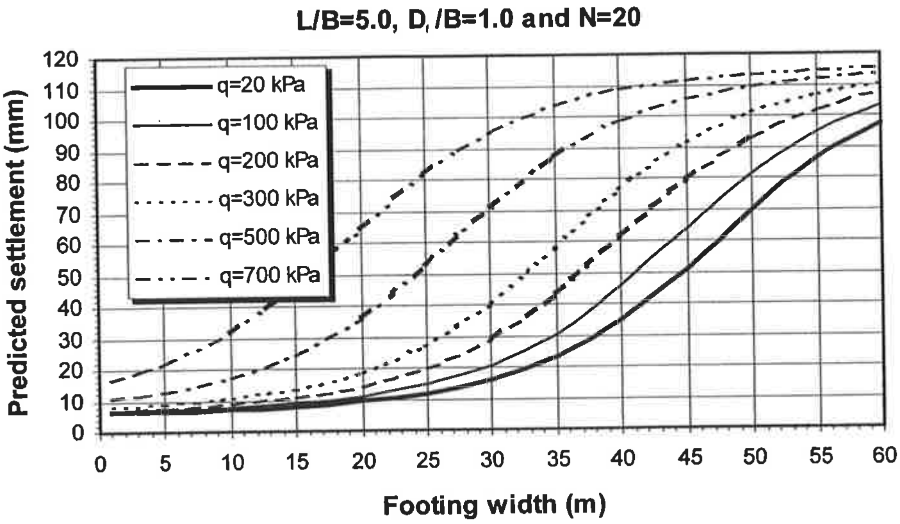

L/B=2.0, D_f/B=1.0 and N=60
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

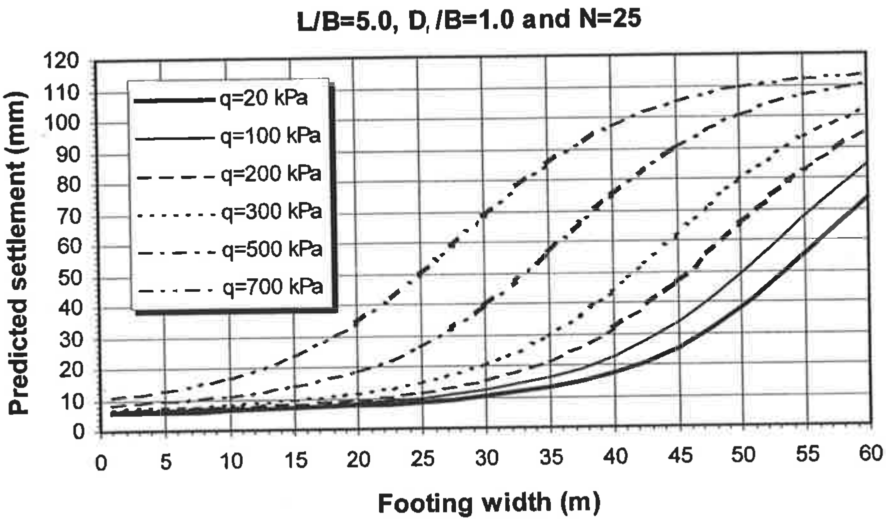
- $L/B = 5.0, D_f/B = 1.0$

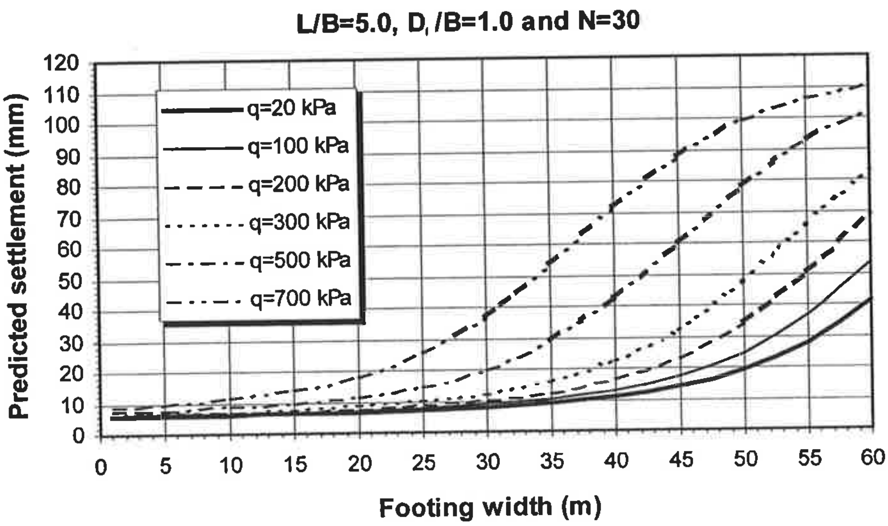

L/B=5.0, D_f /B=1.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=5.0, D_f /B=1.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

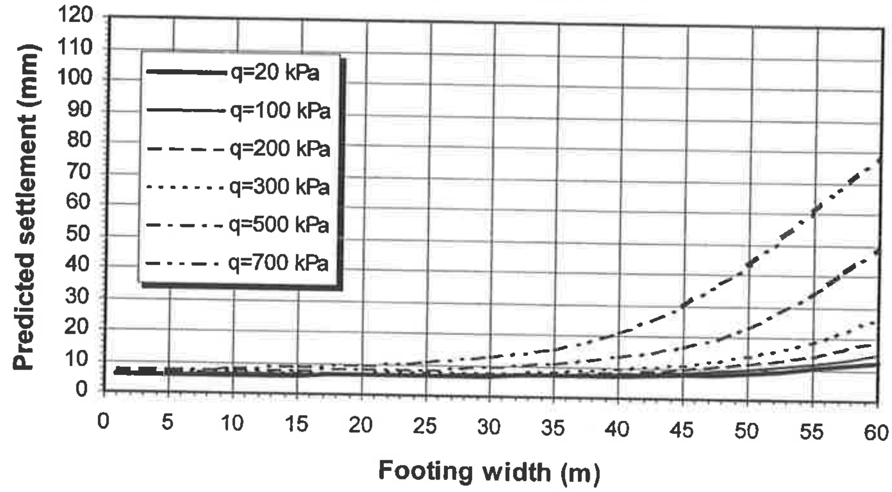

L/B=5.0, D_f /B=1.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=5.0, D_f/B=1.0 and N=20
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
Predicted settlement (mm)
Footing width (m)

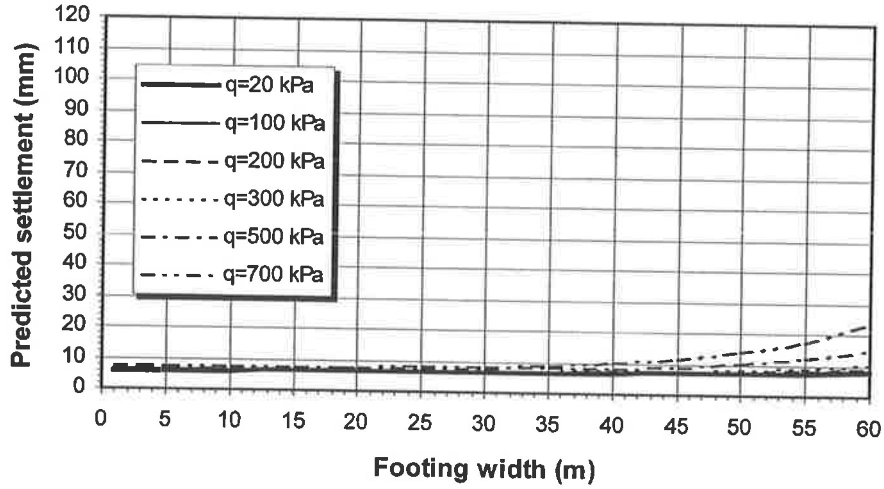

L/B=5.0, D_f/B=1.0 and N=25
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
Predicted settlement (mm)
Footing width (m)


L/B=5.0, D_f/B=1.0 and N=30
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
Predicted settlement (mm)
Footing width (m)

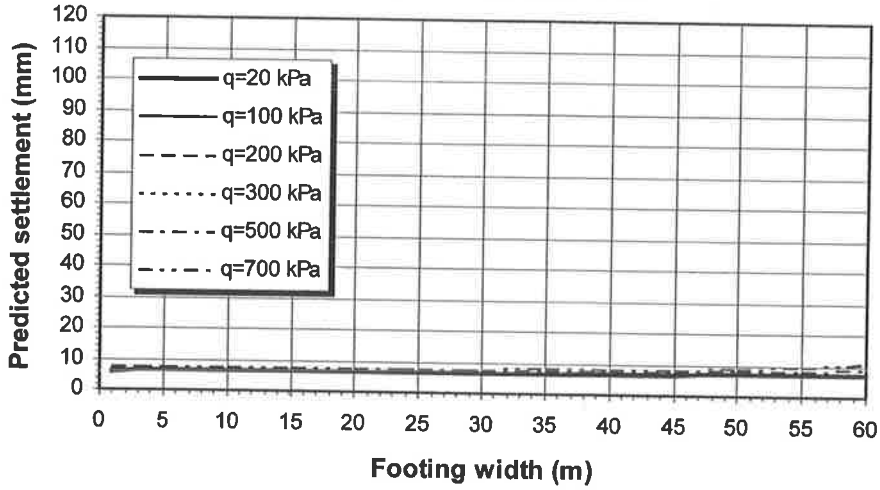
L/B=5.0, D_f/B=1.0 and N=40



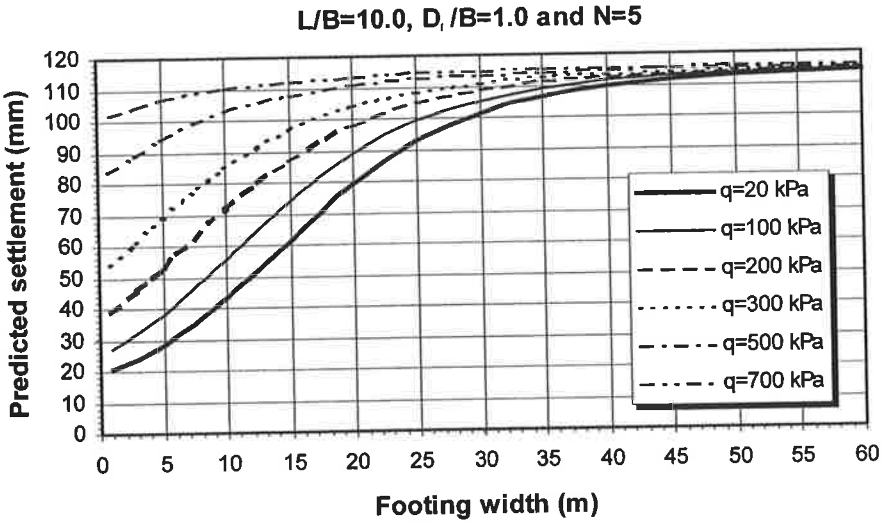
L/B=5.0, D_f/B=1.0 and N=50

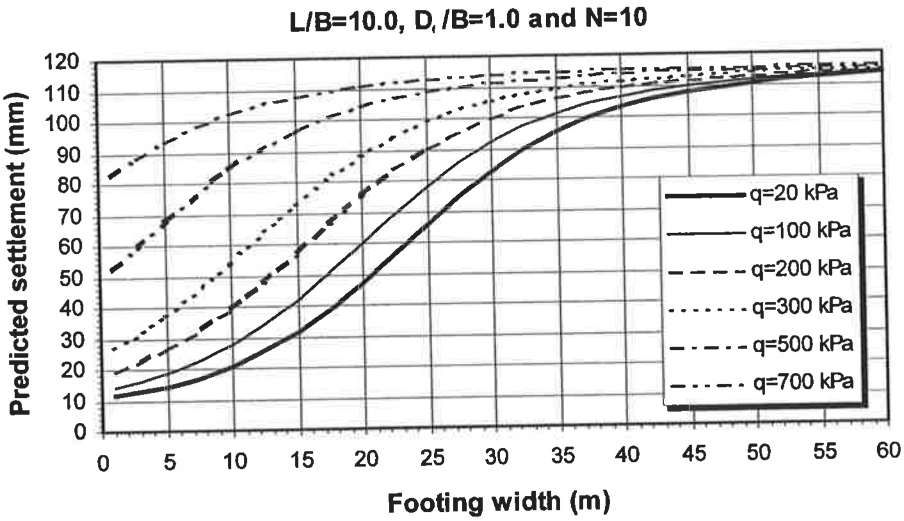


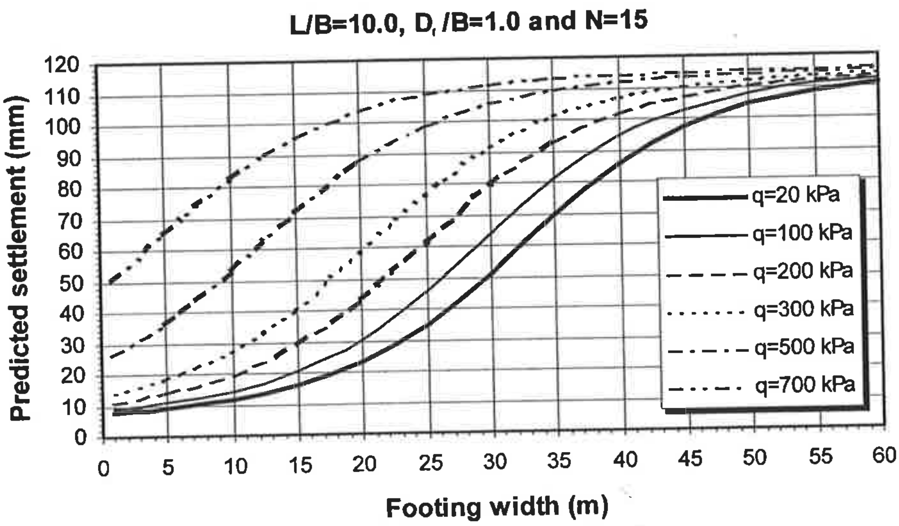
L/B=5.0, D_f/B=1.0 and N=60



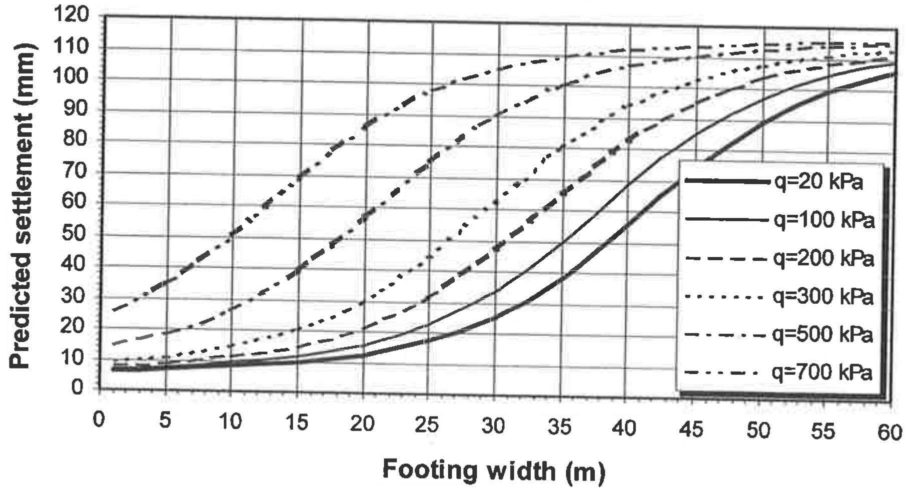
- ***L/B* = 10.0, D_f/*B* = 1.0**


L/B=10.0, D_f/B=1.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

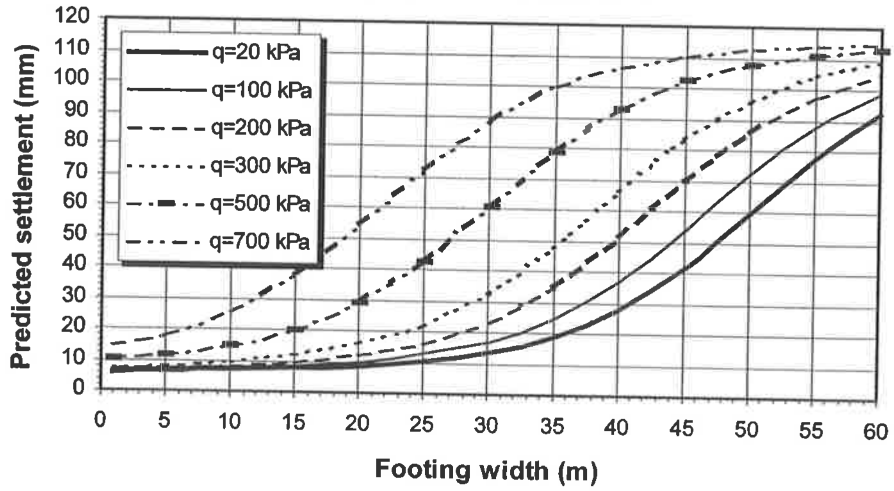

L/B=10.0, D_f/B=1.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=10.0, D_f/B=1.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

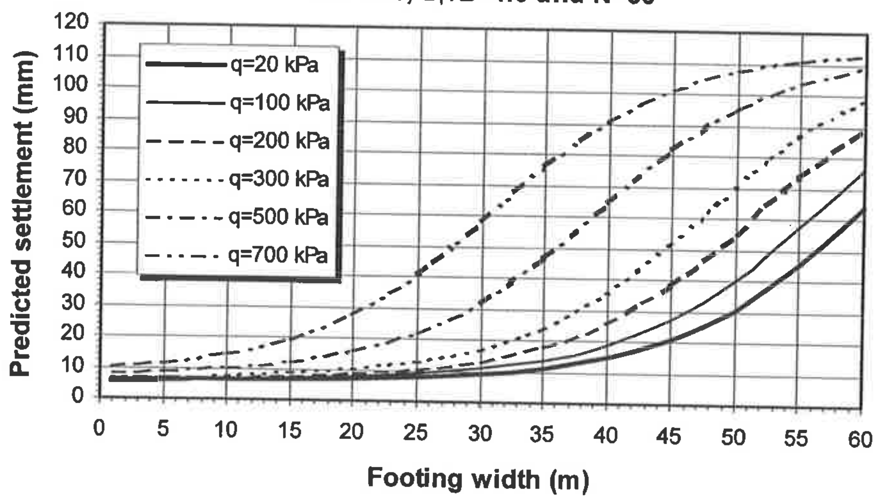
L/B=10.0, D_f/B=1.0 and N=20

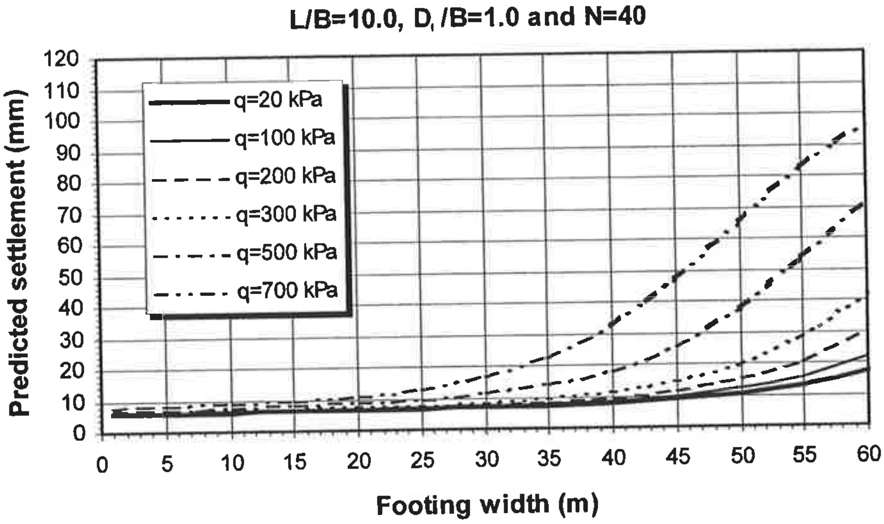


L/B=10.0, D_f/B=1.0 and N=25

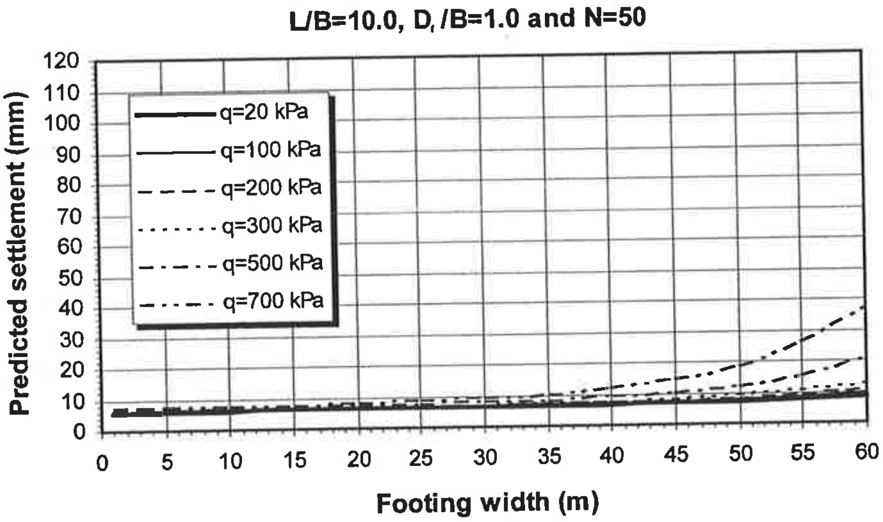


L/B=10.0, D_f/B=1.0 and N=30

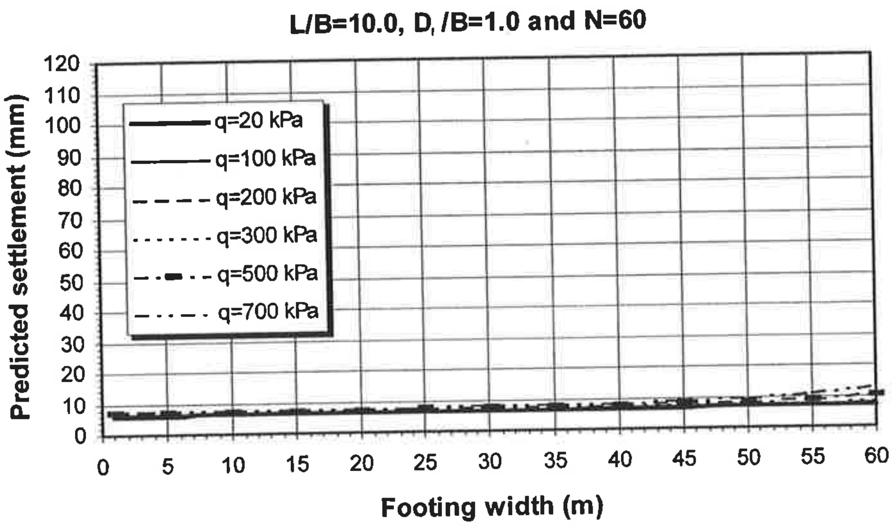




L/B=10.0, D_f/B=1.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

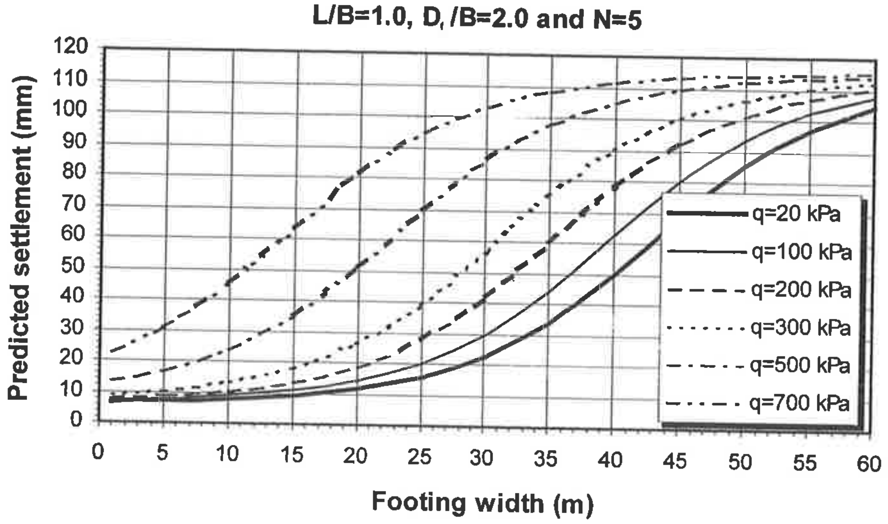


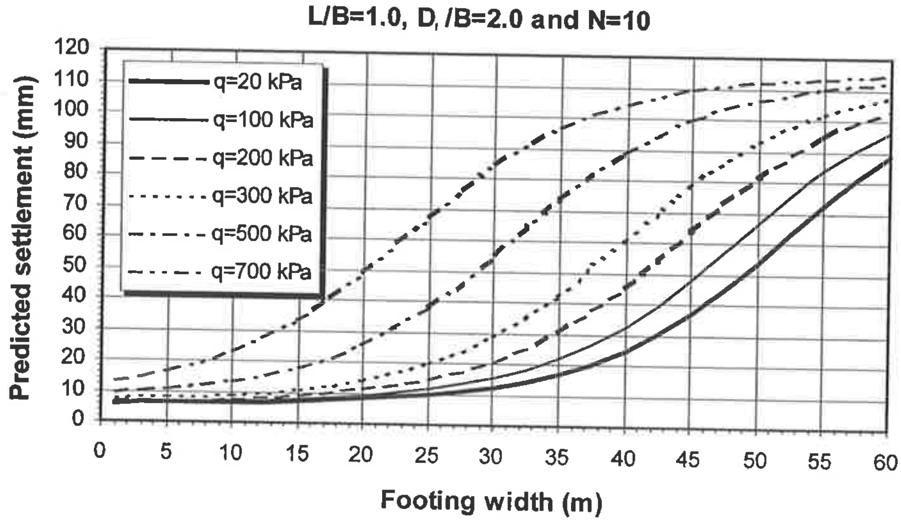
L/B=10.0, D_f/B=1.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

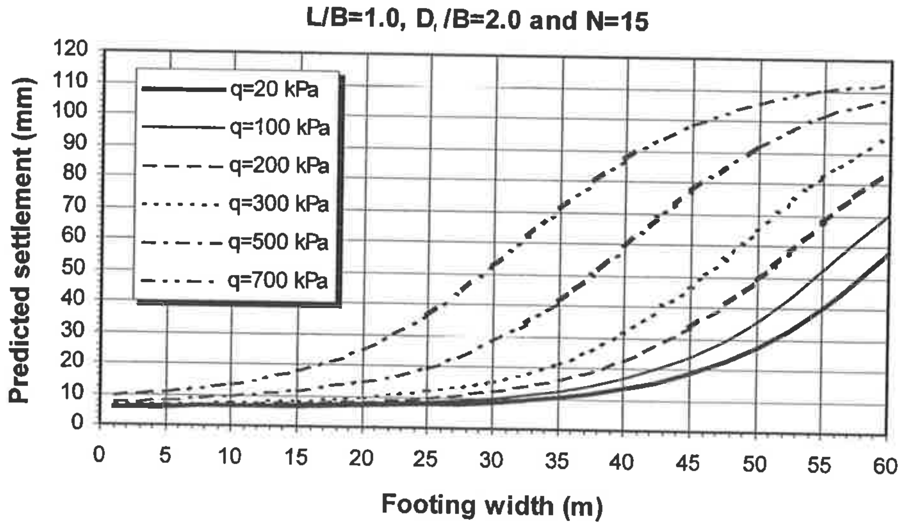


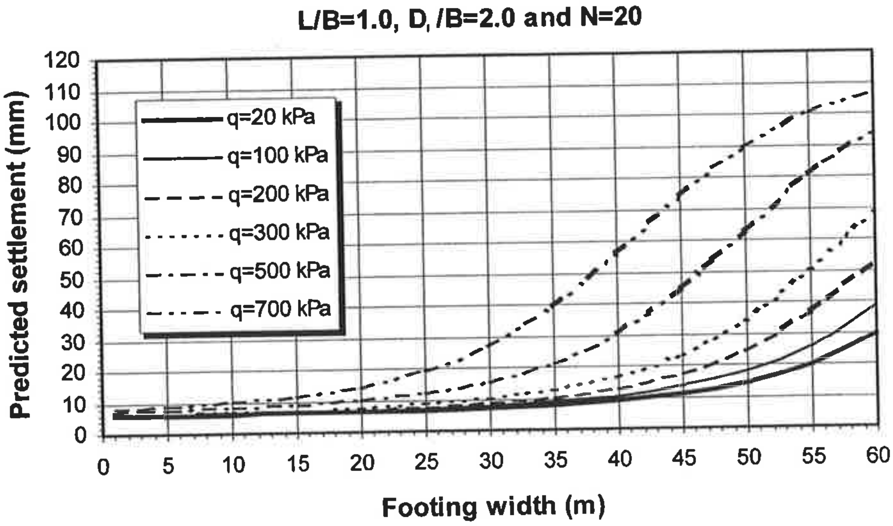
L/B=10.0, D_f/B=1.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

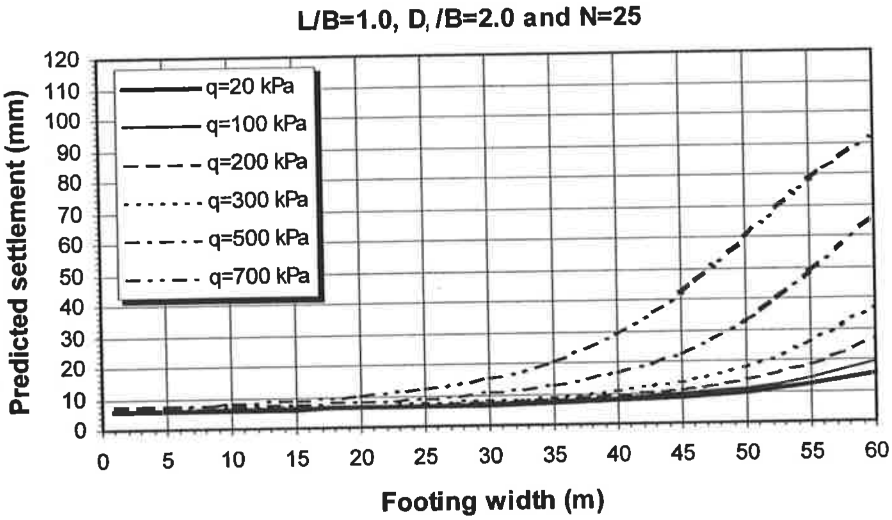
- **$L/B = 1.0$, $D_f/B = 2.0$**

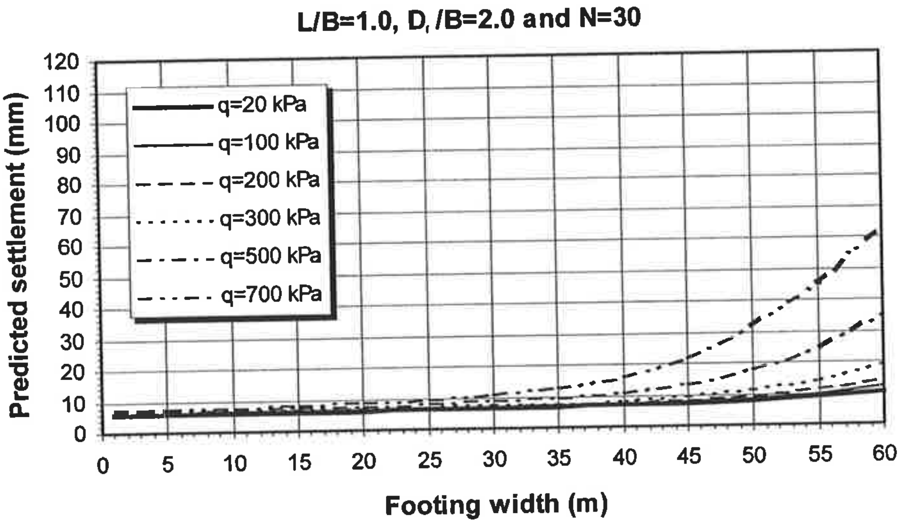

L/B=1.0, Df /B=2.0 and N=5
Predicted settlement (mm)
Footing width (m)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

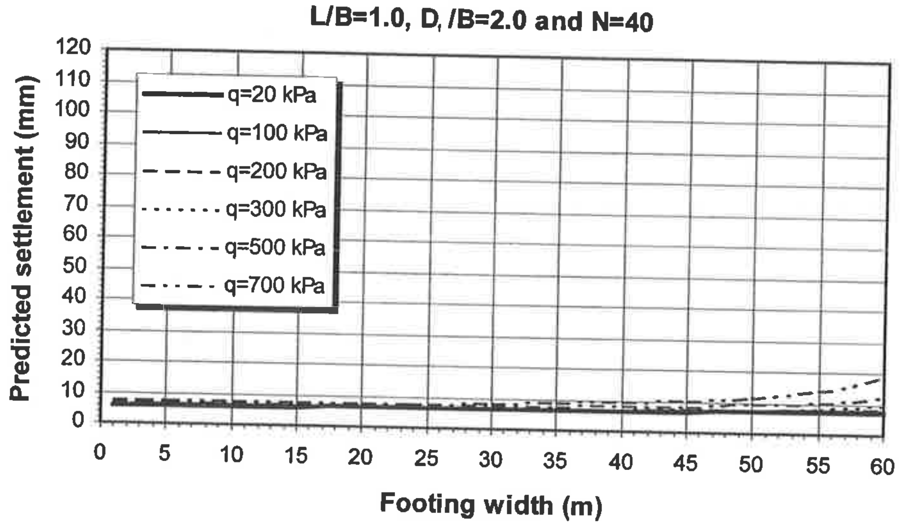

L/B=1.0, Df /B=2.0 and N=10
Predicted settlement (mm)
Footing width (m)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

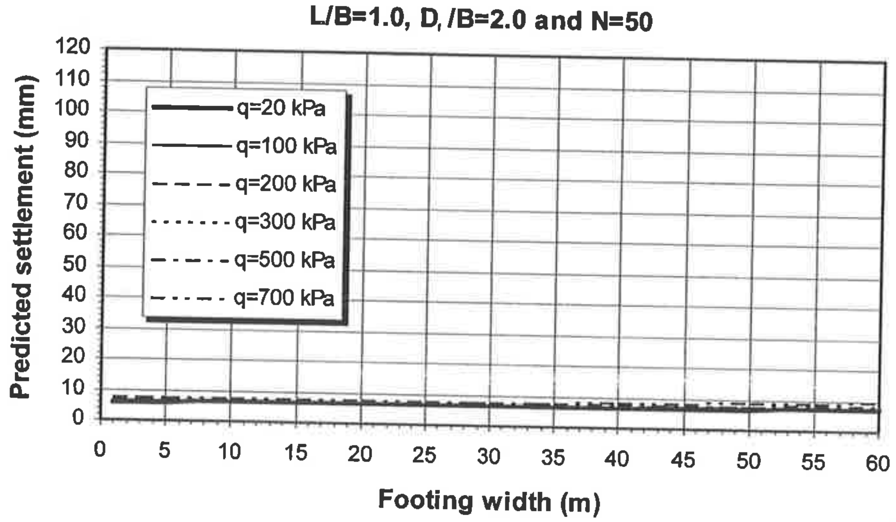

L/B=1.0, Df /B=2.0 and N=15
Predicted settlement (mm)
Footing width (m)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

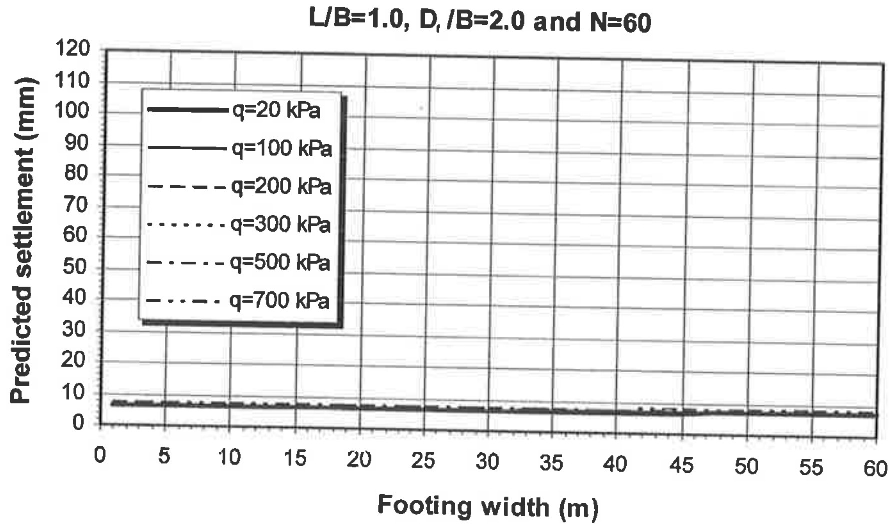

L/B=1.0, Df/B=2.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=1.0, Df/B=2.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

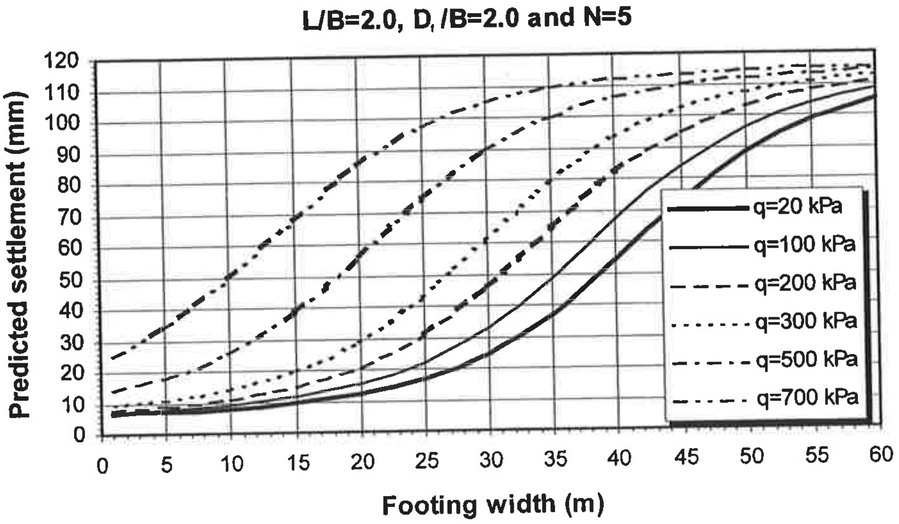

L/B=1.0, Df/B=2.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

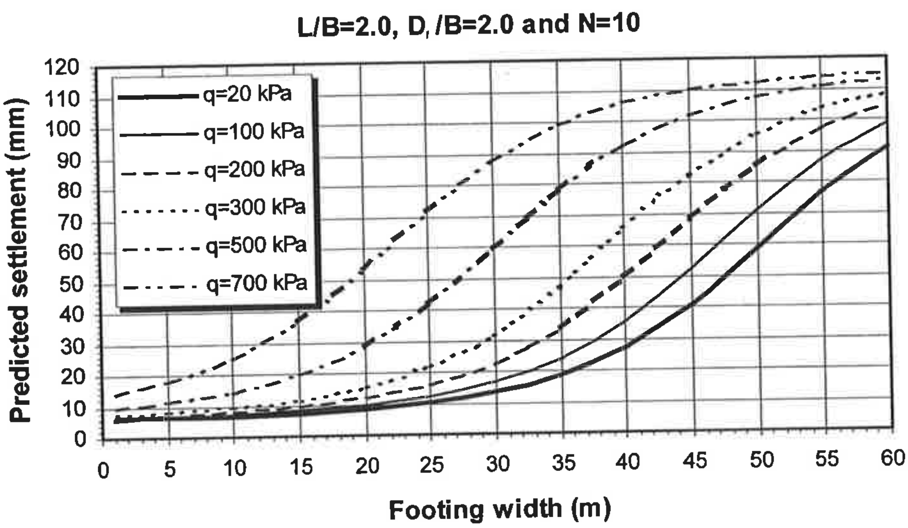

L/B=1.0, D$_f$ /B=2.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

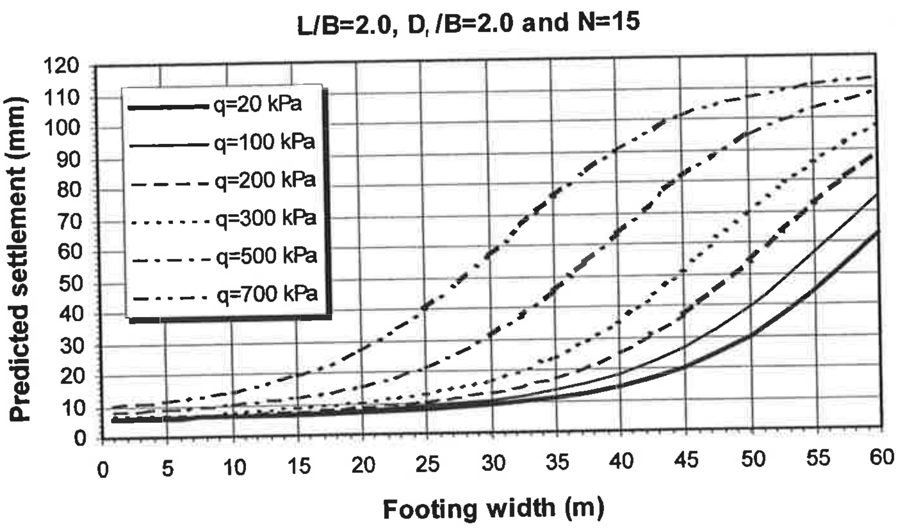

L/B=1.0, D$_f$ /B=2.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

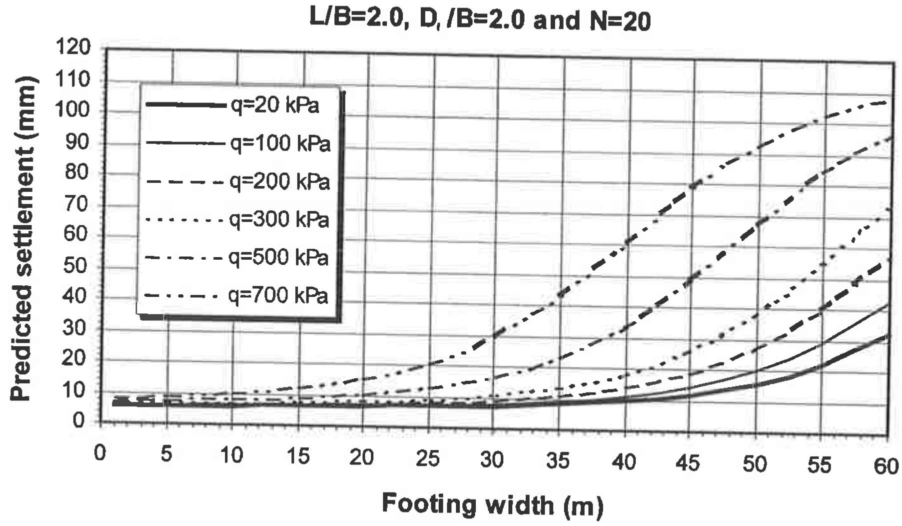

L/B=1.0, D$_f$ /B=2.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

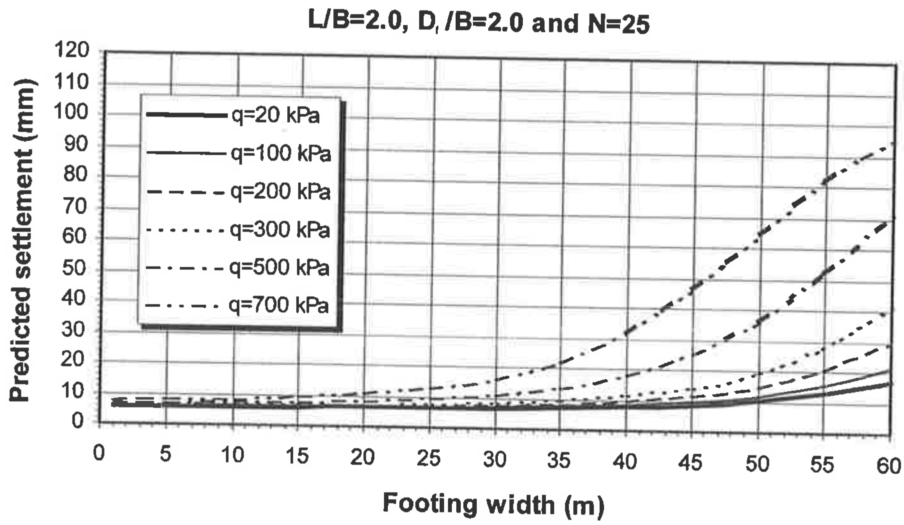
- $L/B = 2.0, D_f/B = 2.0$

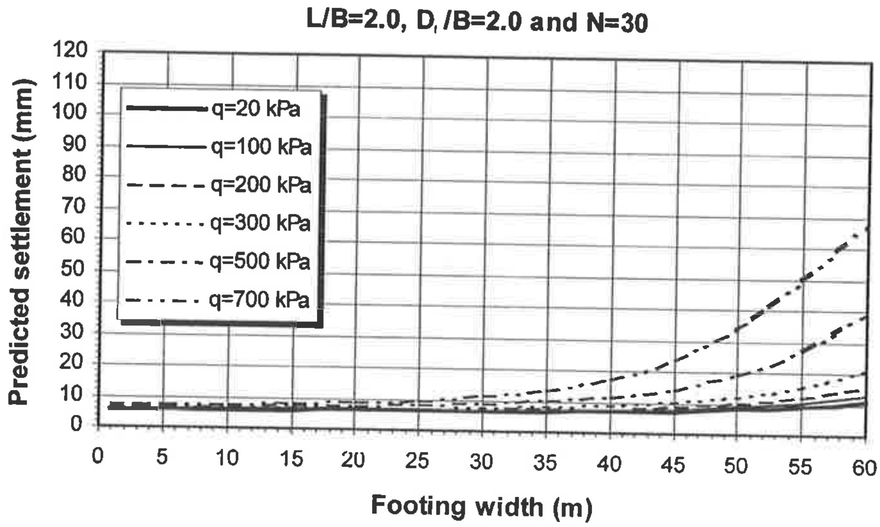

L/B=2.0, Df/B=2.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

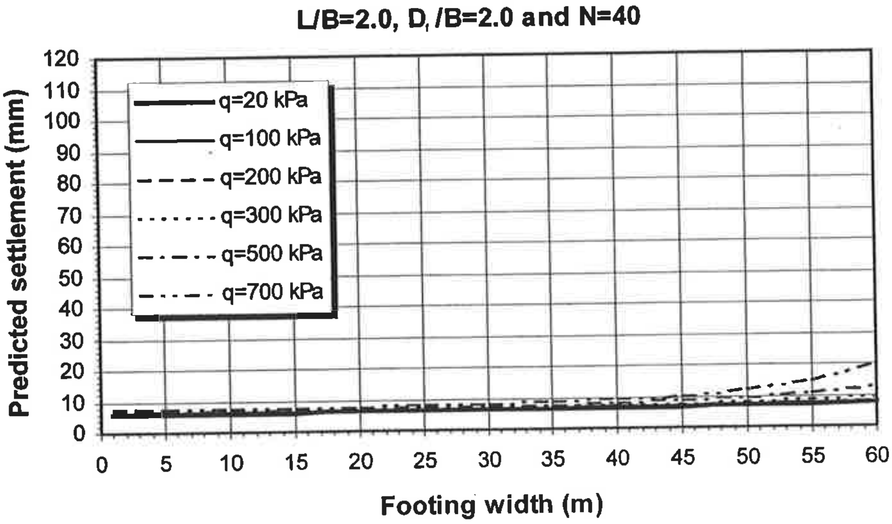

L/B=2.0, Df/B=2.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

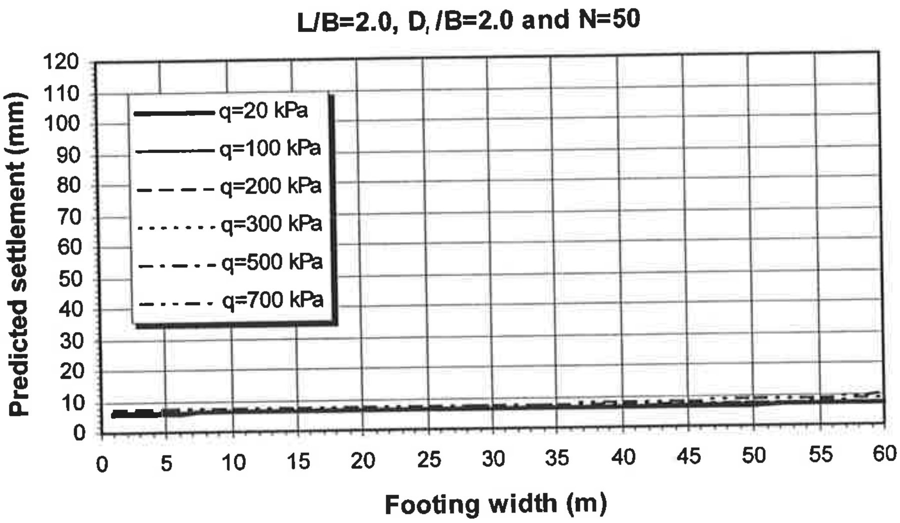

L/B=2.0, Df/B=2.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

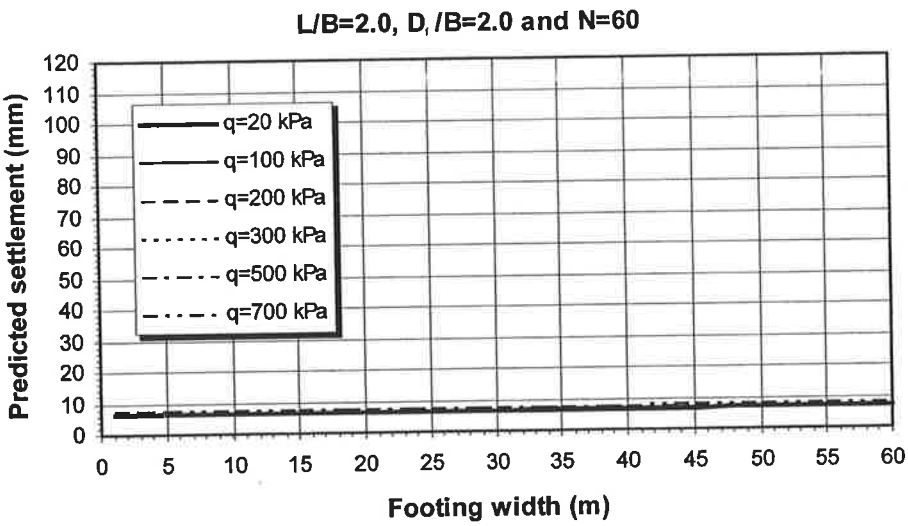

L/B=2.0, D_f/B=2.0 and N=20
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D_f/B=2.0 and N=25
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

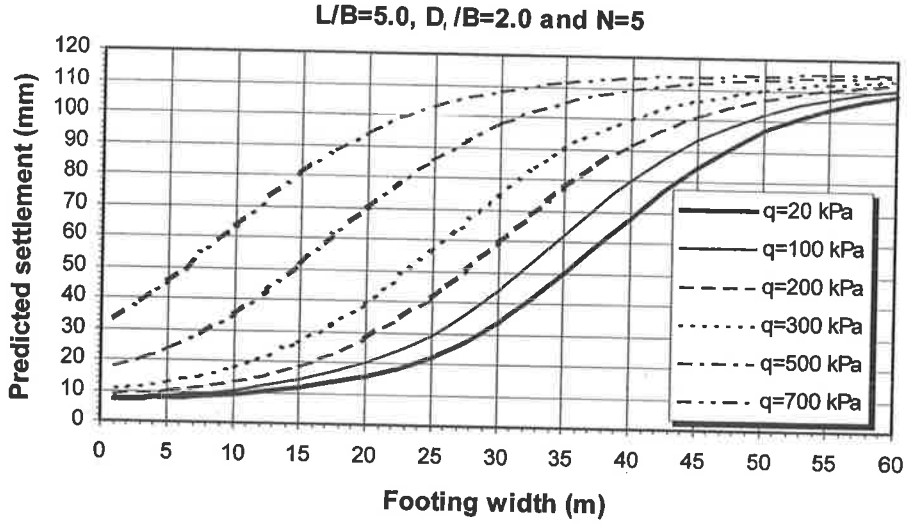

L/B=2.0, D_f/B=2.0 and N=30
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

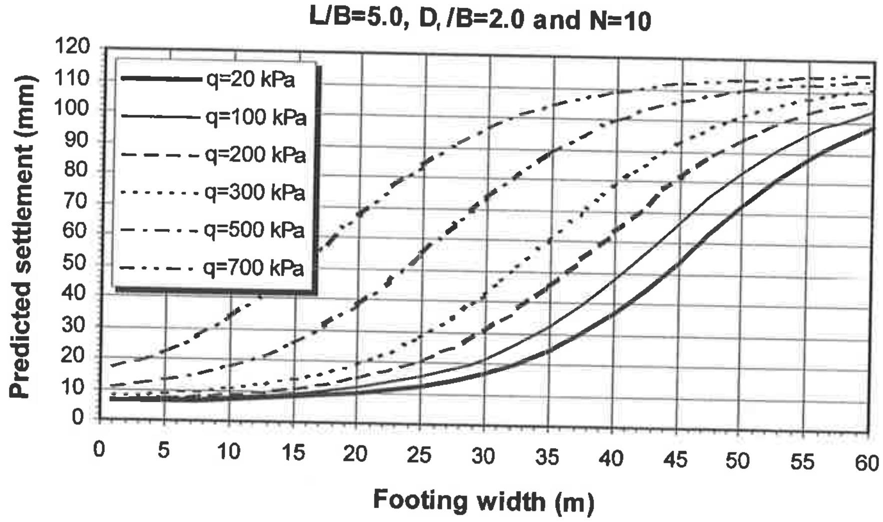

L/B=2.0, Df/B=2.0 and N=40
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

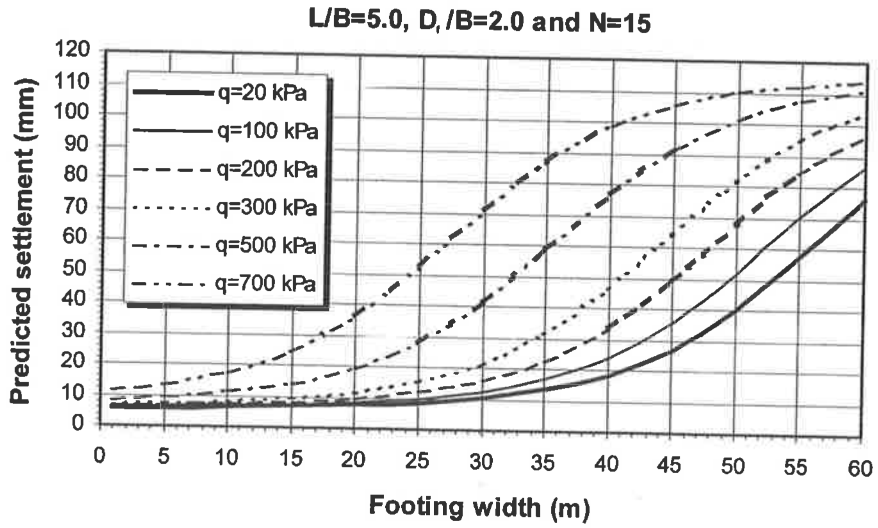

L/B=2.0, Df/B=2.0 and N=50
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)


L/B=2.0, Df/B=2.0 and N=60
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

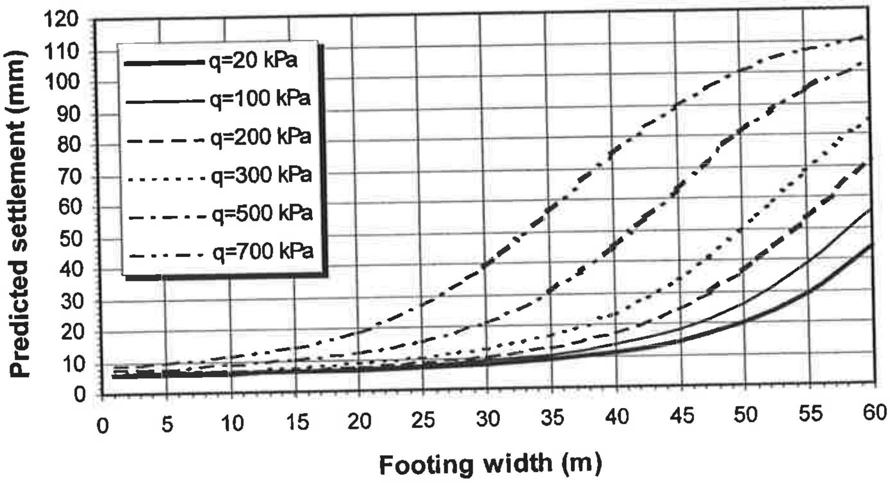
- ***L/B* = 5.0, *D***



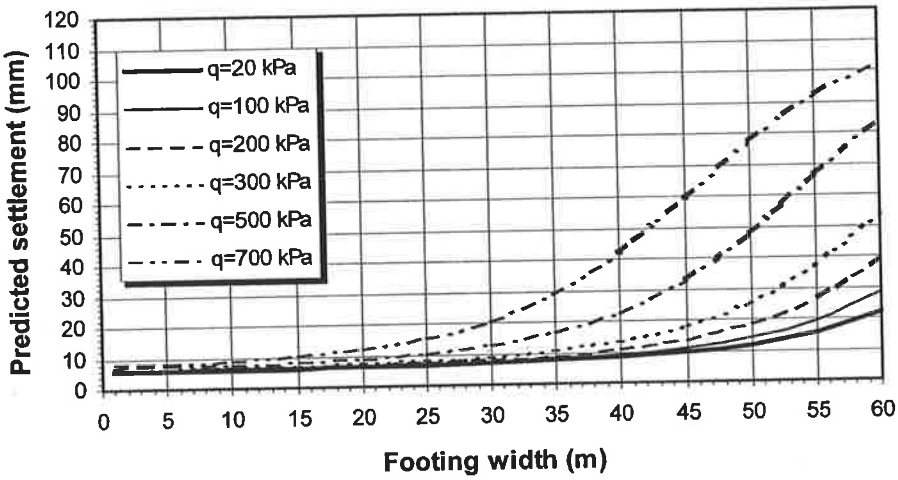




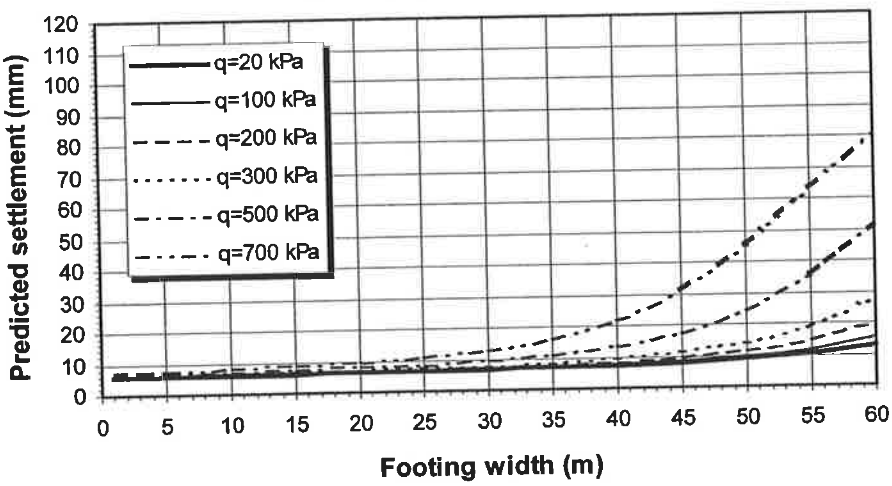
L/B=5.0, D_f/B=2.0 and N=20



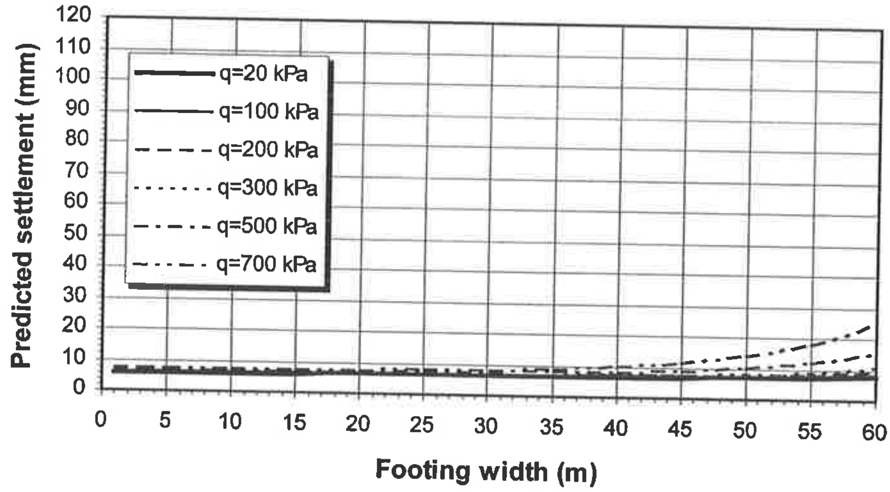
L/B=5.0, D_f/B=2.0 and N=25



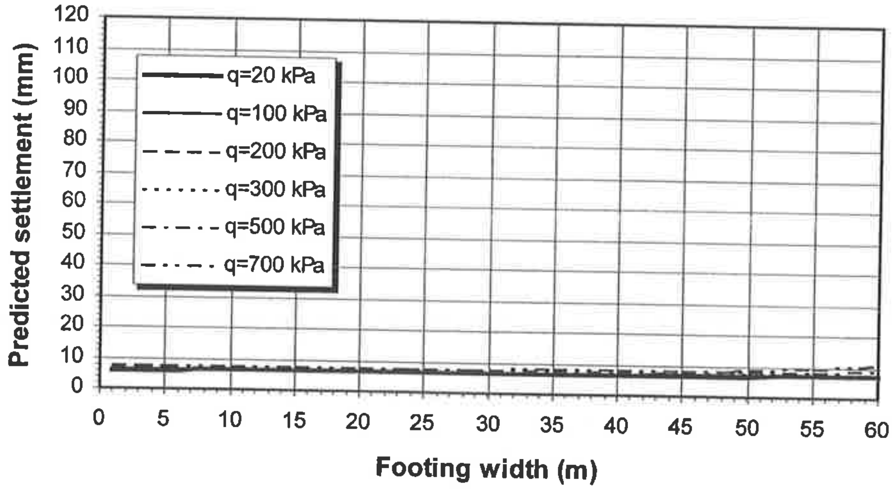
L/B=5.0, D_f/B=2.0 and N=30



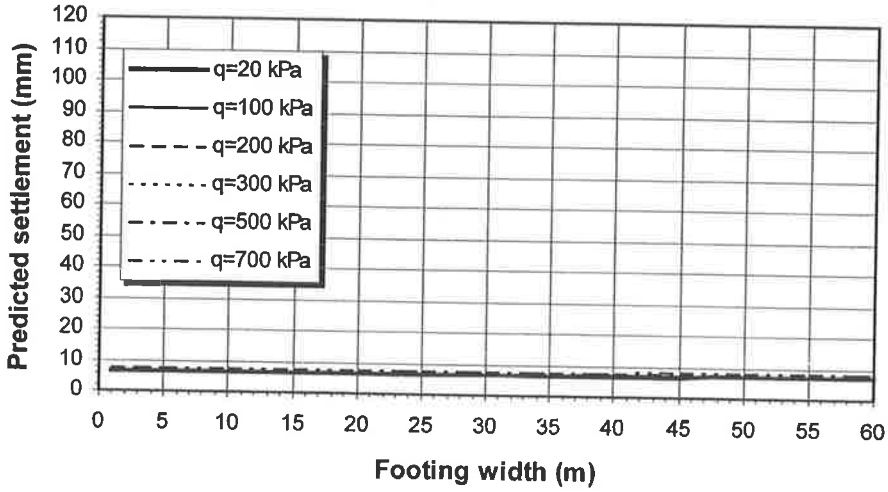
L/B=5.0, D_f/B=2.0 and N=40



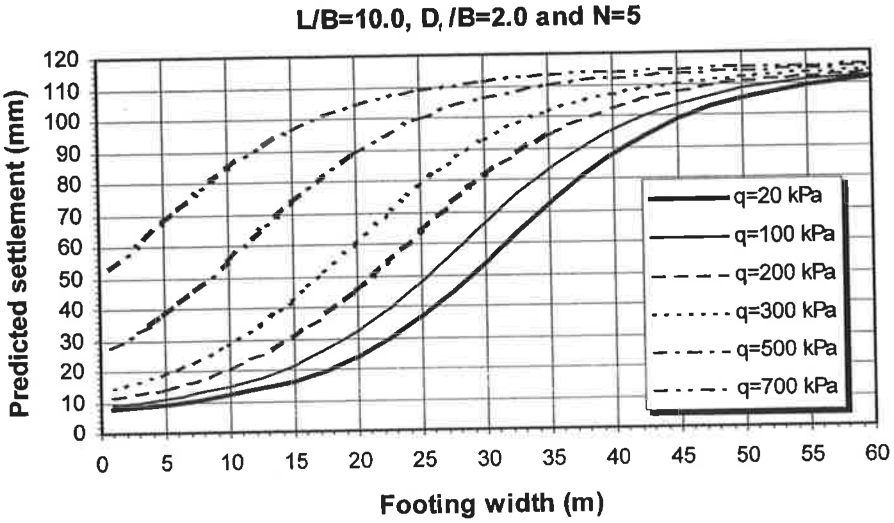
L/B=5.0, D_f/B=2.0 and N=50

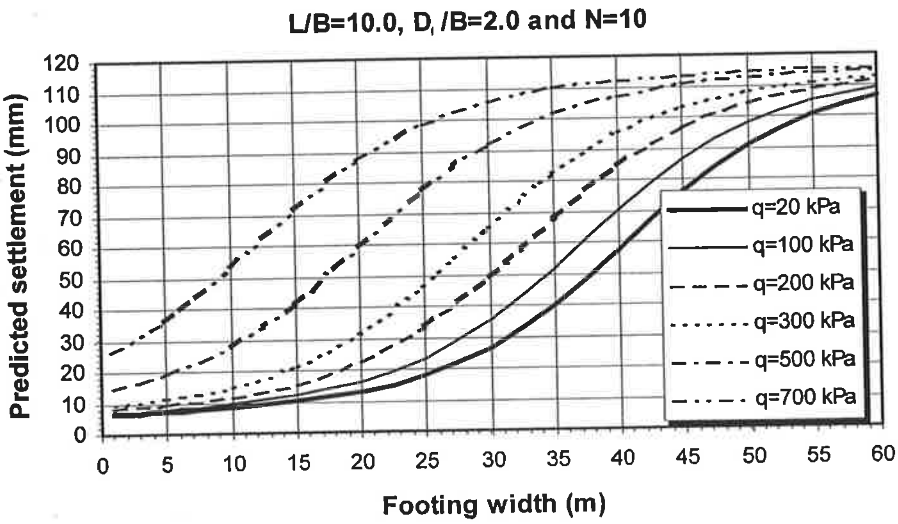


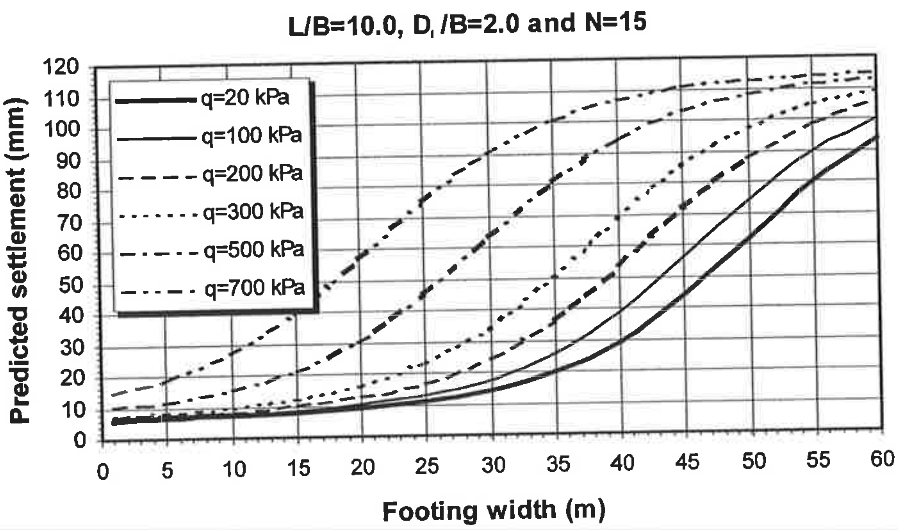
L/B=5.0, D_f/B=2.0 and N=60



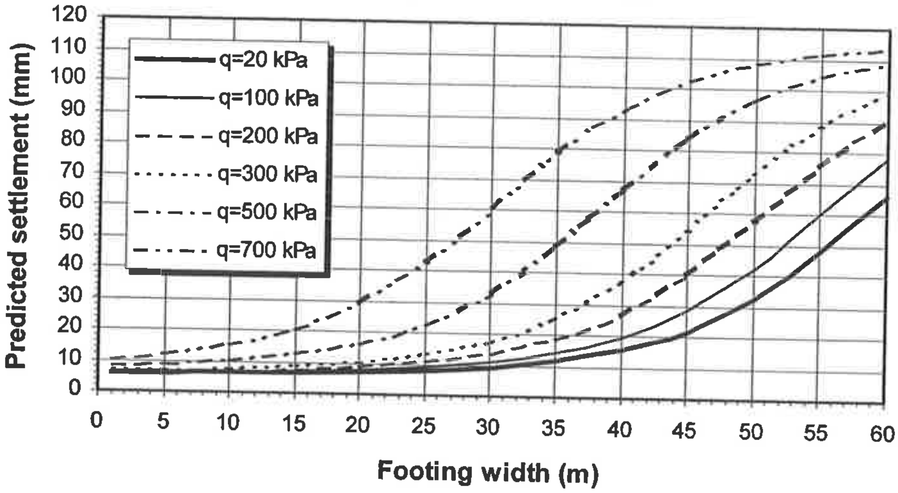
- **$L/B = 10.0, D_f/B = 2.0$**


L/B=10.0, D, /B=2.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

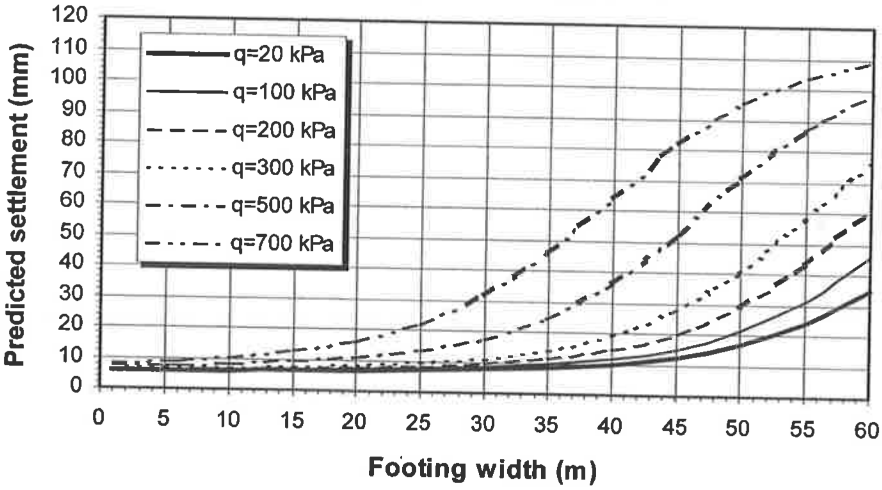

L/B=10.0, D, /B=2.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=10.0, D, /B=2.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

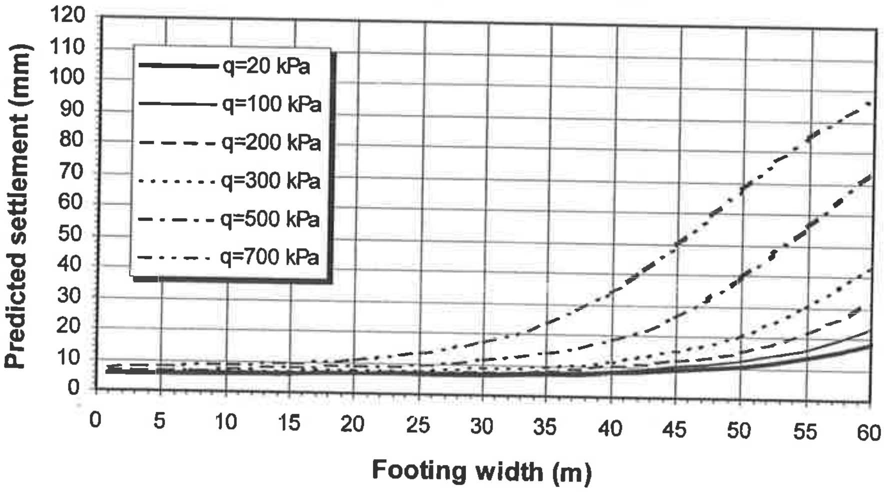
L/B=10.0, D_f/B=2.0 and N=20

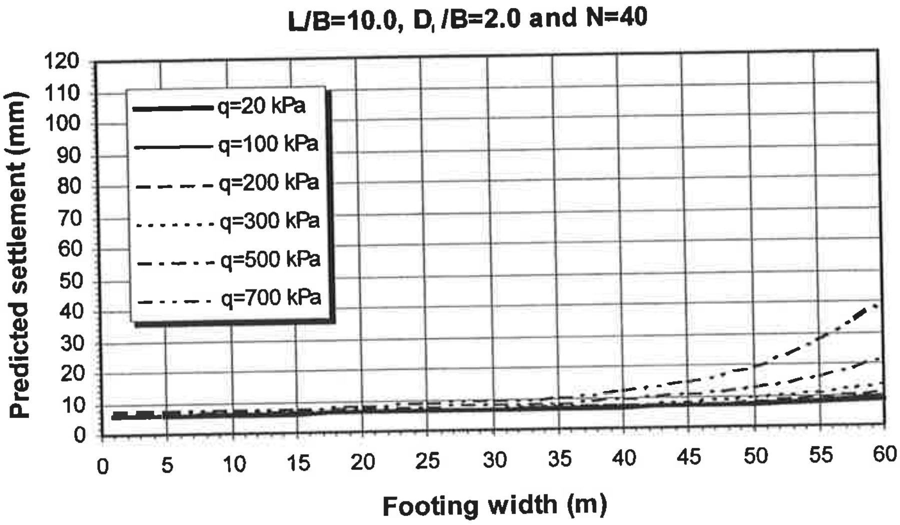


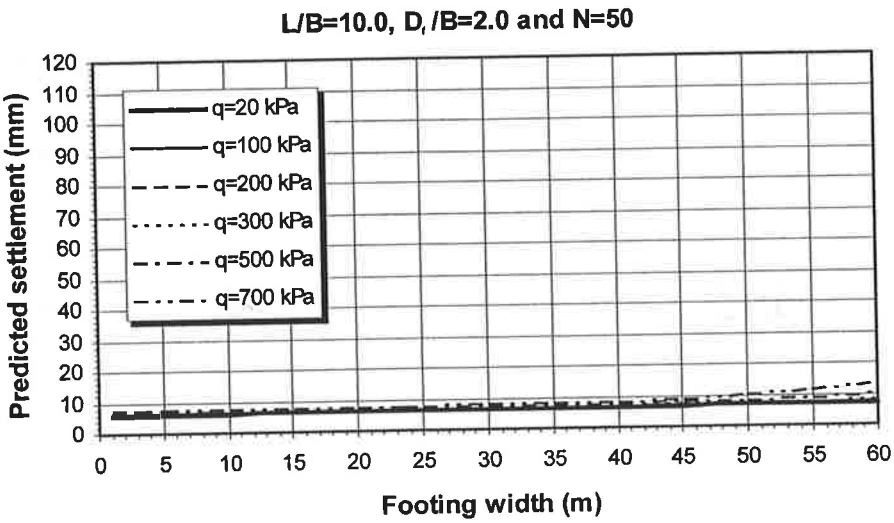
L/B=10.0, D_f/B=2.0 and N=25

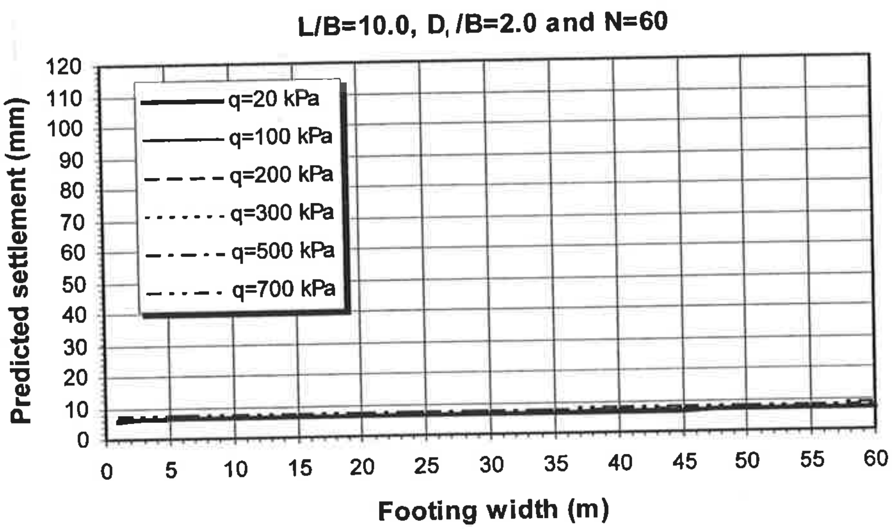


L/B=10.0, D_f/B=2.0 and N=30

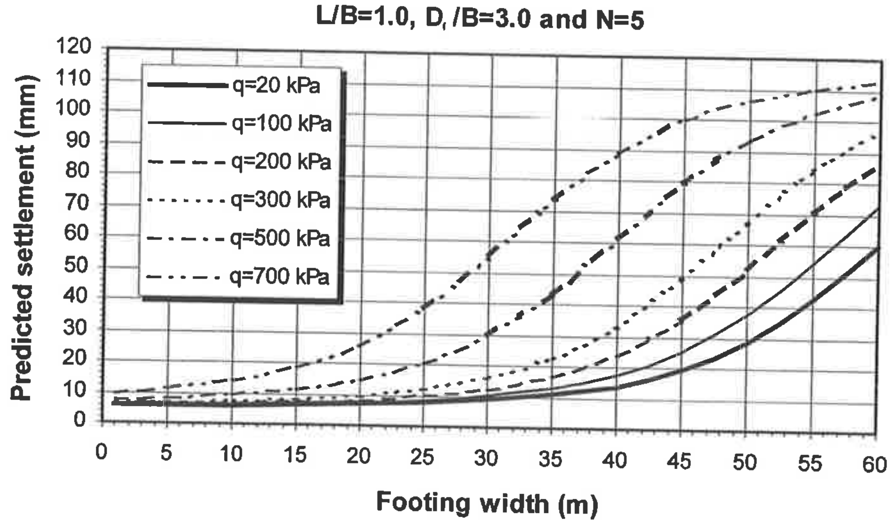


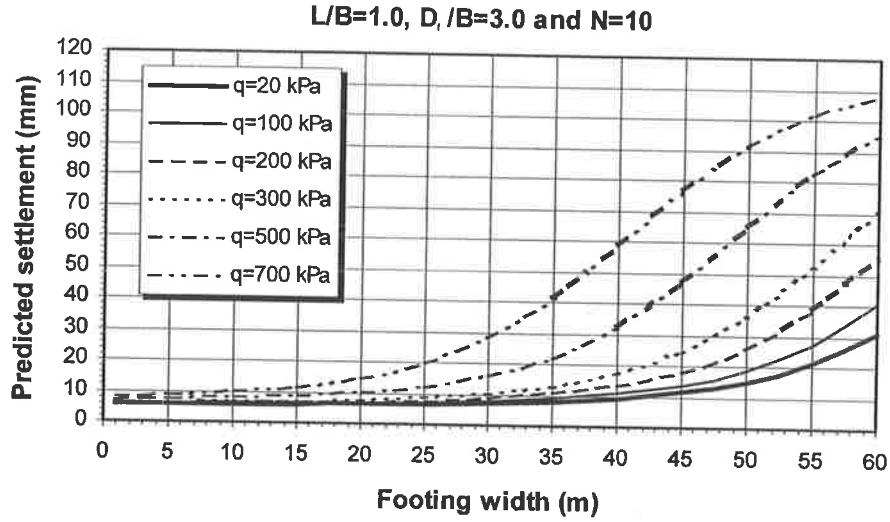

L/B=10.0, D_f/B=2.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

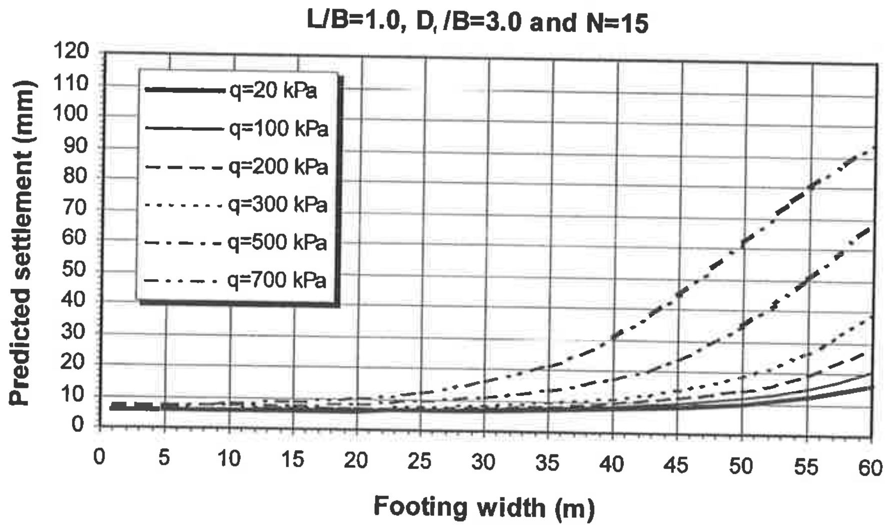

L/B=10.0, D_f/B=2.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60


L/B=10.0, D_f/B=2.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

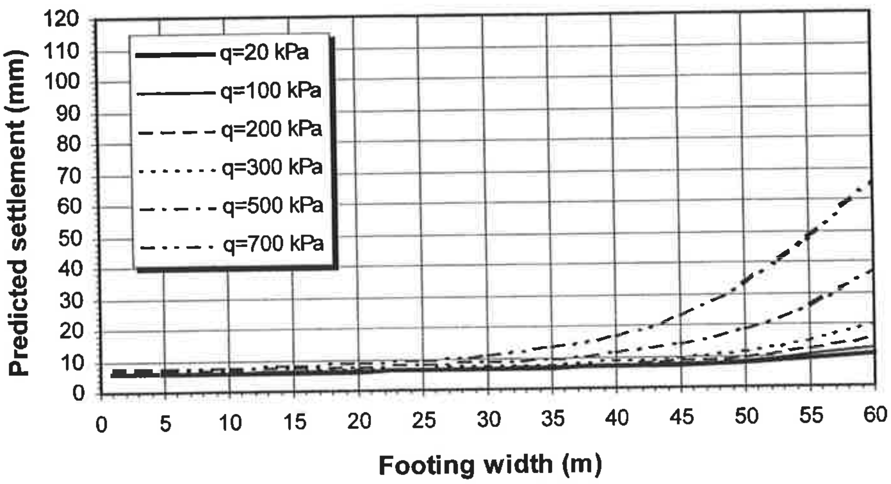
- *L/B* = 1.0, D_f/B = 3.0


L/B=1.0, Df/B=3.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

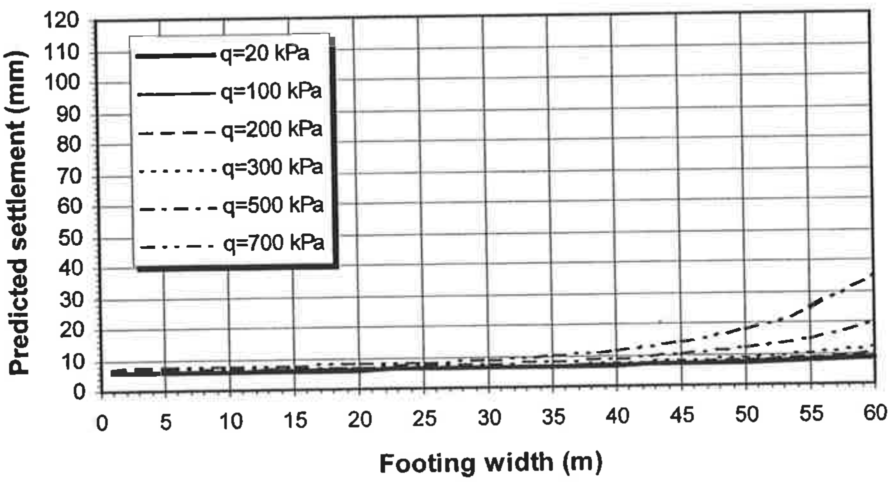

L/B=1.0, Df/B=3.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=1.0, Df/B=3.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

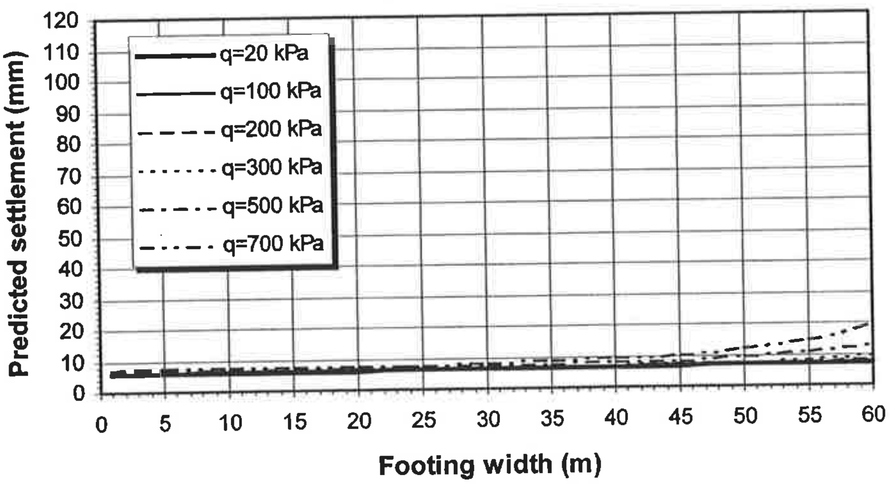
L/B=1.0, D_f/B=3.0 and N=20



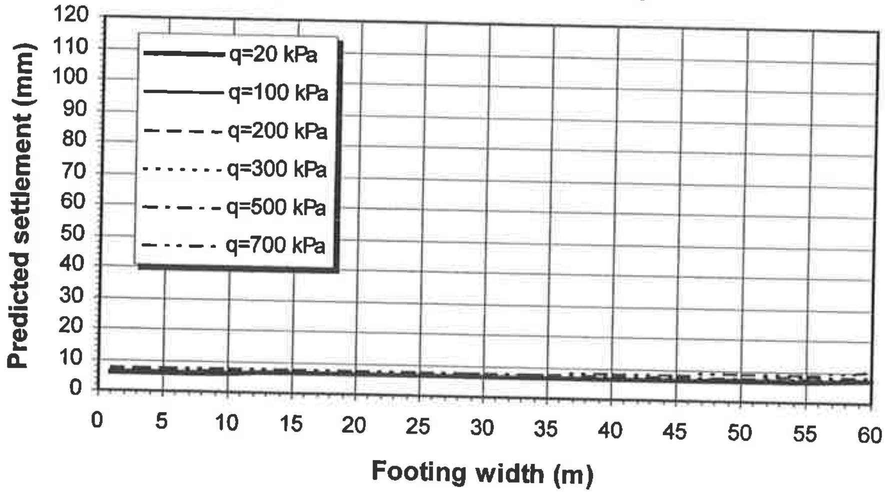
L/B=1.0, D_f/B=3.0 and N=25



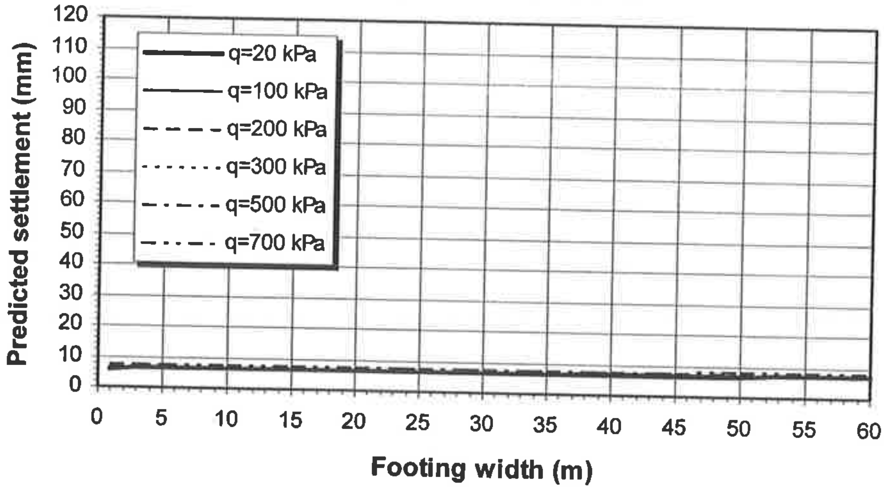
L/B=1.0, D_f/B=3.0 and N=30



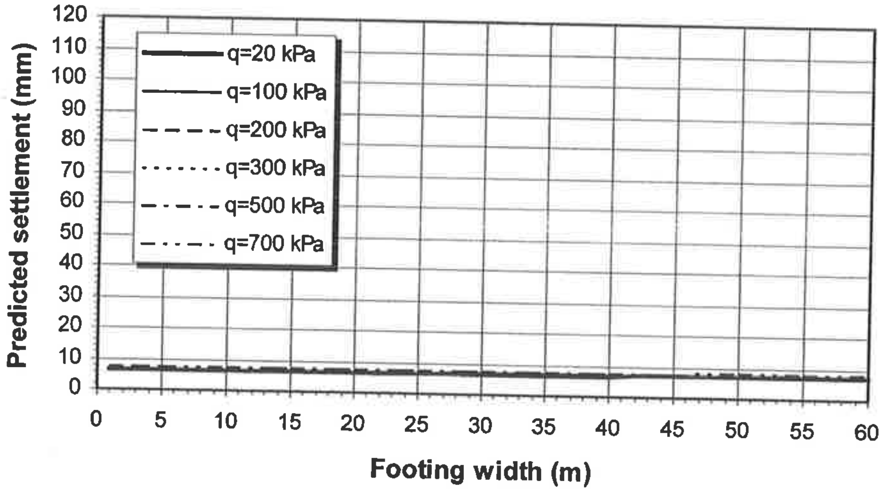
L/B=1.0, D_f/B=3.0 and N=40



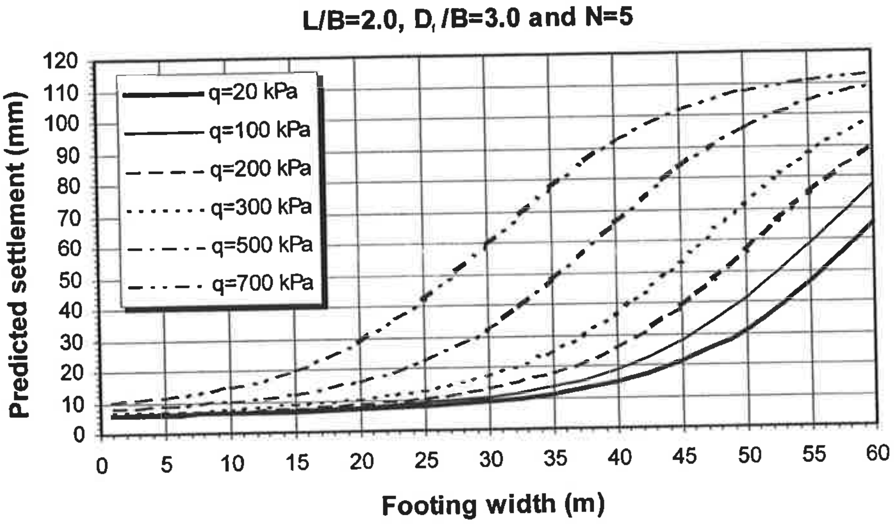
L/B=1.0, D_f/B=3.0 and N=50

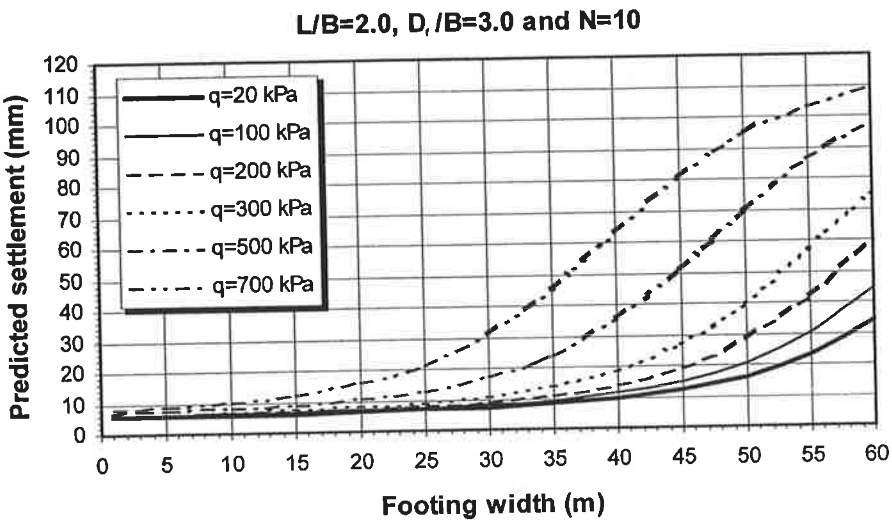


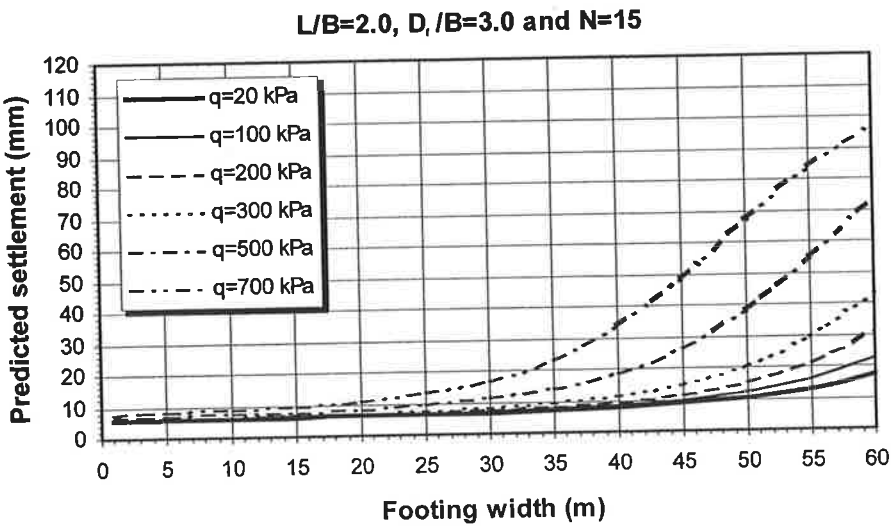
L/B=1.0, D_f/B=3.0 and N=60

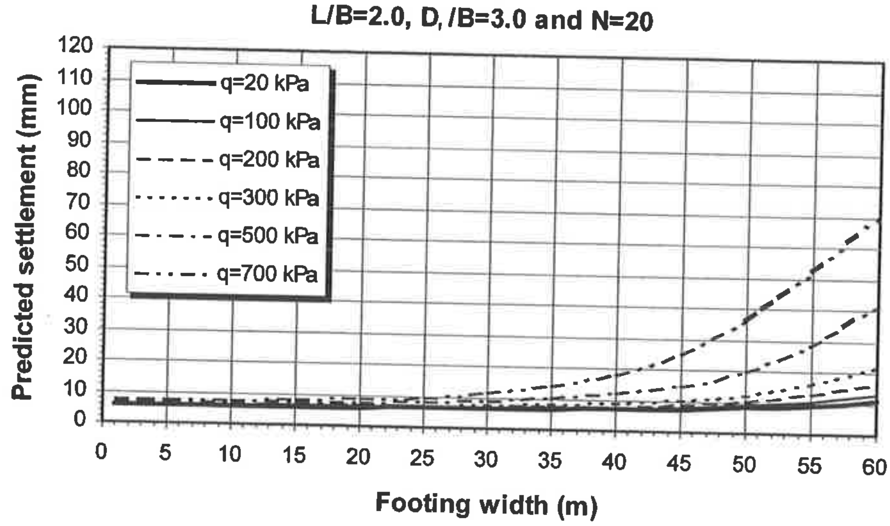


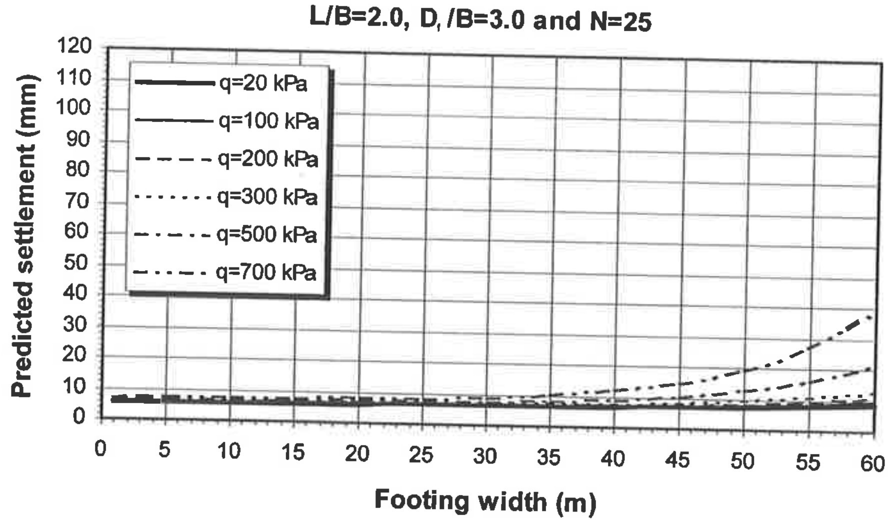
- **$L/B = 2.0, D_f/B = 3.0$**

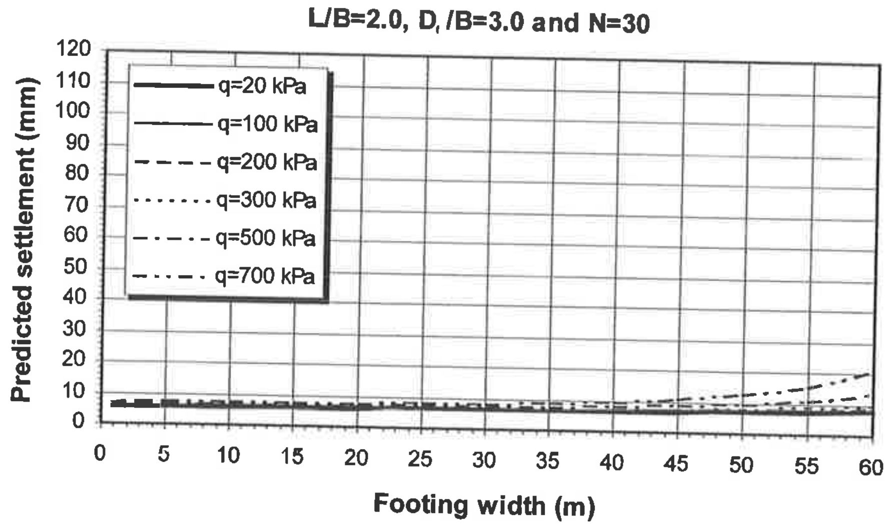

L/B=2.0, Df/B=3.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

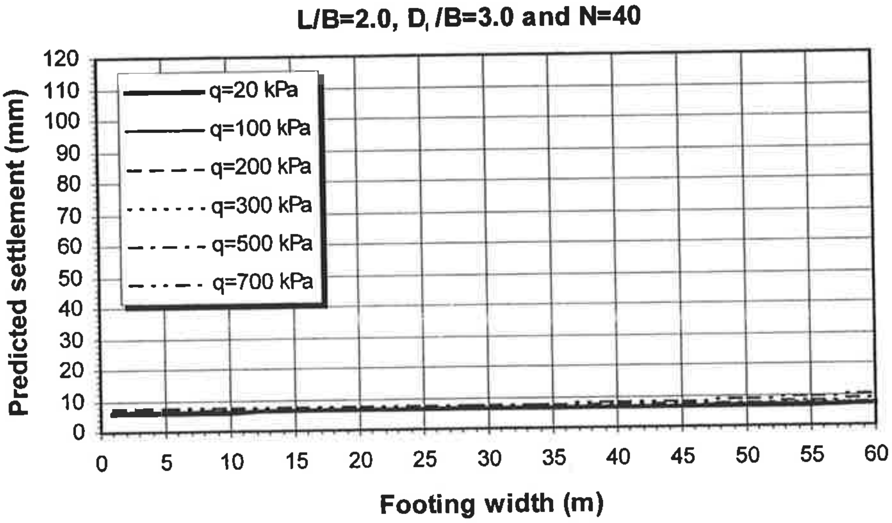

L/B=2.0, Df/B=3.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

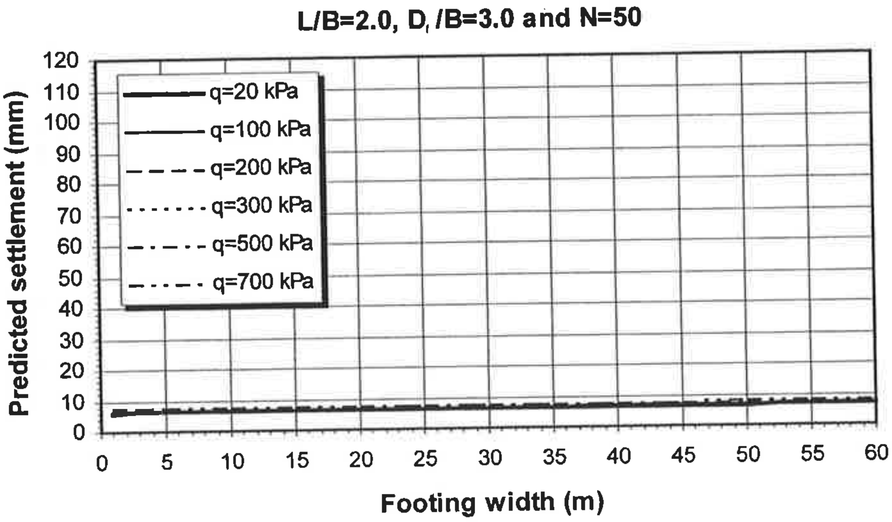

L/B=2.0, Df/B=3.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

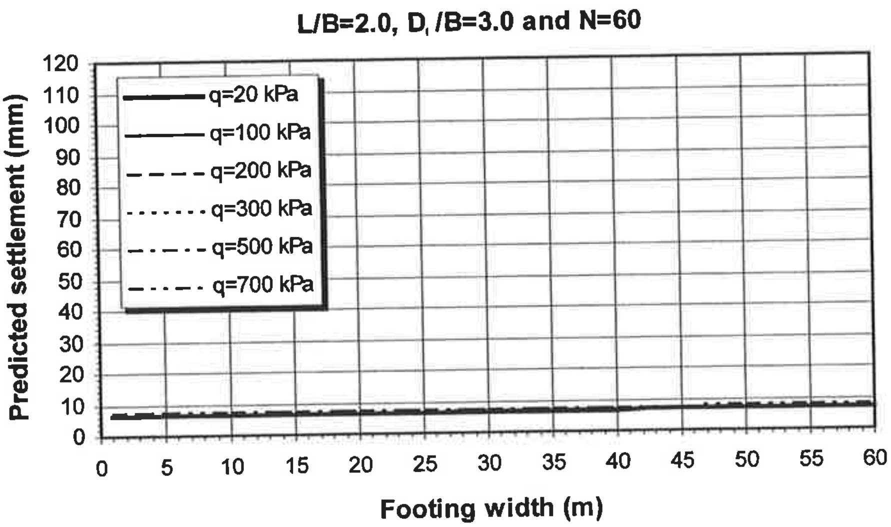

L/B=2.0, D, /B=3.0 and N=20
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D, /B=3.0 and N=25
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

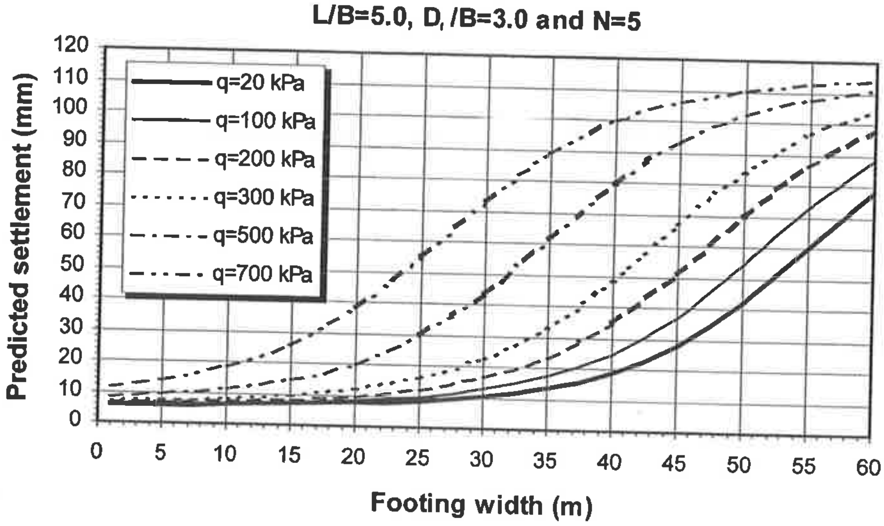

L/B=2.0, D, /B=3.0 and N=30
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

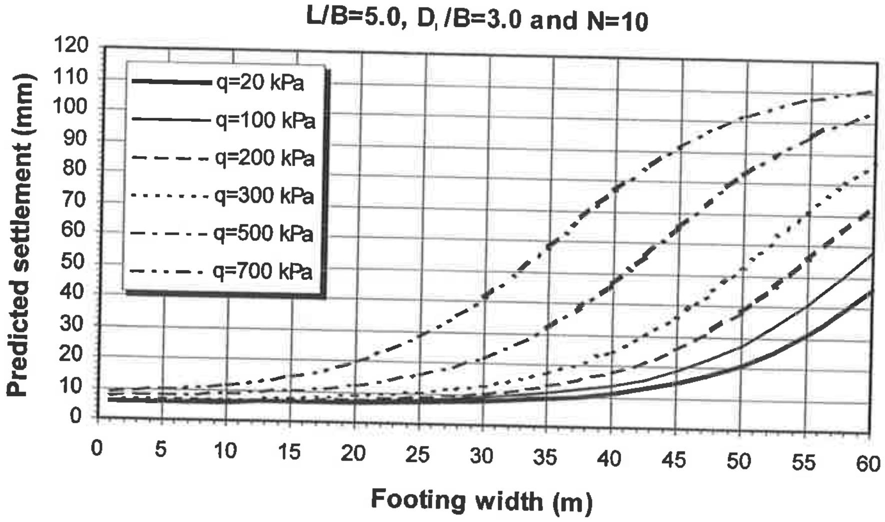

L/B=2.0, D_f/B=3.0 and N=40
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

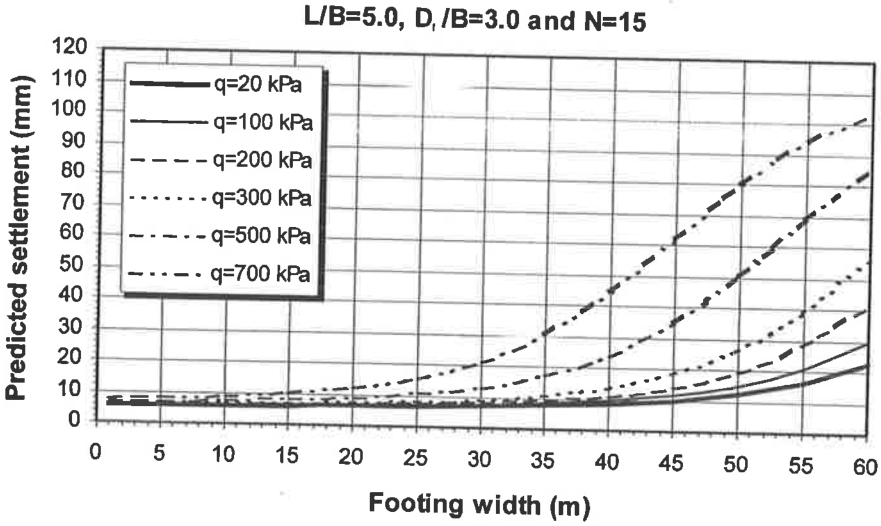

L/B=2.0, D_f/B=3.0 and N=50
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=2.0, D_f/B=3.0 and N=60
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

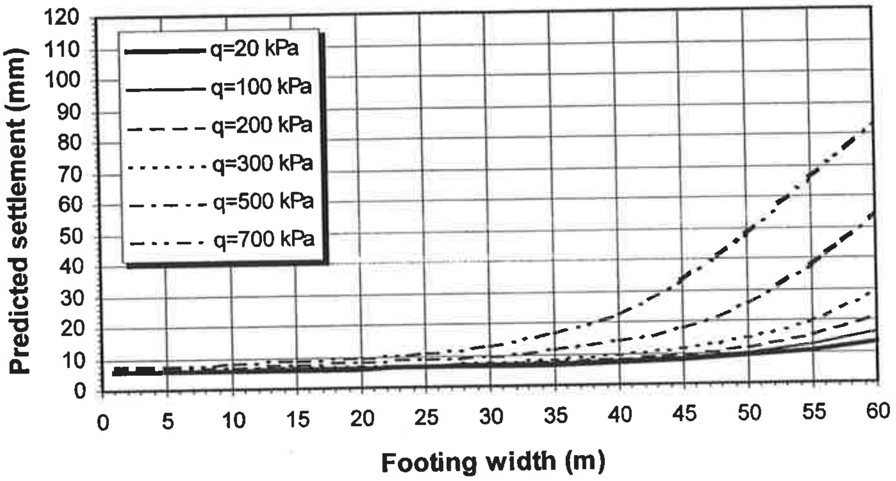
- **$L/B = 5.0$, $D_f/B = 3.0$**


L/B=5.0, D_f/B=3.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

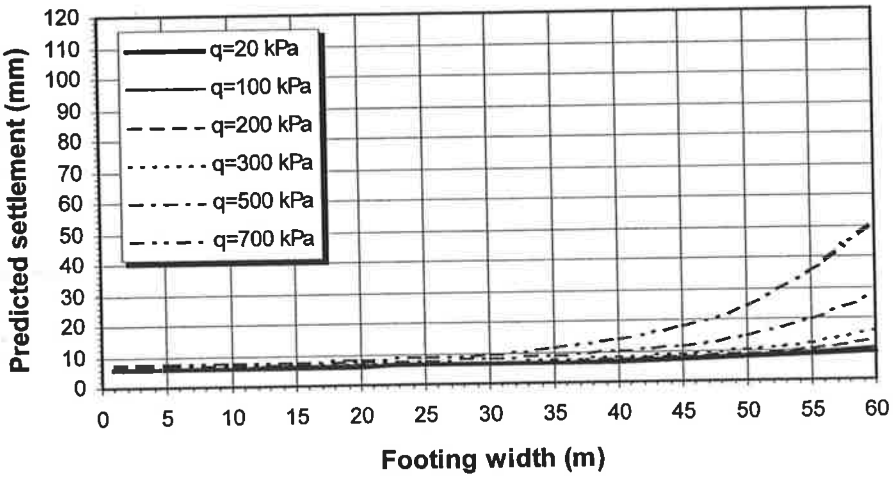

L/B=5.0, D_f/B=3.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa


L/B=5.0, D_f/B=3.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa

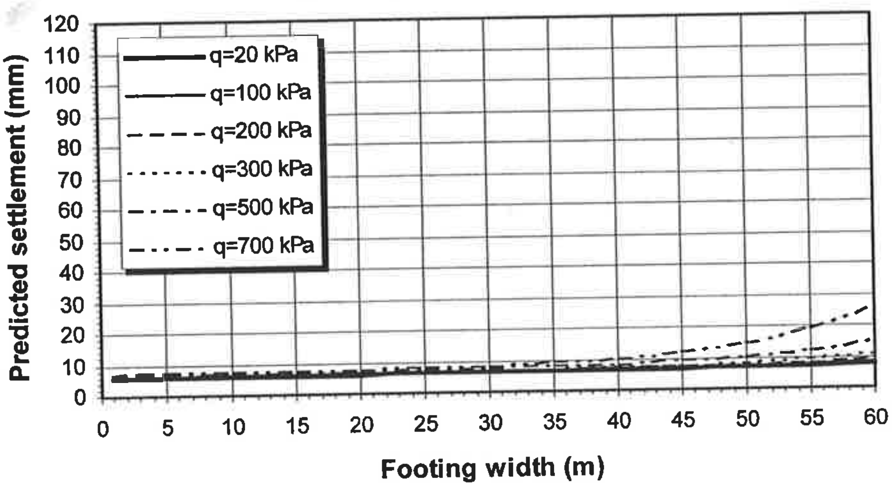
L/B=5.0, D_f/B=3.0 and N=20

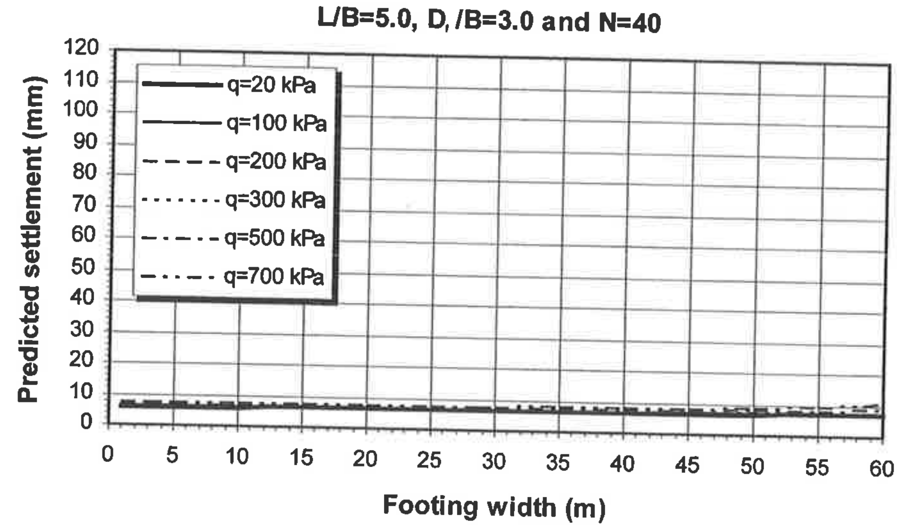


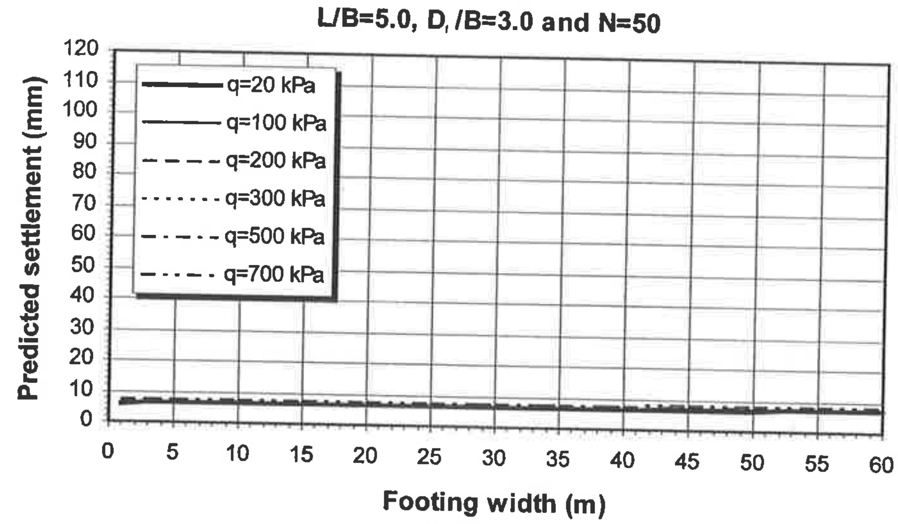
L/B=5.0, D_f/B=3.0 and N=25

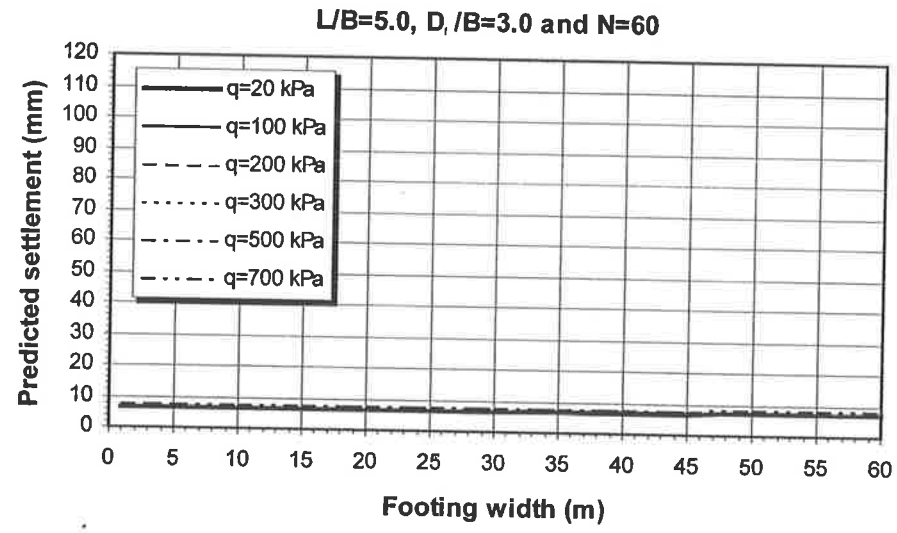


L/B=5.0, D_f/B=3.0 and N=30

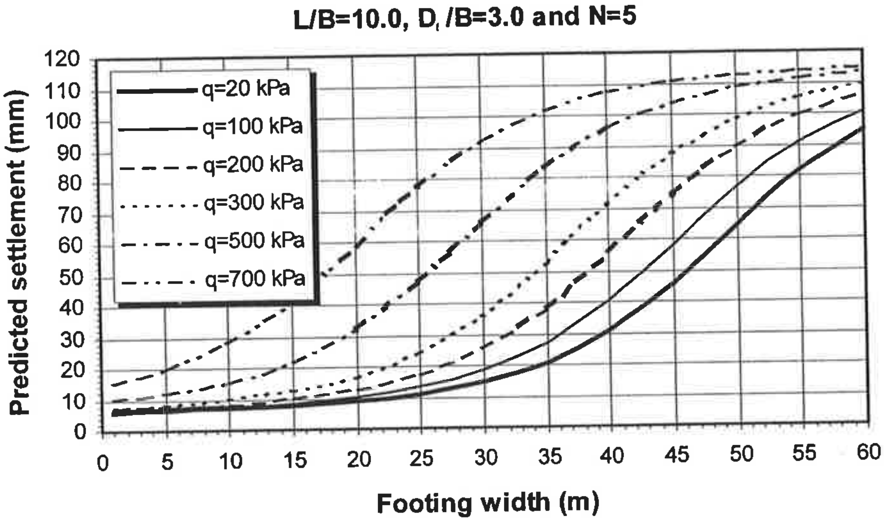


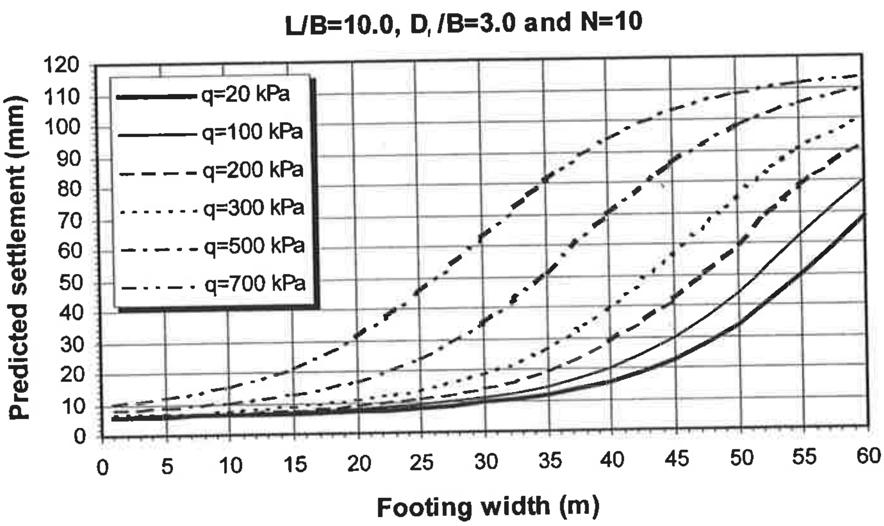

L/B=5.0, D_f/B=3.0 and N=40
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

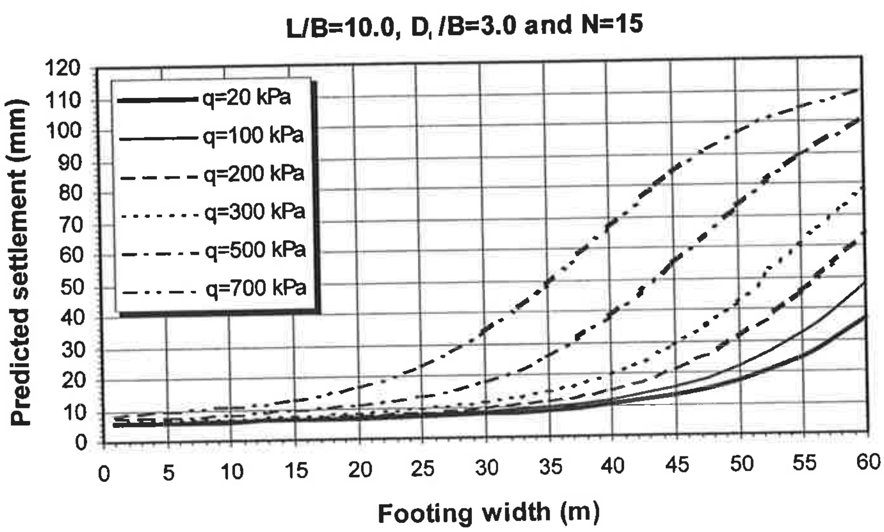

L/B=5.0, D_f/B=3.0 and N=50
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)


L/B=5.0, D_f/B=3.0 and N=60
Predicted settlement (mm)
120
110
100
90
80
70
60
50
40
30
20
10
0
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
5
10
15
20
25
30
35
40
45
50
55
60
Footing width (m)

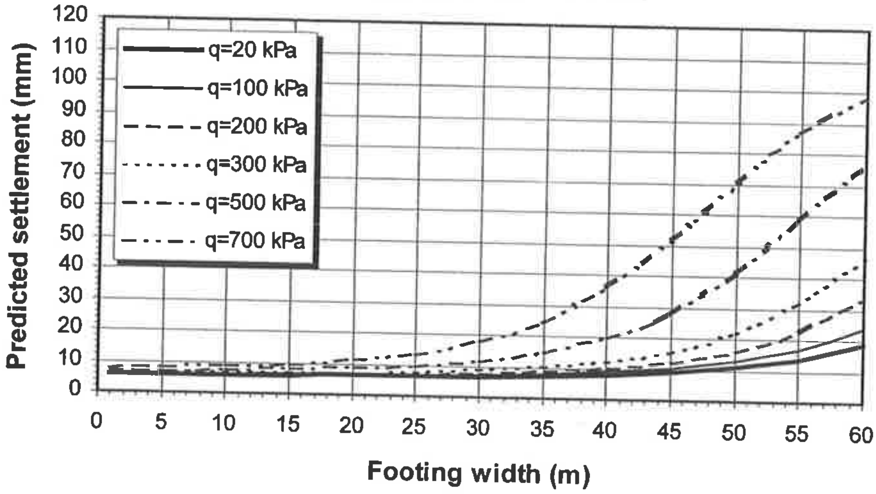
- $L/B = 10.0, D_f/B = 3.0$


L/B=10.0, Df/B=3.0 and N=5
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

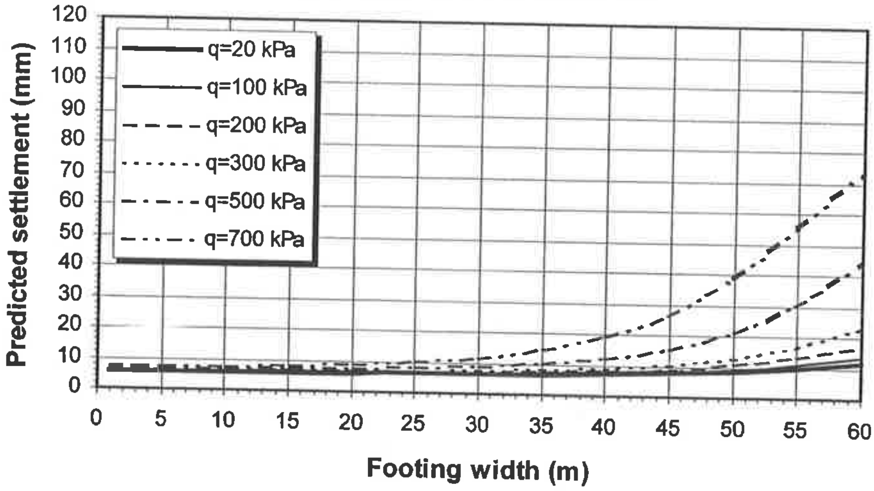

L/B=10.0, Df/B=3.0 and N=10
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60


L/B=10.0, Df/B=3.0 and N=15
Predicted settlement (mm)
Footing width (m)
q=20 kPa
q=100 kPa
q=200 kPa
q=300 kPa
q=500 kPa
q=700 kPa
0
10
20
30
40
50
60
70
80
90
100
110
120
0
5
10
15
20
25
30
35
40
45
50
55
60

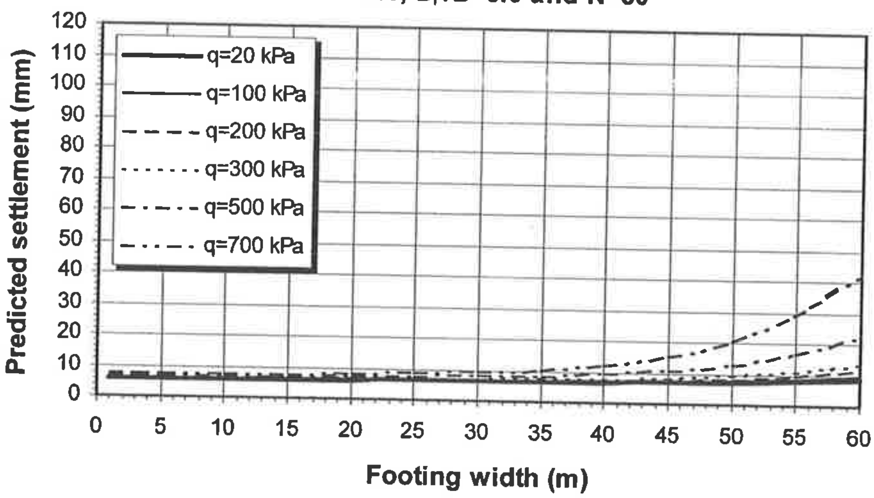
L/B=10.0, D



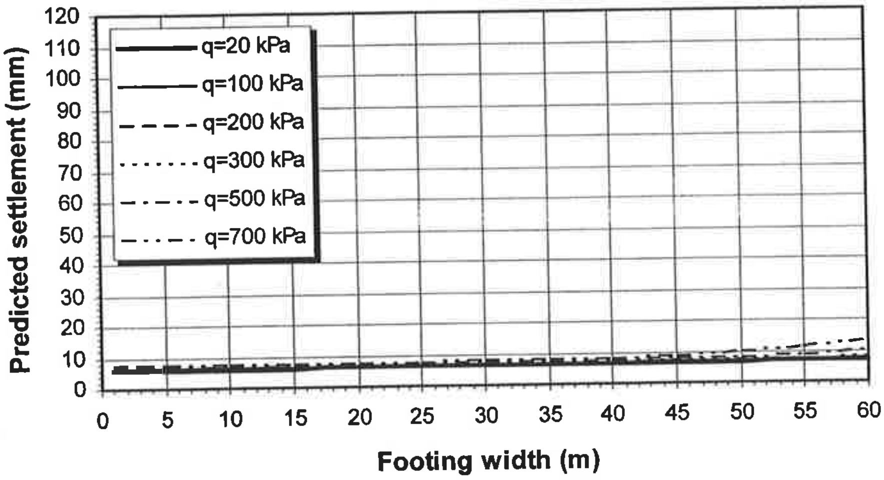
L/B=10.0, D_f/B=3.0 and N=25



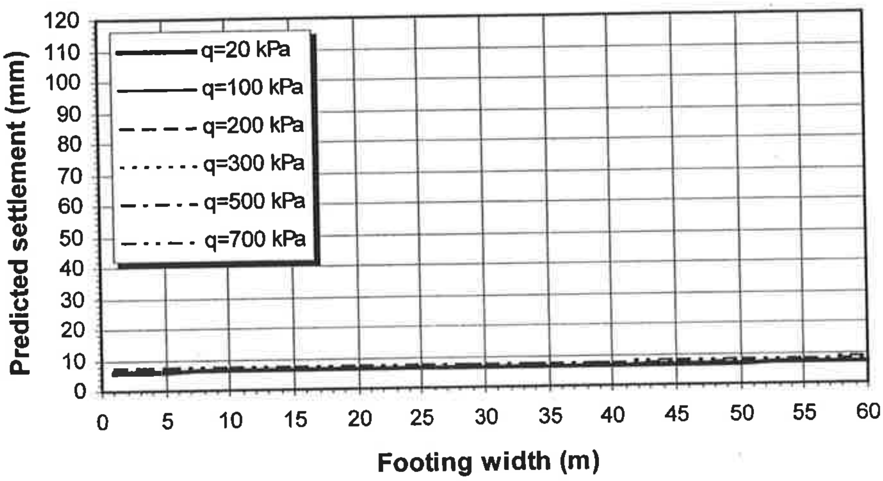
L/B=10.0, D_f/B=3.0 and N=30



L/B=10.0, D_f/B=3.0 and N=40



L/B=10.0, D_f/B=3.0 and N=50



L/B=10.0, D_f/B=3.0 and N=60