Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Intelligence Techniques and In Situ Soil Test Data

R. A. Tharanga Madhushani Ranasinghe
BSc. Eng. (Hons)

THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Civil, Environmental and Mining Engineering

February 2017

To my parents Shelton and Kusum,
my beloved husband Dilanth
and my brother Ravi

Preface

The work described in this thesis was undertaken from November 2012 to February 2017, in the School of Civil, Environmental and Mining Engineering, at the University of Adelaide. Throughout the thesis, all the materials, techniques, concepts and conclusions attained from other external resources have been duly referenced and appropriately acknowledged in the text.

Listed below are the sections of the thesis that designate the work which, to the best of her knowledge and belief, the author claims originality.

In **Chapter 4**:

- the details of the database of cone penetration test (CPT) results that were compiled from previous ground improvement projects associated with rolling dynamic compaction (RDC);
- the complete development of the model to predict the effectiveness of RDC based on CPT data using artificial neural networks (ANNs).

In **Chapter 5**:

- the details of the database of dynamic cone penetrometer (DCP) results that were compiled from previous ground improvement projects associated with RDC;
- the complete development of the model to predict the effectiveness of RDC based on DCP test data using ANNs.

In **Chapter 6**:

- the application of genetic programming (GP) for predicting the effectiveness of RDC based on CPT data;
- the comparison of the results of GP and ANN modelling.

In **Chapter 7**:

- the complete development of the computer program to predict the effectiveness of RDC based on DCP data using GP;
- the comparison of the results of GP and ANN modelling.

In **Chapter 8**:

- the assessment of the developed optimal ANN- and GP-based models and the selection of the most feasible approach of predicting the effectiveness of RDC on different ground conditions with respect to CPT and DCP data;
- the development of a comprehensive set of guidelines for each of the artificial intelligence (AI) techniques, i.e. ANN and GP.

A list of publications that have been prepared as a result of this research is presented below.

Refereed Conference Papers

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Application of Artificial Intelligence Techniques for Rolling Dynamic Compaction. *Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC*, Queenstown, New Zealand, pp. 41-47.

Refereed Journal Papers

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Neural Networks and Cone Penetration Test Data. Soils and Foundations, Submitted for Review.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Application of Artificial Neural Networks for Predicting the Impact of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results. Journal of Rock Mechanics and Geotechnical Engineering, In Press.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). A Genetic Programming Approach for Predicting the Effectiveness of Rolling Dynamic Compaction Using Cone Penetration Test Data. ICE-Ground Improvement, Under Review.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Use of Genetic Programming for the Predictions of Effectiveness of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results. Geotechnical and Geological Engineering, Under Review.

Awards

ARRB/Roads Australia Student Researcher Regional Prize - South Australia and Western Australia, 2016

Abstract

The research presented in this thesis focuses on developing predictive tools to forecast the effectiveness of rolling dynamic compaction (RDC) in different ground conditions. Among many other soil compaction methods, RDC is a widespread technique, which involves impacting the ground with a heavy (6–12 tonnes) non-circular (3-, 4- and 5-sided) module. It provides the construction industry with an improved ground compaction capability, especially with respect to a greater influence depth and a higher speed of compaction, resulting in increased productivity when compared with traditional compaction equipment.

However, to date, no rational means are available for obtaining a *priori* estimation of the degree of densification or the extent of the influence depth by RDC in different ground conditions. In addressing this knowledge gap, the research presented in this thesis develops robust predictive models to forecast the performance of RDC by means of the artificial intelligence (AI) techniques in the form of artificial neural networks (ANNs) and linear genetic programming (LGP), which have already been proven to be successful in a wide variety of forecasting applications in geotechnical engineering aspects. This study is focussed solely on the 4-sided, 8 tonne impact roller (BH-1300) and the AI-based models incorporate comprehensive databases consisting of in situ soil test data; specifically cone penetration test (CPT) and dynamic cone penetration (DCP) test data obtained from many ground improvement projects involving RDC.

Thus, altogether, two distinct sets of optimal models: two involving ANNs – one for the CPT and the other for the DCP; and two LGP models – again, one for the CPT and the other for the DCP – are presented. The accuracy and the reliability of the optimal model predictions are assessed by subjecting them to various performance measures. Furthermore, each of the selected optimal models are examined in a parametric study, by which the generalisation ability and the robustness of the models are confirmed. In addition, the performance of the optimal ANN and LGP-based models, as well as other

aspects, are compared with each other in order to assess the suitability and shortcomings of each. Consequently, a recommendation has been made of the most feasible approach for predicting the effectiveness of RDC in different ground conditions with respect to CPT and DCP test data. The models have also been disseminated via a series of mathematical formulae and/or programming code to facilitate their application in practice.

It is demonstrated that the developed optimal models are accurate and reliable over a range of soil types, and thus, have been recommended with confidence. As such, the developed models provide preliminary estimates of the density improvement in the ground based on the subsurface conditions and the number of roller passes. Therefore, it is considered that the models are beneficial during the pre-planning stages, and may replace, or at the very least augment, the necessity for RDC field trials prior to full-scale construction. In addition, the analyses demonstrate that the AI techniques provide a feasible approach for non-linear modelling involving many parameters, which in turn, further encourages future applications in the broader geotechnical engineering context. Finally, a comprehensive set of guidelines for each of the AI techniques employed in this research, i.e. ANN and LGP, is provided, with the intention of assisting potential and current users of these techniques.

Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: ---- ------ Date: 20/02/2017 ----------

Acknowledgements

I owe an enormous debt of gratitude to my primary supervisor, Prof. Mark Jaksa from the School of Civil, Environmental and Mining Engineering at the University of Adelaide, for his time, patience, guidance and continual support throughout my Ph.D. candidature. His encouragement is of great value to me and this thesis would not have been possible without his commitment and contribution. In addition, I especially, wish to appreciate his prompt assistance and advice through lengthy emails and longer phone conversations, as I was unable to meet him in person as I was residing in Melbourne for most of my candidature. Nonetheless, I am extremely grateful to him for dedicating his precious time and giving me the opportunity for one-to-one discussions, while transiting his travels through Melbourne. In addition, I wish to thank him for the proof reading of this thesis as well as other relevant publications.

My sincere acknowledgement is also extended to my co-supervisors, Dr. Yien Lik Kuo and Dr. Fereydoon Pooya Nejad from the School of Civil, Environmental and Mining Engineering at the University of Adelaide, for their insightful feedback and invaluable support during my research candidature. I thank them for their faith and confidence that motivated me towards the success.

I wish to acknowledge Mr. Stuart Bowes and the staff in Broons Hire (SA) Pty Ltd for their kind assistance, especially in providing access to their impact roller database upon which the developed numerical models are based. In addition, Stuart's kind support with sharing the knowledge, expertise and perception in the industry of rolling dynamic compaction is deeply appreciated. I also wish to thank Mr. Brendan Scott from the University of Adelaide, for his time, assistance and contribution for the development of the in situ soil test databases.

This research was supported under the Australian Research Council's Discovery Projects funding scheme and my candidature was sponsored by the University of

Adelaide via a full fee scholarship. My sincere appreciation is extended to both of these organisations for their generosity and continual support.

I also wish to express my sincere appreciation to the organisations, ARRB Group and Roads Australia, for awarding me the *Student Researcher Regional Prize of South Australia and Western Australia, 2016* and giving me the opportunity to expose my research innovations to industry. In addition, thanks are also due to Dr. Mike Shackleton and Ms. Rita Excell from ARRB Group, for their kindness and generosity.

I wish to thank fellow postgraduate students: Zhongyuan Fu; Andrew Bradley; and Erfan Syamsuddin for their excellent cooperation and friendship throughout my candidature. In addition, I wish to acknowledge the enormous support given to me by all other colleagues in the school, particularly, Mrs. Gayani Fernando, Mrs. Cherranthi Senarratne and Mrs. Nimasha Weerasinghe, who supported me in many ways both physically and emotionally. My appreciation is also extended to their families for making my stay in Adelaide pleasant and enjoyable. In addition, I wish to acknowledge the kind assistance given by the administrative staff, in particular, Ms. Josie Peluso, Ms. Julie Ligertwood and Ms. Ann Smith at the University of Adelaide.

I am also grateful to my husband, Dilanth De Silva, for his endless love, and continual support throughout my research candidature. While being a Ph.D. candidate in Monash University, his selfless sacrifices and the enormous effort to make me comfortable even during challenging situations are deeply appreciated. Of course, it is a great pleasure to share my life with him and it is my happiness to pace the same path together. I also wish to express my profound gratitude for my parents, Shelton and Kusum, and my brother Ravi, for their incredible support and considerable sacrifices, which they have made on my behalf throughout my life journey. I also wish to thank the extended family, especially, my parents-in-law, for their great emotional support over the period.

Contents

Note: All blank pages have been given page numbers.

List of Figures

List of Tables

Notation

A	minimum value of the unscaled dataset;
a	minimum value of the scaled dataset;
B	maximum value of the unscaled dataset;
b	maximum value of the scaled dataset;
D	depth of measurement;
d_j	desired (actual or measured) output;
E	global error function;
$f(.)$	transfer function;
f_s; f_{si}	sleeve friction; sleeve friction prior to compaction;
f_{sig}	sigmoid transfer function;
I	number of input variables;
I_d	index for densification;
I_j	activation level of node j;
n	number of nodes or samples;
P	number of roller passes;
q_c; q_{ci}; q_{cf}	cone tip resistance; cone tip resistance prior to compaction; cone tip resistance after compaction;
R	coefficient of correlation;
R^2	coefficient of determination;
R_f	friction ratio;
t	iteration number;
w_{ji}	connection weight between nodes i and j;
w_{kj}	connection weight between nodes j and k;
x_i	input from node i;
x_{max}	maximum value of variable x;
x_{min}	minimum value of variable x;
x_{scaled}	scaled input variable;
$x_{unscaled}$	unscaled input variable;
y_k	predicted output of node k;

y_j	predicted output of node j;
θ_j	threshold for node j;
θ_k	threshold for node k;
η	learning rate;
μ	momentum term; and
σ	standard deviation.

Abbreviations

AI	artificial intelligence;
AIMGP	automatic induction of machine code by genetic programming;
ANN	artificial neural network;
CBR	California bearing ratio;
CGP	Cartesian genetic programming;
CIS	continuous induced settlement;
CNN	computational neural network;
CPT	cone penetration test;
CSWS	continuous surface wave system;
CV	coefficient of variation;
DCP	dynamic cone penetrometer test;
EA	evolutionary algorithm;
EPC	earth pressure cell;
FEM	finite element method;
GA	genetic algorithm;
GE	grammatical evolution;
GEP	gene-expression programming;
GP	genetic programming;
LGP	linear genetic programming;
MAE	mean absolute error;
MAPE	mean absolute percentage error;
MARE	mean absolute relative error;
MASW	multichannel analysis of surface wave;
MDD	maximum dry density;
MEP	multi-expression programming;
MLP	multi-layer perceptron;
MLR	multivariate linear regression;
MSE	mean squared error;
NNN	natural neural network;

OMC	optimum moisture content;
PPT	Perth penetrometer test;
RDC	rolling dynamic compaction;
RE	relative error;
RMSE	root mean square error;
SASW	spectral analysis of surface wave;
SD	standard deviation;
SE	standard error;
SOM	self-organizing map;
SPT	standard penetrometer test;
SRA	simple regression analysis;
SSE	sum of squared error;
TFR	transition fine content ratio;
TGP	tree-based genetic programming.

Chapter 1.

Introduction

1.1 INTRODUCTION

The ground, which involves soils and rock, by its nature, exhibits varied and uncertain behaviour because of its formation and variability. However, sometimes, these ground variabilities impose limitations upon which constructions are affected. Thus, geotechnical engineering often deals with problematic soil conditions, where ground improvement techniques are often necessary. Ground improvement provides modification or alteration of existing site foundation soils or earth structures in order to facilitate better performance under operational conditions (Army, 1999). As identified by Munfakh and Wyllie (2000), the main objectives of ground improvement are to:

- increase the bearing capacity (i.e. reduce settlement);
- control deformations and accelerate consolidation;
- provide lateral stability;
- form seepage cut-off and environmental control; and
- increase resistance to liquefaction.

Currently, there are more than 30 different ground improvement techniques (Phear and Harris, 2008), which can be broadly categorised into 5 main groups: removal, compaction, consolidation, modification and load transfer. Amongst these, compaction has found to be an appropriate, cost effective and feasible approach for a number of projects associated with ground-related risks. Soil compaction is a form of ground improvement technique, whereby the ground is physically compressed by means of applied mechanically energy. It often involves the rearrangement of soil structure, whilst the strength of the material is increased with a reduction in porosity and hydraulic conductivity. Soils are compacted by means of 4 different types of compaction effort: vibration, kneading, pressure and impact, which can be further categorised into two principal types of compactive force: static and dynamic. Static compaction involves the densification of the ground by means of a downward force on the ground surface by the self-weight of the machinery, whereas dynamic compaction is characterised with a kinetically-driven downward force additional to the

equipment's static weight. The most common types of static compaction equipment are circular rollers, which usually employ drums, pad feet and pneumatic multi-tyres, whilst dynamic compaction makes use of heavy tamping, vibratory drums and plates, rammers, vibroflotation and rolling dynamic compaction.

Impact rolling, generically known as rolling dynamic compaction (RDC), is a well-established method of soil compaction where the soil densification is achieved by means of high energy impacts. RDC involves heavy (6–12 tonnes) non-circular modules (3-, 4- and 5-sided), which rotate about their corners and fall to the ground as they are drawn forward behind a tractor. Consequently, RDC enables the impact roller to impart a greater amount of compressive energy on to the soil and thus, RDC often provides an alternative to the traditional approaches of ground improvement, with superior compaction capabilities. As such:

- RDC is effective, in that it improves the ground to a greater depth – more than 1 m beneath the ground surface and sometimes deeper than 3 m in some soils; compared to conventional static and vibratory compaction, where the influence depths are generally less than 0.5 m (Clifford, 1976, 1978b; Avalle and Carter, 2005; Jaksa et al., 2012);

- RDC can compact thicker lifts, in excess of 0.5 m, which is considerably greater than the traditional layer thicknesses of approximately 0.3 m, which enhances RDC's cost effectiveness (Avalle, 2006; Scott and Jaksa, 2012); and

- RDC is also very efficient when employed in large and open sites, as it traverses the ground at speeds of 9–12 km/h, which is much faster than traditional compaction methods, using a vibratory roller, for example, which travels at a speed of 4 km/h (Pinard, 1999).

To date, wider experience with RDC has been gained from impact roller applications in a variety of fields particularly in civil and mining projects, pavement rehabilitation

and in the agricultural sector. In addition, a number of field studies conducted to date has added to the existing body of knowledge on the mechanics of RDC, its energy transfer and especially its effectiveness and zone of influence in a range of ground conditions. Moreover, research has been directed towards investigating the factors affecting the efficacy of RDC, whereas optimization methods have also been suggested.

However, there still exists a considerable gap in knowledge in relation to the prior estimation of the degree of densification and the extent of the influence depth by RDC in different ground conditions. This often results in the use of RDC being based on intuition and past experience from previous work undertaken in similar soils and site conditions. In addition, the ground improvement projects that identified RDC as a possible application for earthworks compaction, generally require a detailed trial program in order to affirm the method specification or for verification of RDC performance, which usually incurs a non-trivial cost and time commitment. Thus, there exists an urgent need for a rational method for the prior assessment of RDC with respect to various influencing factors. Such predictive models and empirical guides are readily available for many of the more common compaction practices. For instance, empirical guides are available to estimate the depth and ground conditions under which heavy tamping is applicable (Lukas, 1995). A reliable and accurate predictive model for RDC, applicable in a range of ground conditions, will enable geotechnical engineers to make *a priori* estimates of the effectiveness and the depth of influence associated with impact rolling. Indeed, forecasting the influence of RDC is complex due to the heterogeneous nature of the ground and the various site-specific factors that can potentially affect the improvement process.

In addressing the above knowledge gap, the present study aims to develop an accurate and robust predictive model for prior estimation of the effectiveness of RDC by means of the artificial intelligence (AI) techniques in the form of artificial neural networks (ANNs) and genetic programming (GP), which have been shown to be of value in the broader geotechnical engineering context (Shahin et al., 2005a; Das et al., 2011; Alavi and Gandomi, 2012; Alavi et al., 2013). Attention is focussed on the 4-sided, 8 tonne

'Broons impact roller' (BH-1300), which is the most common 4-sided application of RDC in practice. It is intended that the model will provide additional, *a priori* information to supplement field trials undertaken on site prior to ground improvement. It is not expected that such AI models will replace or devalue site-specific field trials; rather they will provide a worthwhile additional tool for ground improvement projects involving RDC. To date, no such predictive models exist for RDC, neither empirical, theoretical nor numerical.

1.2 AIMS AND SCOPE OF THESIS

This research aims to investigate and quantify the effectiveness of RDC in a range of ground conditions and seeks to establish predictive tools, which are based on AI techniques. This study is focussed solely on the 4-sided, 8 tonne impact roller (BH-1300) and the AI-based models are developed using an extensive database consisting of in situ soil test data; specifically cone penetration test (CPT) and dynamic cone penetration (DCP) test data obtained from many ground improvement projects involving RDC.

Specifically, this thesis aims to:

- develop an extensive database of actual, measured test records of ground improvement using RDC, together with a broad range of governing soil parameters;

- explore the use of ANNs and GP for a reliable prediction of the ground improvement by RDC over a wide range of soil parameters;

- compare the accuracy and reliability of the two AI techniques (ANNs and GP);

- evaluate the influence of a range of input parameters and assess their relative importance with respect to RDC ground improvement;

- where possible, produce a simple and practical formula, based on the optimum AI model results, to forecast the performance of RDC that will improve the application of RDC in practice; and

- establish a set of general guidelines for developing ANN and GP models to facilitate their further application in the future.

1.3 LAYOUT OF THE THESIS

This thesis details the research undertaken to develop predictive tools which quantify the effectiveness of RDC using AI techniques, as outlined in the following chapters. The thesis is presented in the form of a 'thesis by publication', where each of the major contributions has either been published in relevant international, peer reviewed journals, or has been submitted for review. The details of these are provided below.

Chapter 2 provides a brief overview of the conventional soil compaction methods followed by a detailed assessment of RDC, highlighting its advantages, applications and limitations. A review of the existing literature regarding the estimation of the effectiveness and zone of influence of RDC is also given. In addition, the factors affecting the efficacy of RDC, which are relevant to the numerical modelling undertaken in this study, are presented. In discussing the current practice of estimating the effectiveness of RDC, an example of an impact roller compaction trial is described and thus, the need for a predictive tool for *a priori* prediction of the influence of RDC is highlighted. The field tests and measurement techniques used for the verification of ground improvement by RDC are then briefly discussed, as these test data will be incorporated in the subsequent AI model development.

Chapter 3 provides a detailed review of the two AI techniques, ANNs and GP, which are the focus of the numerical modelling undertaken in this study, by providing background information and their structure and operation. Later in this chapter, the relative success or otherwise of ANNs and GP in fields related to ground compaction is examined to assess their utilisation in solving real-world problems.

Chapter 4 presents the first of the publications resulting from this research and involves the development of an ANN model based on CPT data. The paper first describes the CPT database, which consists of data related to both the pre- and post-compaction conditions that were compiled from previous ground improvement projects associated with RDC. Following this, the details of the ANN model development in the form of multi-layer perceptrons (MLPs) that are trained with the back-propagation algorithm are presented. The performances of the resulting models are compared and the selection of the optimum model is then discussed. In addition, a parametric study that ensures the generalisation ability of the selected optimum model, followed by the formulation of design equations based on the optimum model parameters, are described. Finally, modelling issues and issues encountered throughout the study are discussed.

Chapter 5 presents the second publication from this research, which involves the application of ANNs for predicting the impact of RDC using DCP test results. The paper initially describes the compilation of the database in order to provide an appreciation of the data that are included in the ANN modelling. This database provides information relating to the soil properties and strength data in terms of DCP test results that were compiled from a range of ground improvement projects involving the 4-sided impact roller. In this chapter, the steps involved in the ANN modelling process are outlined with a comprehensive description of each stage, together with available options. The results of the model optimisation are presented, followed by the behaviour of the optimal network when assessed for robustness using a parametric study. Numerical equations that facilitate the dissemination and deployment of the optimal ANN model to predict the level of ground improvement derived from RDC are detailed. Subsequently, a sensitivity analysis, which identifies the relative importance of the factors that are significant to ground improvement predictions in different soil conditions, is presented.

Chapter 6 is comprised of the third publication resulting from this study and details the application of GP for predicting the effectiveness of RDC using CPT test results. The emphasis of this paper is placed on a particular variant of GP, namely

linear genetic programming (LGP). The same database as described in Chapter 4 that comprised in situ soil strength data in the form CPT results is again utilised. The details of the optimal LGP-based model are then presented along with the performance analysis results. Later in this chapter, the robustness of the optimal model is investigated using a parametric study, followed by the explicit formulation of an LGP-based numerical equation. Subsequently, a comparison of the results of GP modelling with those of the previous ANN modelling is presented followed by a thorough assessment of the reliability of the developed models.

Chapter 7 presents the fourth publication, which implements the use of GP for the predictions of effectiveness of RDC using DCP test results. The same DCP database that presented in the Chapter 5 is employed for the development of the LGP-based model in this paper. The results of the optimal LGP model, along with a comparison with those obtained from the existing ANN model, are summarised. Following this, the details of parametric study and sensitivity analysis are discussed. In addition, the selected optimal LGP model for predicting the effectiveness of RDC is presented in terms of the C language computer code.

Chapter 8 presents a summary of this research that synthesises the optimal AI models discussed in the previous chapters. The optimal ANN and GP models, with respect to CPT and DCP data, are compared with each other and the most feasible approach relevant to each dataset is recommended for future applications. The discussion of the models obtained using the ANN and GP techniques for the CPT and DCP databases has also been included in a conference paper, which is presented separately in Appendix C. Later in this chapter, a set of general guidelines for developing ANN and GP models are established based on the experience gained from the present research. Finally, recommendations for future work are presented.

Chapter 2.

Literature Review:
Rolling Dynamic Compaction

2.1 INTRODUCTION

This chapter provides the background information and a review of the relevant literature relating to rolling dynamic compaction (RDC). The following aspects of RDC are examined: its equipment and operation, its characteristics and applications, and its advantages and limitations. In addition, a detailed assessment of previous studies that have quantified the efficacy of RDC is provided followed by a treatment of the field trials and in situ test methods that are currently utilised in practice. Each of these aspects is relevant to the analyses and developments that follow in later chapters.

2.2 GROUND IMPROVEMENT

An increasing proportion of construction challenges geotechnical engineers to use land optimally, even land previously considered to be geotechnically unsuitable due to poor soil conditions, such as expansive soils, soft or compressible soils, non-engineered fills, collapsing and softening soils. In such circumstances, these problematic soils need to be treated using ground improvement techniques. Ground improvement is basically considered as the controlled alteration of the state, nature or mass behaviour of ground materials (Mitchell and Jardine, 2002) in order to achieve a satisfactory behaviour of the ground in terms of strength, bearing capacity, settlement and lateral stability. Over the years, a number of different ground improvement methods have been developed, which can be categorised into densification, consolidation, weight reduction, reinforcement, chemical treatment, electro treatment, thermal stabilisation and biotechnical stabilisation (Munfakh and Wyllie, 2000). Among these, densification of the ground by means of soil compaction is by far the most frequent method.

Soil compaction is the method of increasing the soil density by means of mechanically applied energy. A rapid volume reduction takes place in the soil during the compaction process due to pore air expulsion, which results in particle rearrangement and sometimes crushing. Consequently, soil strength is improved, enhancing the stability and stiffness, whereby the post-construction and differential settlements of the ground

are reduced. At the same time, a denser state of soils reduces the hydraulic conductivity (Avalle, 2004d) so that compacted soils are less vulnerable to water seepage, contraction and swelling.

A number of methods are widely used in ground compaction, which can be subdivided into two categories as static and dynamic compaction, based on the type of compaction force being applied. Static compaction is simply the densification of soils with the application of a downward force on the ground surface by the self-weight of the machine. The compactive effort in static compaction can be pressure and/or kneading. Some of the static compaction machinery include drum, pad foot, pneumatic multi-tyre and sheepsfoot rollers (Figure 2.1). In addition to the compaction of soil layers, static compaction rollers are also used for finishing operations as they are able to achieve a smooth ground surface (e.g. smooth drum rollers). On the other hand, in dynamic compaction, in addition to the equipment's self-weight, a kinetically-driven downward force is applied on the ground surface. Dynamic compaction makes use of heavy tamping, vibratory drums and plates, rammers, vibroflotation and rolling dynamic compaction (Figure 2.1). As these machines operate, the additional downward force, either by vibration or an impact mechanism, effectively rearrange the soil particles into a denser state. However, most of the static and dynamic compaction methods, except heavy tamping and rolling dynamic compaction, are only capable of compacting near surface soil layers, where the compaction depth is limited to between 0.2 m and 0.5 m (Scott and Jaksa, 2012).

2.3 ROLLING DYNAMIC COMPACTION (RDC)

Rolling dynamic compaction (RDC) is a soil compaction method that has become increasingly popular in the construction industry over the past few decades. This method employs a non-circular (3-, 4- and 5-sided) heavy module (6 to 12 tonnes) which is drawn behind a tractor and, which distinguishes RDC from other conventional compaction methods.

(a) (b) (c)

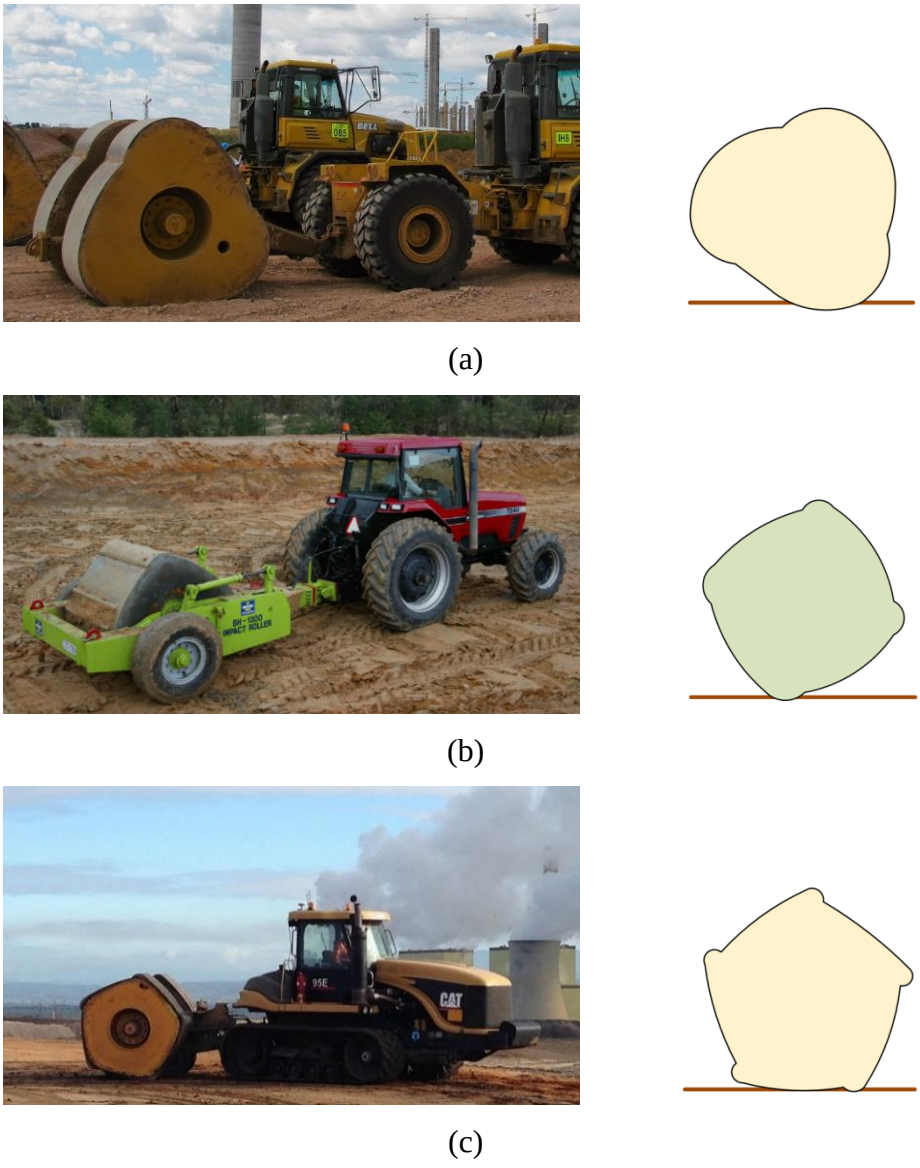
(d) (e) (f)

Figure 2.1 Various ground compaction methods: (a) drum rollers; (b) pad foot rollers; (c) pneumatic multi-tyre rollers; (d) rolling dynamic compaction; (e) heavy tamping; and (f) vibrating plates.

The innovative concept of a non-circular roller emerged during the 1930s, when the principles of soil mechanics were being formalised (Clifford, 1978b). Later, in the early 1950s, the first practical prototype impact roller became successful in the treatment of road subgrades underlain by collapsible sands in South Africa (Clifford, 1978b). Over the years, a number of advancements have been incorporated on the module design but its value was not fully realised until the mid-1980s. Since then RDC has become a more viable method amongst the various ground improvement techniques currently in widespread use. At present, RDC is commercially available in different compactor module designs, implemented worldwide, with two leading companies: Landpac Technologies Pty Ltd. employing 3- and 5-sided modules; and Broons Pty Ltd. manufacturing and operating the 4-sided module. Figure 2.2 illustrates the different module shapes of the RDC compactors in current use.

Broons Pty Ltd. (2016) manufactures and operates 4 variants of the 4-sided RDC module of RDC with specifications as outlined in Table 2.1. This study is focused on the 4-sided, 8 tonne 'impact roller' (BH-1300) which is, by a long measure, the most

often used in ground improvement projects. Since its introduction in Australia in 1984 this module has been applied to a wide variety of compaction work related to landfills, sub-base and subgrade compaction, as well as agricultural and mine site applications. As a result, a large amount of data is available from many project records that can be utilised to develop numerical models to predict the application of RDC.



(a)

(b)

(c)

Figure 2.2 Different impact roller modules: (a) 3-sided; (b) 4-sided; and (c) 5-sided.

Table 2.1 Broons module characteristics (Broons Hire, 2016).

Description		Compaction details					Dimensions			
Specifications	Model	Operating weight (ton)	Drum width (m)	Drum module weight (ton)	Lift height (mm)	Operating Zones*	Length (m)	Width (m)	Height (m) – module flat	Turning Circle (m)
	BH-1300	13.8	1.3	8	150	F, RF, SB, SG	4.5	2.61	1.5	12
	BH-1300 HD	18.2	1.3	12	150	F, RF, SB, SG	4.8	3	1.5	12
	BH-1300 MS	14.3	1.3	8.5	150	M, R, F	4.8	2.61	1.5	12
	BH-1950 MS	18.2	1.3	12	150	M, R, F	4.8	3.31	1.5	12
* F = Fill, M= Mining, R = Rock, RF = Refuse, SB = Sub-base, SG = Subgrade										

2.3.1 Characteristics of the 4-sided Impact Roller

The 4-sided impact roller module consists of a steel casing which is completely filled with concrete to produce the non-circular solid mass with rounded corners, as shown previously in Figure 2.2(b). A feature of the 4-sided impact roller is that it incorporates a double-linkage spring system, which again can be seen in Figure 2.2(b) and is connected to the impact roller frame and the module's axle. The double-spring linkage system provides additional energy to initiate rolling and compact the ground. When the impact roller traverses the ground, the module rotates eccentrically about its corners and falls to the adjacent face of the square-shaped mass resulting in a series of high amplitude impact blows delivered onto the ground at a low frequency of 90 to 130 blows per minute (Pinard, 1999). As the module undergoes the lifting phase, the spring system compresses and when the energy stored in spring system is released, the module is accelerated to the downward phase, whereby an impact blow is imparted onto the ground. Thereafter, the impact roller mass remains at rest on the ground surface for a brief amount of time (dwell-time) allowing the double linkages to relax, whilst the tractor moves forward relative to the compactor module (Clifford and Bowes, 1995). As such, the revolution of the roller mass continues and impact blows are delivered onto the ground at regular intervals. In addition, to enable the roller to be

transported to and from the site without damaging adjoining pavements and infrastructure, the roller is fitted with a hydraulic ram that lifts the module clear of the ground.

2.3.2 Energy Transformation and Improvement in Soil Density

During the compaction process, the compactive effort is derived from three sources: (1) main potential energy from the static self-weight of the module; (2) additional potential energy from being lifted about its corners; and (3) kinetic energy developed from being drawn along the ground. However, the computation of the energy transferred to the ground essentially requires several factors to be rationalised including: the indentation effect at the leading corner of the module; possible losses due to friction forces; and effects of the spring system (Clifford and Bowes, 1995; Avalle, 2004d). Nonetheless, it has been identified that with the impact roller mechanism has the potential to fully transfer the kinetic energy to the ground as indicated by the cessation of the drum motion immediately after an impact blow (Avalle, 2004d). Additionally, several field studies have investigated the amount of energy transferred to the ground by the action of impact rolling. For example, Avalle et al. (2009) conducted a study to quantify the effects of the 4-sided impact roller in terms of the energy imparted into the ground and to investigate the ground response. The testing regime comprised load cells and accelerometers embedded in ground along the impact roller path, whereby pressures and ground decelerations were measured. In this investigation, difficulties were encountered in controlling the reproducibility of the impacts, as it is almost impossible to ensure that the module strikes the exact same location on the ground, with each pass. Nevertheless, the study demonstrated that the pressure distribution beneath the impact module is non-uniform.

As described earlier, due to the combination of potential and kinetic energy derived from the impact mechanism, together with the large mass of the module, RDC produces a greater amount of compactive effort than traditional (i.e. non-dynamic) compaction techniques. Consequently, the soil beneath the surface is densified into a state of lower void ratio by expelling the pore air and fluid, as mentioned above. The

process of densification by RDC was explained by Clifford (1978a). As the impact blows are delivered onto the ground surface, a denser band of soil is created immediately below the ground. The thickness of this band varies depending upon a number of factors, such as soil type, moisture content, the number of roller passes, and this will be discussed further later in this chapter. With the compaction is in progress, this denser band starts to expand in thickness. When it reaches the surface level, this soil band acts as a mass of dense material, where the energy waves resulting from the impact blows are radiated to the deeper layers in the ground. This phenomenon is illustrated in Figure 2.3.

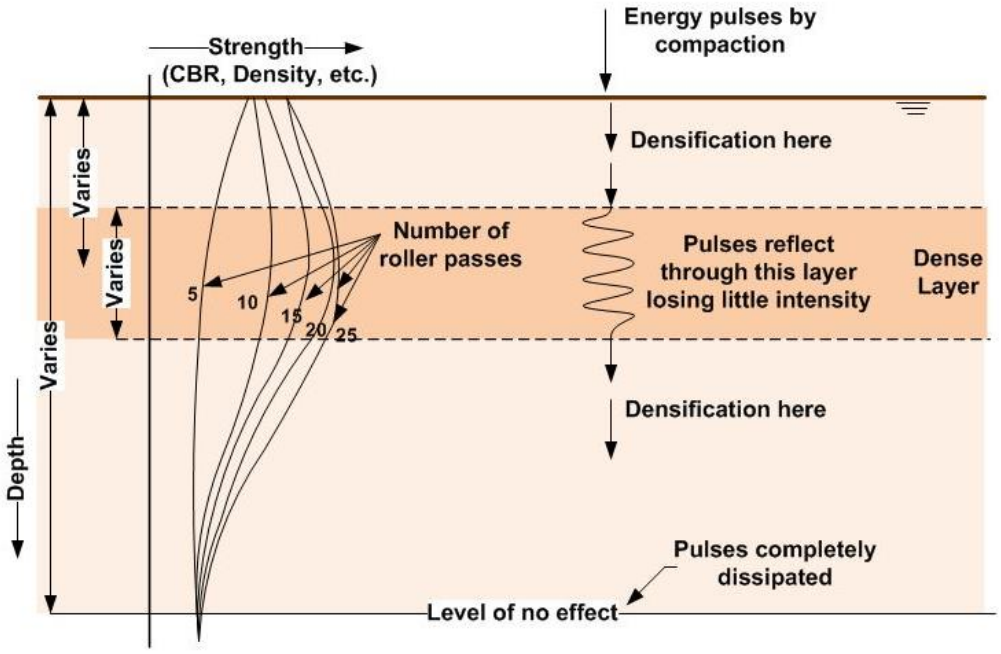


Figure 2.3 Dense layer build-up by impact compaction from Clifford (1978a).

2.3.3 Comparison of Impact Compaction with Traditional Counterparts

As a result of the greater amount of energy imparted to the ground from the impact mechanism, RDC often leads to a deeper influence depths; i.e. in excess of 3 m below the ground surface in some soils (Avalle and Carter, 2005), which is much deeper than conventional static and vibratory compaction (Clegg and Berrangé, 1971; Clifford,

1976, 1978b), where the influence depths are generally less than 0.5 m below the ground. This was demonstrated by Pinard (1999) using a nail/hammer analogy, as illustrated in Figure 2.4. It is evident that RDC accounts for high load intensity and a large surface area of the compactor module in comparison to the static and vibratory compactors, coupled with a high amplitude/low frequency operation mode which results in RDC developing deeper influence depths than those of other conventional compaction methods (Pinard, 1999).

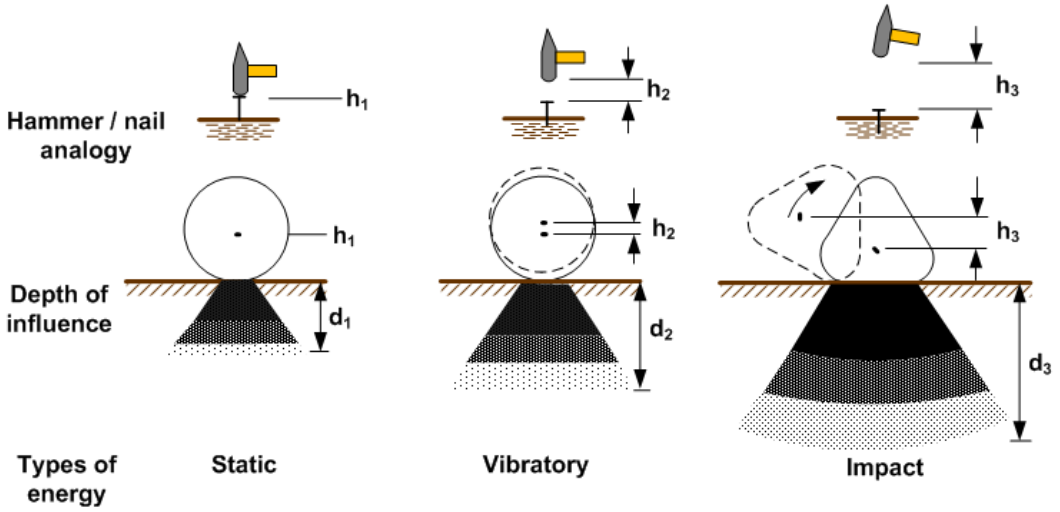


Figure 2.4 Effectiveness of different compaction methods with respect to the nail/hammer analogy [modified from Pinard (1999)].

The effectiveness of RDC has been compared with traditional methods in several previous trial studies in South Africa. For example, Clifford (1976) conducted a comparison of the performance of impact rollers and vibratory rollers (4.5 and 9 tonne static mass) on a site filled with marine sand. He compared the performance data in terms of surface settlements corresponding to 50 roller passes, as illustrated in Figure 2.5, and concluded that the impact roller influences the ground to a depth at least 3 m, whereas the depth influence is approximately 1.8 and 2.2 m below the ground for 4.5 and 9 tonne vibratory rollers, respectively.

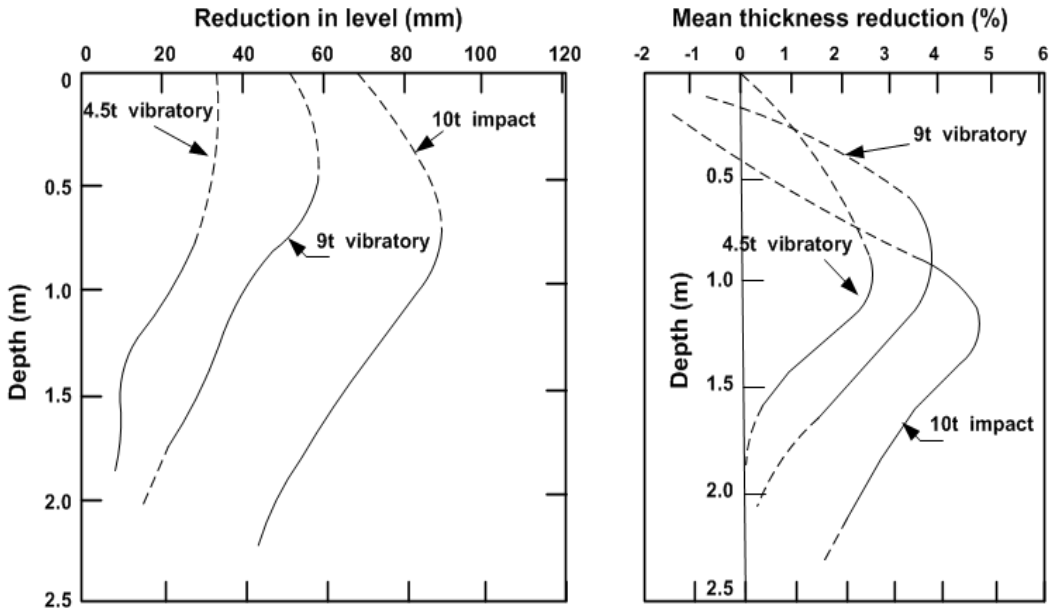


Figure 2.5 Comparison of impact compaction with vibratory rollers - level reduction in hydraulic fill of marine sand [modified from Clifford (1976)].

As described by Pinard (1999), the following graph (Figure 2.6) compares the in-depth effectiveness in terms of the dynamic stiffness of the soil with respect to several impact roller modules together with a conventional vibratory roller. It is evident that impact roller modules result in higher compacted layer stiffness than the vibratory roller.

Furthermore, RDC is capable of inducing significant settlements over poorly compacted ground. In an earlier study conducted by Clegg and Berrangé (1971), the settlements imposed by a pentagonal shaped impact roller, with respect to the number of roller passes, were compared with those obtained from several other conventional static and vibratory rollers. As demonstrated in Figure 2.7, it is evident that the impact roller induces significantly greater subgrade settlements when compared to the other roller types.

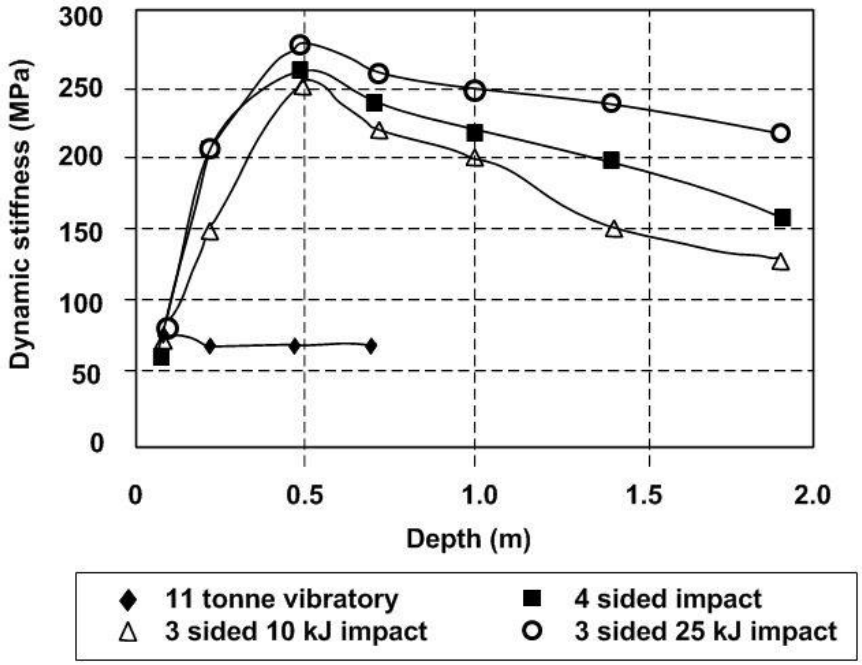


Figure 2.6 Variation of dynamic stiffness of soil from different compaction methods [modified from Pinard (1999)].

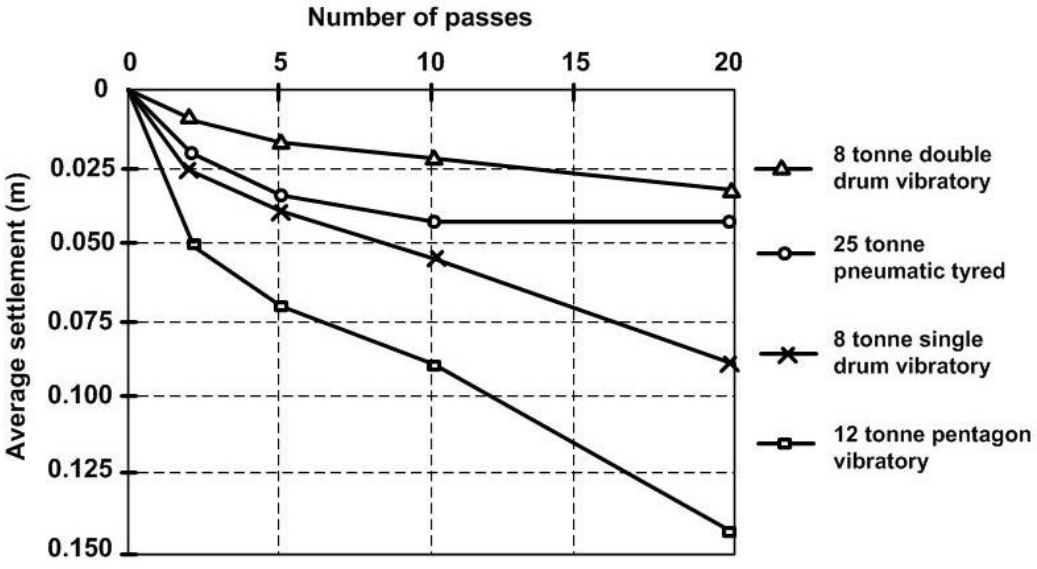


Figure 2.7 The variation of average settlement induced by different roller modules [modified from Clegg and Berrangé (1971)].

2.3.4 Advantages, Applications and Limitations

The major benefit of implementing RDC is its effectiveness; the capability of influencing the ground to a greater depth, i.e. in excess of 3 m below the ground surface in some soils, when compared to conventional soil compaction methods. The effectiveness is discussed in detail later in this chapter. Moreover, the economics of the use of RDC is also favourable since the module traverses the ground at a speed of between 9 to 12 km/h, which is substantially greater than traditional static and vibratory rollers that operate at approximately 4 km/h (Pinard, 1999). This creates approximately two module impacts over the ground each second. Thus, RDC is particularly efficient when employed in large and open sites. For example, case studies conducted by Avalle and Carter (2005) and Bouazza and Avalle (2006) showed that the use of RDC resulted in a cost-effective and environmentally sustainable approach, where it was required to improve a large open area (approximately 2.2 ha) overlying a former landfill. These inherent characteristics of RDC make it very effective for many compaction applications. Moreover, it also appears that prudent use of RDC can provide significant cost savings in the civil construction sector. Thus, it has proven to be successful in a variety of applications within Australia and worldwide, particularly in civil, mining and agricultural applications, as discussed below.

The deeper compactive effort and faster operating speed make RDC very effective for bulk earthworks. Thus, RDC is extensively used for the in situ densification of existing fills, such as on brownfield sites or former industrial land, raised land, landfills and earth embankments. The utilisation of RDC for these applications can be widely seen in developments related to Australian metropolitan areas, as well as in the United Kingdom, Hong Kong and the Netherlands (Avalle, 2004a). Moreover, RDC is frequently adopted in land reclamation projects, where usually dredged sand is compacted to form a densified soil base of significant thickness in order to reduce differential settlements. In addition, RDC is found to be beneficial for landfill treatment due to its capability of effectively compacting the extremely variable, heterogeneous waste fills that generally lead to substantial differential settlements (Avalle and Carter, 2005; Bouazza and Avalle, 2006).

Impact rolling is also extensively used for subgrade compaction, where density improvement and a more uniform subgrade can be achieved from the ground surface, without excavation of material. It can compact thicker lifts, in excess of 500 mm (Scott and Jaksa, 2012), which is considerably greater than the traditional layer thicknesses of between 200 mm and 500 mm (Avalle, 2006) and also it can operate with larger particle sizes. Furthermore, RDC facilitates subgrade preparation by inducing significant settlements over poorly compacted ground. In addition, RDC is used for proof rolling, as it is readily able to identify soft spots. Sealed or stabilised base courses, even concrete bases in existing pavements, can also be fragmented efficiently using impact rolling, leaving material suitable for sub-base where new base and wearing courses can be laid without further earthworks (Avalle, 2004a). There have been many earthworks and pavement construction related projects that have been carried out successfully using the square impact roller, such as the reconstruction of rural roads in South Australia, Adelaide Airport and the Port River Expressway (Avalle, 2004a; Avalle and Grounds, 2004).

RDC is also useful for a number of applications in the mining industry, including the construction and improvement of haul roads and cross-falls in mine sites, as the high speed and deep compaction effect facilitates the use of thicker layers and coarser materials. In addition, RDC has been productively utilised for construction of tailing dams (Avalle, 2006), compaction of the capping over waste rocks (Scott and Jaksa, 2012), compaction of bulk earthworks of mine spoil materials and to induce fracturing of the surface layers in rock quarries, instead of using drill-and-blast techniques (Avalle, 2006). Moreover, significant value has been demonstrated in rubblising rocky materials in mine haul roads and tip heads to create smaller particles, thereby extending the life of tyres of mine haul trucks, which are particularly costly to replace. Some of the applications of RDC are illustrated in Figure 2.8.

Impact rollers have also been shown to be beneficial for agriculturally related applications. Due to RDC, the relative density of the ground is improved and at the same time the material's permeability is also reduced as a result of reduction of the pore air volume. Thus, RDC is being used for in situ density improvement of the bases

of water storage facilities and canal banks, whilst reducing infiltration, so as to minimise adverse environmental effects (Avalle, 2004d). Moreover, reduction in ground permeability facilitates water retention, which is significant in preserving limited water resources. Several field based studies, for example, Auzins and Southcott (1999), has investigated the potential benefits that arise from the use of impact roller in these applications. In addition, (Avalle, 2004d) reported several case studies, where infiltration tests are conducted in conjunction with impact rolling to quantify the soil permeability with respect to number of impact roller passes. The results showed a significant reduction in infiltration rate into the ground after a few roller passes. For example, in channel improvement works carried out in Marthaguy Irrigation in NSW, the infiltration rate of 210 mm/hour was reduced to a no measurable rate after 1 hour after only 5 impact roller passes.

(a) (b)

(c) (d)

Figure 2.8 RDC applications: (a) compaction of subgrade; (b) proof rolling and demolition of existing pavements; (c) rubbilising rocky materials in mine haul roads; and (d) mine haul road maintenance.

As presented above, there are many benefits of RDC, however, there are also several limitations of adopting this technology. One is the disturbance to the ground surface layer, which occurs as a consequence of the action of the module. After RDC, the ground surface becomes undulating and the upper 100 to 200 mm, approximately, is loosened due to the kneading and shearing effect caused by the shape of the module (Avalle, 2004b, 2006). Most of the displacements occur vertically, but lateral displacements are also likely to occur around the area of impact (Clifford, 1978b). Thus, the top surface layer is not optimally compacted, when using RDC and often a grader and a drum roller are needed to finish the ground surface. However, on the other hand, the resulting surface corrugations provide a measure of interlock between adjacent soil layers, which helps to overcome lateral shearing effects.

In addition, the module travel speed is also a consideration. As mentioned, the square impact roller is typically towed at an optimal speed of 9 – 12 km/h. This creates approximately two module impacts over the ground each second (Avalle et al., 2009). However, slower speeds of less than 9 km/h cause the compactor module to slide over the ground without turning about its corners due to the insufficient momentum, whilst at speeds above 12 km/h, the module tends to bounce causing uneven compaction, as well as the driver discomfort. This may also cause damages to the damping system and the mechanical components of the roller (Avalle et al., 2009) and undue fatigue to the tractor. Given these reasons, RDC is usually constrained to occupy in open spaces. When it is used on small and complex earthworks sites, RDC becomes inefficient and ineffective because of the inability of the module to achieve the speed requirements to operate in full momentum (Clifford, 1978b) and because of the required turning circle of approximately 10 m due to the equipment and the speed of operation.

Because of the not insignificant ground vibrations that are generated by the impact roller operation, potential hazards exist to nearby sensitive structures and buried services in built up areas. In such situations, continuous monitoring of ground vibrations is usually specified. Avalle (2007a) suggested that a buffer zone of at least 2 to 5 m be adopted for industrial buildings, whilst a 7 to 20 m zone be used for residential buildings during impact rolling with the 8 tonne, 4-sided module.

2.3.5 Effectiveness and Zone of Influence

There are several instances in the literature where RDC has been examined experimentally through field-based studies with the intention of investigating the degree of densification and the extent of influence depth in different ground conditions. Some of the recent research studies that have evaluated the effectiveness of the 4-sided impact roller (BH-1300) are briefly discussed below in chronological order.

Avalle and Carter (2005) evaluated the ground improvement caused by impact rolling from a field study conducted in an industrial site in Banksmeadow, Sydney consisting of Botany sands, which were described to comprise uniformly graded fine-to-medium sized quartz grains. In investigating the depth of influence, the results of the cone penetration test (CPT) were compared with respect to number of roller passes. As a further measure, Avalle and Carter (2005) analysed the CPT results in terms of the improvement index for densification (I_d) as per Equation 2.1, which utilises the specific energy of CPT penetration represented by cone tip resistance (q_c), where q_{cf} and q_{ci} are the values of q_c after and before compaction, respectively.

$$I_d = \left(\frac{q_{cf}}{q_{ci}}\right) - 1 \qquad (2.1)$$

Based on the comparative results of I_d obtained from this study at 3 locations after 5 and 20 passes, as shown in Figure 2.9, Avalle and Carter (2005) confirmed that there was an improvement in strength after impact rolling to depths of at least 3 m below the surface. At the same time, a degree of inconsistency was observed in the degree of improvement at Location 1, which is likely the result of natural variations in the ground.

Bierbaum et al. (2010) also attempted to investigate the influence zone of the impact roller in an uncompacted area of land in Gillman, South Australia. In their study, the research correlated the influence depth of RDC with the energy transferred during impact rolling in terms of vertical stress measurements, assuming minimal ground disturbance. Based on the measurements from earth pressure cells (EPCs) that were

embedded in the ground at different depths, they suggested that the influence depth with respect to 20 passes of the impact roller continued below 1.8 m. However, as the deepest embedment of the EPC was limited to a depth of 1.8 m, this study was unable to capture the influence of RDC extending deeper into the ground. Nonetheless, Bierbaum et al. (2010) conducted several additional CPTs and observed a noticeable improvement in the cone tip resistance measurements with respect to the number of roller passes and thus, verified that the influence depth of RDC could extend up to 3 to 3.5 m below the ground.

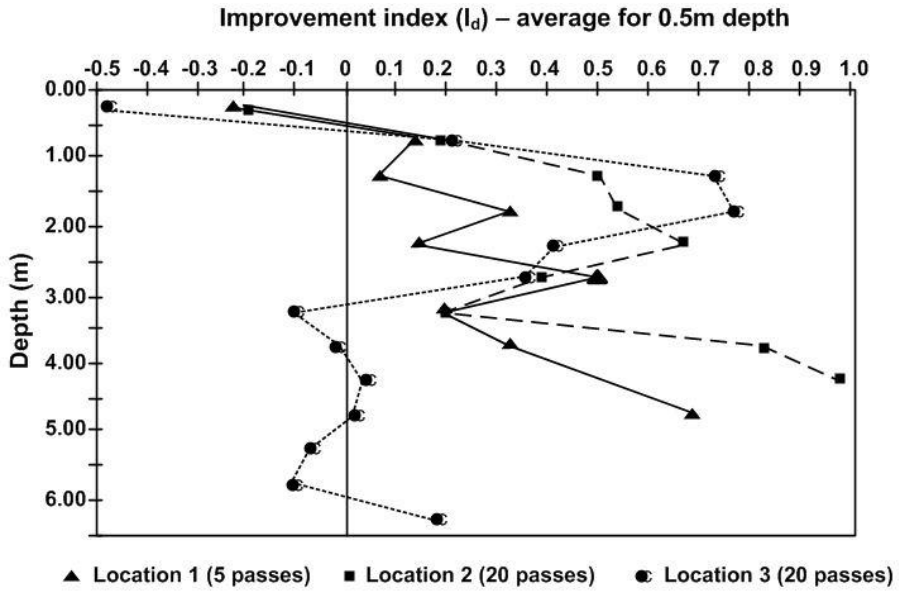


Figure 2.9 Variation of I_d with respect to different number of impact roller passes [modified from Avalle and Carter (2005)].

In addition, Jaksa et al. (2012) quantified the influence zone, i.e. both the vertical and lateral extent of the 4-sided impact roller, by means of a field-based case study. However, this study was considered to be an extension of the field work previously undertaken by Avalle et al. (2009), where the effectiveness of RDC was investigated in terms of the energy imparted into the ground measured using load cells. The field trial was conducted at the Iron Duke mine in Whyalla, South Australia and involved a series of EPCs embedded in the ground at different depths below the surface. As illustrated in Figure 2.10, the superimposed EPC data demonstrated a deterioration in maximum dynamic pressure as the depth increases and thus, Jaksa et al. (2012)

confirmed that the influence of impact blows decays with the depth, as expected. Moreover, the results of the averaged peak pressure measurements of EPCs at different depths obtained in this study (Figure 2.11) ensured that a significant and quantifiable improvement in ground has occurred mostly within the upper 2 m beneath the ground surface. This study also observed that the influence depth extends beyond 3.85 m below the ground.

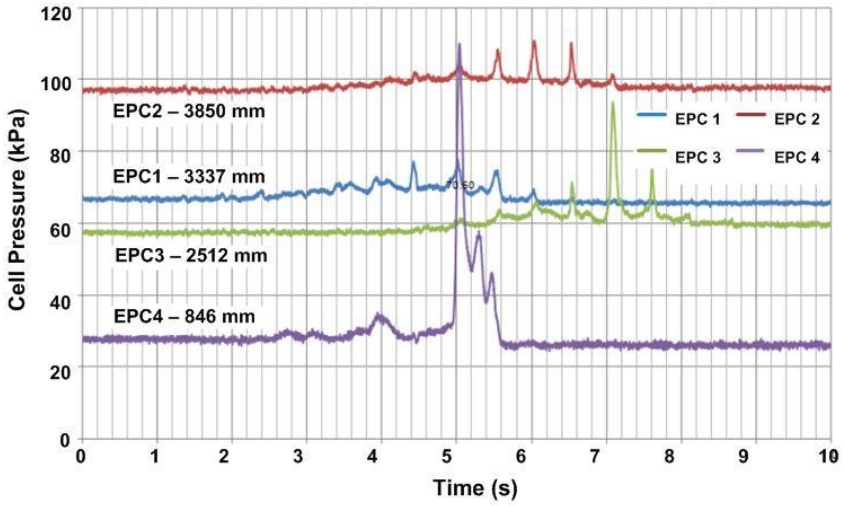


Figure 2.10 Super-imposed results of EPCs embedded at different depths as obtained by Jaksa et al. (2012).

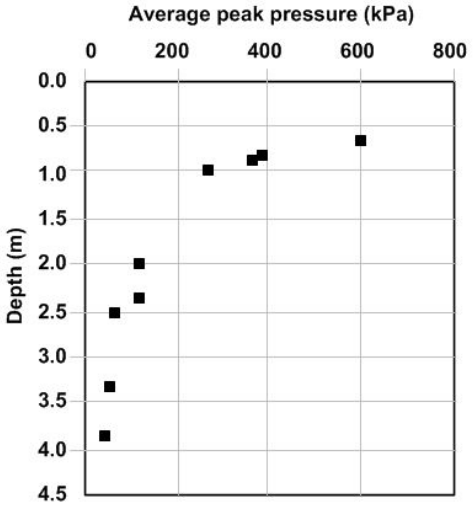


Figure 2.11 Resulting average peak pressure variation as obtained by Jaksa et al. (2012).

More recently, Scott and Jaksa (2014) conducted a field study involving a compaction trial that investigated both the vertical extent and lateral zone of influence of RDC again by means of the CPT. The site was described to be comprised of predominantly quartzose and carbonate sand (thickness of the compacted fill was 1 m) underlain by natural soil of stiff to hard silty clay. Based on the measurements of CPTs undertaken to a minimum depth of 2 m, they found that the influence depth extended to at least 1.75 m below the surface after 10 impact roller passes. In addition, they found that the nature of the profile obtained from the CPT measurements, i.e. q_c versus depth (Figure 2.12), was unlikely to be attributed to the RDC compaction trial, however, it was suggested that the results may have been influenced by sensing the soil interface ahead of the cone tip.

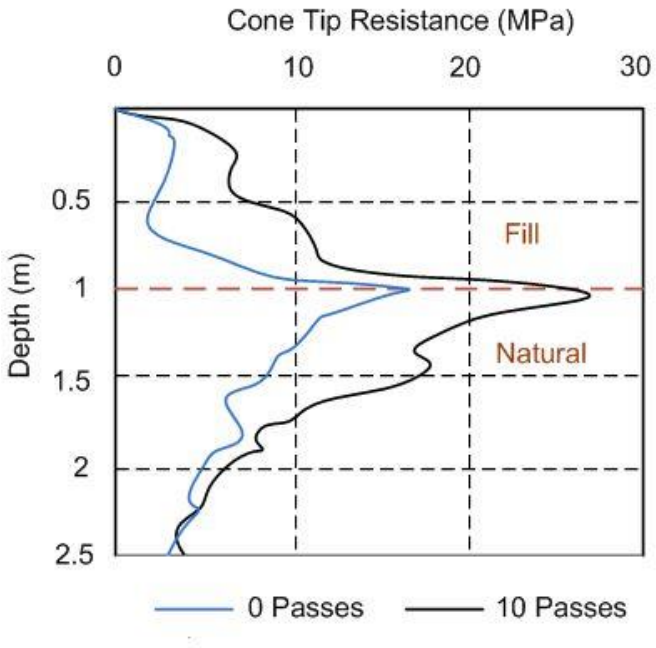


Figure 2.12 The variation of cone tip resistance with depth before and after 10 passes of impact roller as obtained by Scott and Jaksa (2014).

In addition to the above field-based research studies, to date many field trials have been conducted prior to full-scale compaction associated with various construction projects to ascertain the site-specific operational parameters, especially the number of

roller passes to satisfy the project specifications. The details of these field trials are discussed later in this chapter. However, comparatively, there has been very little research directed so far towards the application of numerical modelling for evaluating the effectiveness of RDC. To date, Kim (2010), Bierbaum et al. (2010) and Bradley et al. (2012) have investigated the influence zone of the impact roller using the finite element method (FEM).

In particular, Kim (2010) used the FEM to account for the depths of influence of RDC with respect to different shapes of the impact roller module, i.e. cylindrical, triangular, Landpac (3-sided module), pentagonal and octagonal shapes having an identical drum weight of 12 tonnes. This study considered the depth of influence to be the depth where the vertical stress decreases to one-tenth of the stress at the surface. The numerical model demonstrated that the width of the contact area between the roller module and the soil is directly related to the influence depth, in addition to the higher energy induced by impact blows. Thus, based on the comparison of results, as shown in Figure 2.13, this study affirmed the efficacy of utilising non-circular rollers for in-depth compaction over cylindrical shaped modules. However, this numerical simulation did not account for the effects of the water table, different soil types and the wave absorption and reflection effects, which restricts its wider acceptance. In addition, the study adopted a relatively simple soil constitutive model which was not calibrated against field-based measurements.

The research conducted by Bierbaum et al. (2010), in their final year undergraduate research study, simulated the stress distribution in the ground resulting from the 4-sided impact roller, both statically and dynamically. Static analysis accounted for the stresses induced by the self-weight of the roller module, whilst the dynamic analysis considered the pressure induced by the impact blows. However, the software used in this study, i.e. Midas GTS, incorporated very limited capabilities of dynamic FE analysis and restricted the simulations of multiple strikes of the roller module. However, based on their findings, Bierbaum et al. (2010) confirmed that dynamic effects induce greater influence depths than those obtained by static compaction.

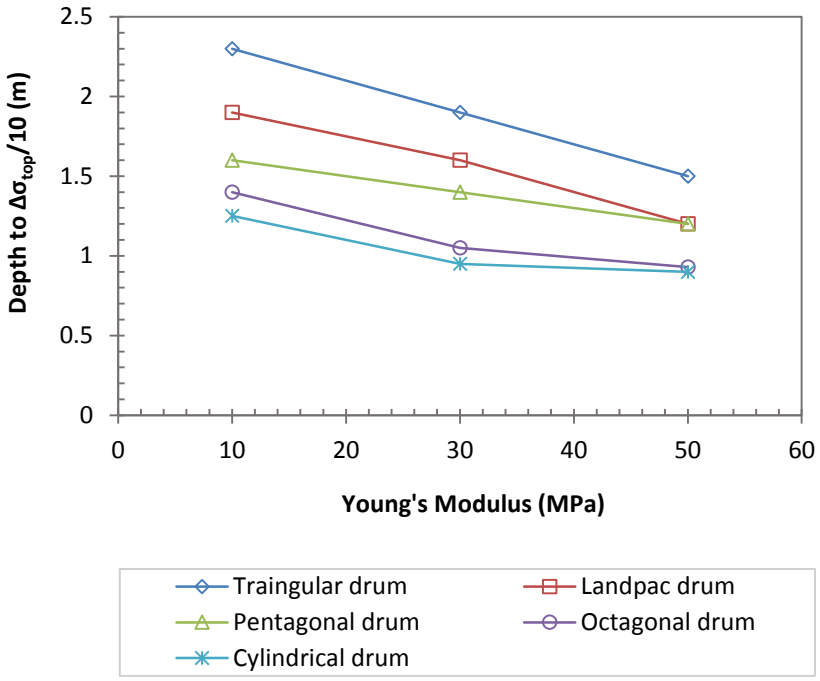


Figure 2.13 The depth of influence of different drum shapes as obtained by Kim (2010).

Bradley et al. (2012) also simulated the stress distributions obtained from impact rolling, as well as static loading. In quantifying the effectiveness of RDC, this study proposed a redefinition of the depth of influence, which is the depth of soil affected by the load imposed at the ground surface; generally, but certainly not uniformly, using 10% of the peak stress as the limit. However, Bradley et al. (2012) defined the *improvement depth*, which is the depth over which the soil undergoes significant improvement in density and shear strength due to RDC. The proposed FE model was found to predict soil stresses reasonably accurately for both static and dynamic conditions. As the FE model was validated against the field data obtained by Mentha et al. (2011), it was shown that RDC was most effective for depths of 0.8 to 3 m below the surface, where the soil density increased with greater numbers of roller passes. However, these numerical simulations of RDC were found to have several limitations in the form of assumptions, errors in laboratory testing and modelling methodology.

In addition to the field-based studies and numerical simulations described above, a number of other researchers have attempted to study the influence depth in relation to other impact roller variants, i.e. 3- and 5-sided, and these are summarised in Table 2.2.

Table 2.2 Influence depths of different impact roller modules.

Research	Impact roller module	Soil type	Depth of influence
Clifford (1978a)	MK. II, IV, V 4 sided	Hydraulic sand and industrial rubble	~ 3 m
Kelly (2000)	3 sided Landpac roller	Botany sand	~ 4 m
Kelly (2000)	Combination of 5- and 3-sided Landpac roller	Botany sand	4 – 5 m
Avalle (2007b)	12 tonne, 4 sided Broons roller	Calcareous sand	1.5 – 7 m

2.3.6 Factors Affecting the Efficacy of Rolling Dynamic Compaction

It is evident from the RDC case studies described above, that the degree of densification or the extent of the influence depth in different ground conditions is unpredictable because it is likely to be affected by various site-specific factors and especially by the heterogeneous nature of the soil. In addition, the complex nature of the operation of the impact roller, as well as the consequent behaviour of the ground, has meant that the performance of RDC is complex. It has been identified that the degree of soil compaction depends upon key factors including: the inherent physical properties of the soil, such as dry density, moisture content, soil type and gradation; the thickness of the soil layer being compacted; and the compactive effort. These factors are briefly described below with the associated examples from several case studies.

2.3.6.1 Soil Type and Gradation

Among the compaction variables that are most influential on the degree of densification, soil type and gradation are more prominent. Soils used for earthwork compaction vary considerably from project to project, particularly in terms of their geotechnical characteristics and particle size range. However, one of the strengths of RDC is that it can work with a wide range of earthworks materials (Avalle, 2004b). Generally, filling or in situ subgrade material may consist of large oversized particles such as rock or rubble, or well or poorly graded coarse- or fine-grained materials, or sometimes organic materials. Although, it is an obvious advantage that the impact roller is able to work on a variety of soil types but, the estimation of its efficacy is challenging (Avalle, 2004b). Nonetheless, several studies have confirmed the RDC behaviour and its applicability on different types of soils.

For example, as reported by Avalle (2007b), impact rolling with the 4-sided, 12 tonne module, on a site consisting of a coarse-grained calcareous sand, caused significant improvement as measured by the CPT to a depth more than 7 m after 40 passes. In addition, a case study conducted by Scott and Jaksa (2012) at a mine site verified the efficacy of the 4-sided, 8 tonne impact roller on coarse-grained filling consisting of tailing materials representative of a well graded sand with some gravel (80% sand sized, 14% gravel sized and 6% clay sized). In situ density tests of this study verified the RDC performance (95% of maximum modified dry density) to be acceptable after 6 passes on 900 mm thick lifts or 10 passes up to 1,100 mm thick lifts. Kelly (2000) presented another two case studies based on field trials, i.e. one employed the 3-sided module and the other a combination of 3- and 5-sided modules, where both sites were found to be comprised of coarse-grained soils; i.e. natural marine sand and a reclaimed sand deposit, respectively. These case studies demonstrated a uniform increase in soil strength to depths of more than 4 m below the surface based on CPT results and thus affirmed the suitability of RDC for the compaction of weak, loose sands.

In addition to the above highlighted case studies on coarse-grained soils, the impact roller has the potential of compacting fine-grained materials, as RDC also involves, to

some extent, a kneading action. The applicability of impact rolling on fine-grained soils has been verified by several studies in the literature. For example, according to Avalle (2007b), 98% compaction was achieved after 15 passes of the 4-sided, 8 tonne module on a 700 mm thick fine-grained alluvial soil, i.e. mainly sandy clay/clayey sand. Another successful application of RDC on fine-grained soils was reported by Kelly (2000), which involved a field trial that employed the 3-sided impact roller on an uncontrolled, variable clay fill. In this case study, RDC was found to be effective 4 m below the surface with a significant improvement in shear strength in the fill material.

From the observations of field trials on different types of soils, it has been observed that RDC results in a significant increase in the in situ strength to a considerable depth in coarse-grained materials, more so than in fine-grained soils. This has also been observed in the numerical study conducted by Bierbaum et al. (2010), where a greater amount of stress was measured in coarser material than in fine-grained soil. This may be attributed to the fact that coarser materials drain better and the pore pressures dissipate far more rapidly than in fine-grained soils and thus, allowing the particles to rearrange more easily into a denser state. According to Terashi and Juran (2000) dynamic compaction works best on dry coarse-grained soils due to better water drainage than in fine-grained soils. Moreover, it has been specified that loose soils are susceptible to over compaction more readily in fine-grained silts and clays and thus, to overcome such situations, the necessary amount of impact energy (e.g. number of roller passes, and tamper weight and drop height) needs to be established in advance.

2.3.6.2 Compactive Effort

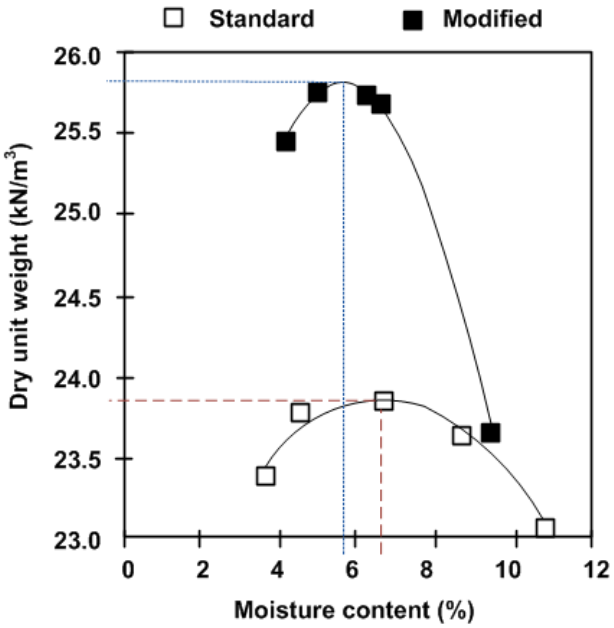
The compactive effort is also an important factor that affects the degree of soil compaction. Compactive effort is the amount of energy imparted to the ground, which can be considered as a function of lift thickness, number of roller passes, machine speed, weight and height of the drop. However, for a given compaction situation, most of these parameters are fixed for a particular module type and thus, the compactive effort is determined solely on the number of roller passes.

In a recent study conducted by Scott et al. (2012) the applicability of the standard Proctor compaction test (Standards Association of Australia, 2003b) and modified Proctor compaction test (Standards Association of Australia, 2003a) was examined with respect to RDC. The characteristics of these tests are compared in Table 2.3 below.

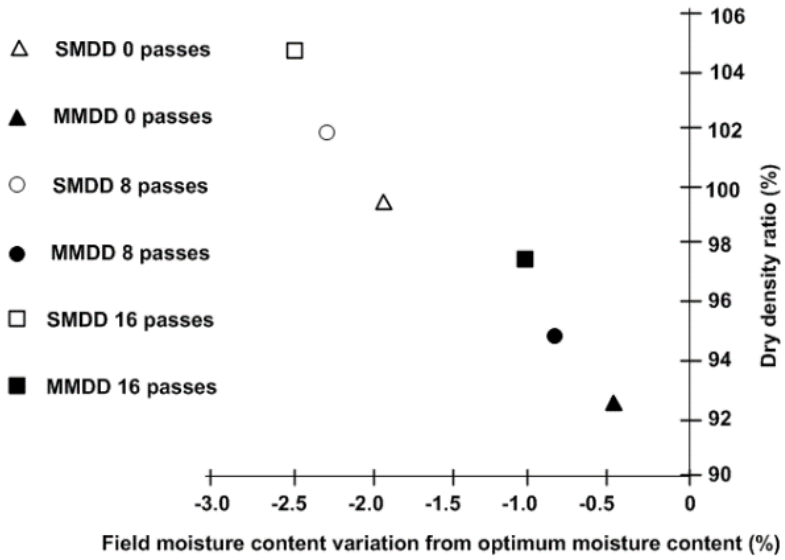
In this investigation, Scott et al. (2012) conducted field density tests on a site comprised of iron tailings (14% gravel sized, 80% sand sized and 6% clay sized particles) with respect to 0, 8 and 16 impact roller passes and the results were compared with standard and modified Proctor compaction test results. As shown in the Figure 2.14, the resulting variation in dry density ratio, with respect to the difference between the field and optimum moisture content, confirmed that the actual density versus moisture relationship is more represented by the modified test than the standard test. Thus, Scott et al. (2012) concluded that compactive effort of RDC is best measured by the modified Proctor compaction test. For this site condition, 16 passes of the impact roller represented 98% compaction, with respect to maximum modified dry density.

Table 2.3 Comparison of imparted energy for standard and modified Proctor compaction tests.

Test	Standard Proctor	Modified Proctor
Hammer weight	2.7 kg	4.9 kg
Drop height	300 mm	450 mm
Energy imparted per blow	7.94 J	21.62 J
No. of soil layers	3	5
No. of blows per layer	25	25
Energy imparted per unit volume	596 kJ/m^3	2703 kJ/m^3



(a)



(b)

Figure 2.14 Dry density vs optimum moisture content: (a) comparison of laboratory standard and modified Proctor results; and (b) comparison with respect to 0, 8, 16 roller passes [modified from Scott et al. (2012)].

[SMDD - standard maximum dry density, MMDD - modified maximum dry density]

2.3.6.3 Dry Density and Moisture Content

The moisture content and the dry density are also extremely important among the factors associated with compaction. It is well known that soil type, moisture content, compactive effort and dry density are related to one another via a series of compaction curves, as shown above in Figure 2.14(a). For example, as shown in Figure 2.14(b), the field trial conducted by Scott et al. (2012) indicated that a greater dry density can be achieved for lower moisture content with a higher compactive effort (represented by the number of roller passes) and this is beneficial in earthworks compaction.

However, as shown in the Figure 2.15, with the increasing amount of compactive effort in terms of number of blows, the gain in dry density decreases and the additional blows begin to have little or no effect on the dry density. Thus, it is essential to maintain the right amount of compactive effort that is both effective and efficient in terms of the optimal moisture content and maximum dry density.

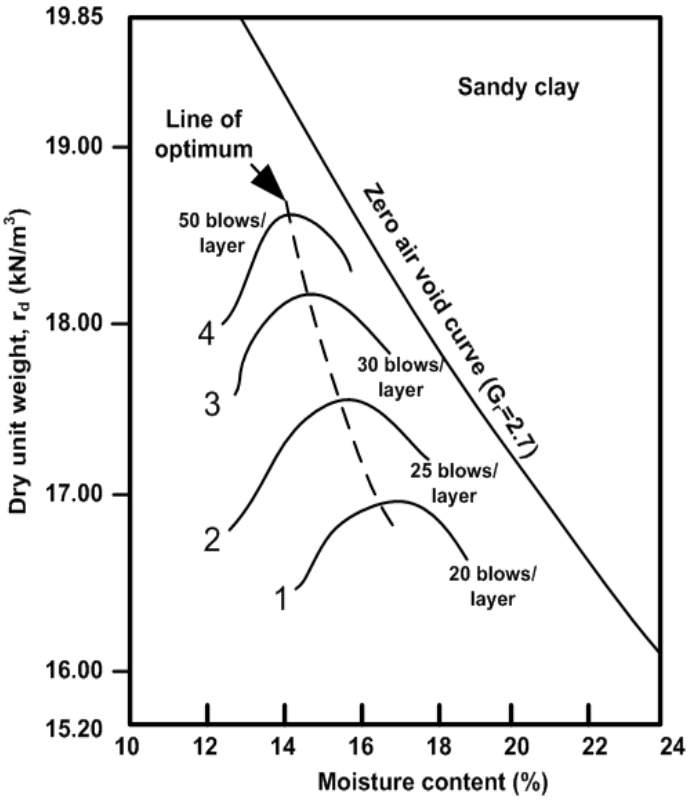


Figure 2.15 Variation of dry density with number of roller passes.

As RDC imparts comparatively greater energy to the ground than that by conventional rollers, the impact roller is capable of achieving the required dry density over a wide range of moisture contents. Nevertheless, RDC is less effective if the soil is too dry of the optimum moisture level (Scott and Jaksa, 2014). Broons Hire (2016) suggests that 2% below the laboratory tested optimal moisture content is the most effective. The reduced dependency of RDC on moisture content is advantageous and expands its applicability, especially in the agricultural sector in semi-arid areas in achieving a uniform subgrade.

2.3.7 Current Practice of Estimating RDC Effectiveness

To date, there exists no rational means, theoretical or empirical, for obtaining a prior estimation of the degree of densification or the extent of the influence depth by RDC in different ground conditions. Moreover, the complexity and the interrelationships of the factors described above, that affect the efficacy of RDC, are not yet fully understood. Subsequently, the performance design and application of RDC currently relies heavily on the geotechnical engineer's experience-based judgement. Moreover, field trials are often carried out on site prior to the full-scale compaction work.

2.3.7.1 Field Trials

A detailed field trial program is usually undertaken in advance of many earthworks projects to ascertain the relevant operational parameters, especially the optimal number of impact roller passes needed to achieve the required percentage of maximum dry density in a site. In another words, a field trial is often carried out to define the method specification that stipulates the methods to be used for the earthworks construction at a particular site (Avalle, 2004c). A number of detailed reviews on field trials are available in the literature (Barrett and Wrench, 1984; Avalle and Young, 2004; Avalle, 2007b; Nash, 2010; Scott and Jaksa, 2012).

In a typical field trial, a testing pad is arranged, which is representative of the full-scale operation of the compaction process. For example, Figure 2.16 illustrates a typical

schematic diagram of the testing pad adopted for the compaction trial conducted by (Scott and Jaksa, 2012), which was undertaken to construct a proposed tailings dam embankment at the mine site. As shown, the trial pad was sufficiently large with the dimensions of 120 m in length and 25 m in width, including ramp areas, so that the impact roller could achieve its optimal speed when it reached the middle section (25 m × 50 m) of the trial pad, where all the tests were undertaken. The test pad consisted of a single layer with various thicknesses, from 0.5 to 1.5 m, since one of the research objectives was to determine the lift thicknesses that yielded optimal compaction characteristics. In addition, this particular test pad accommodated 9 impact rolling lanes, comprising 3 separate zones of 10, 20 and 30 passes of the 4-sided impact roller.

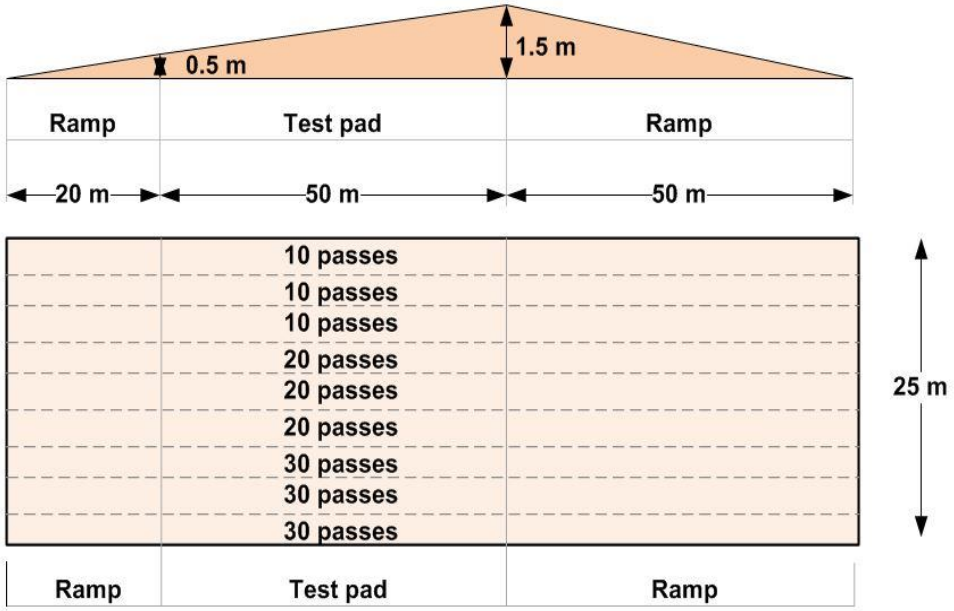


Figure 2.16 Schematic diagram of a trial pad
[modified from Scott and Jaksa (2012)].

The compaction trials often involve a testing or monitoring regime, which seeks to verify the efficacy of RDC with respect to the required number of roller passes and lift thickness to achieve a performance specification of minimum dry density or maximum surface settlement, for example. Thus, in the field trial conducted by Scott and Jaksa (2012), the verification of RDC involved a combination of surface settlement surveys,

soil sampling and in situ tests, such as penetrometer and field density testing, that were undertaken after different numbers of module passes. A detailed description of the test methods and measurements techniques involved in RDC assessment are described in the next section. Based on density test results obtained from different numbers of roller passes (Figure 2.17), Scott and Jaksa (2012) established a relationship between the dry density ratio and compactive effort, in terms of the number of roller passes and lift thickness. Thus, it was recommended to implement a minimum of 10 impact roller passes on layers of up to 850 mm in thickness or a minimum of 30 passes on layers of up to 1,000 mm thick, in order to achieve 98% compaction with respect to maximum standard dry density at this particular site.

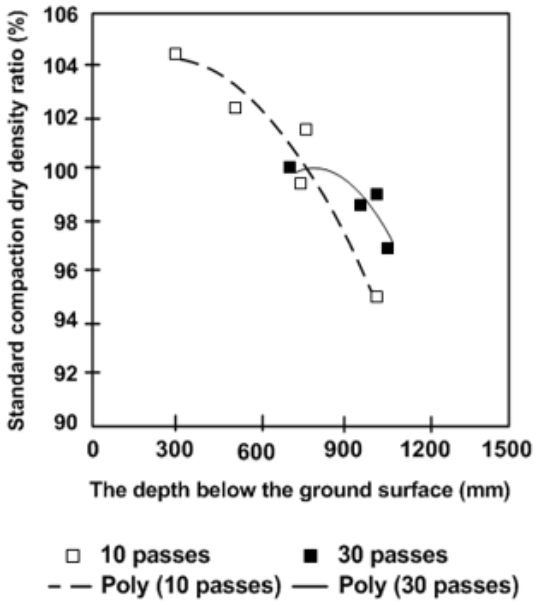


Figure 2.17 Results of density tests after different number of impact roller passes as obtained by Scott and Jaksa (2012).

As such, field trials are particularly useful as a technique for providing a suitable framework for site-specific methods for impact rolling with an appropriate validation strategy (Avalle and Young, 2004). Furthermore, a detailed and carefully structured field trial confirms the effectiveness and increases confidence in utilising RDC for a particular compaction project. However, field trials associated with RDC often incur a non-trivial cost and time, which is disadvantageous to its wider acceptance.

2.3.7.2 Test Methods and Measurement Techniques

As mentioned above, unlike in conventional compaction, which employs thin soil layers with relatively smaller particles placed in a relatively controlled manner, RDC compacts thicker lifts with larger particle sizes and is able to accommodate a wide range of material types, including crushed rock, rocky mine spoil and waste materials. Although these are potential benefits of RDC, it is challenging to specify an appropriate and cost-effective testing regime for the verification of efficacy of RDC (Avalle, 2004b, 2004c, 2007b) over such a wide range of ground conditions. However, an appropriate testing regime for field trials is usually chosen in relation to the nature of project and type of application, the site location and subsurface conditions, the availability of test equipment and, most importantly, budgetary constraints. The test methods are selected in collaboration with the geotechnical engineer and client and mostly depends on their preferences and previous experience with impact rolling and ground improvement projects. These test methods can be categorised as soil classification, surface settlement surveys, in situ density tests and geophysical tests. The most commonly used test methods and their applicability are briefly described below.

Soil classification testing

Classification of the soil is an important task before commencing any geotechnical engineering application, where a detailed soil investigation is performed to assess the underlying soil profile, its physical characteristics and engineering behaviour. The particle size distribution of the soil is determined by sieve analysis and, sometimes, hydrometer analysis tests, whilst Atterberg limit tests evaluate the consistency limits of the sub-grade material. Finally, these tests are also used to classify the soils into the specific Unified Soil Classification System classes.

Surface settlement monitoring

There are several techniques used for surface settlement monitoring. Total stations, for example, have been utilised in several field trials, such as the field trial for an industrial

project in Banksmeadow, Sydney (Avalle and Carter, 2005), are a fast and efficient method for obtaining surface settlement measurements before and after compaction over large tracts of land. In addition, direct and simple conventional methods, such as string lines and automatic level measurements are also used in practise, such as was adopted for the Adelaide Airport development (Avalle and Grounds, 2004). Although surface settlement monitoring is considered to be an effective means of assessing the efficacy of RDC, the surface irregularities left after the impact rolling is the major difficulty. In addition, whist surface settlement provides an indication of in situ density improvement, it is not possible to determine density improvement with depth, nor quantify density in any respect, due to lateral spreading and surface irregularities. In order to account for such surface undulations, the ground is sometimes trimmed using a grader prior to the surface settlement monitoring. Alternatively, only low spots are measured in subsequent passes. Level readings are generally measured in a grid over the site and are averaged with respect to different number of roller passes. The rate of increase of overall average settlements are indicative of effective refusal, which is described to be the number of passes at which no further significant measurable settlement occurs (Avalle and McKenzie, 2005).

In addition to these methods, Landpac technologies (Landpac International, 2016) utilises a proprietary technology for monitoring of compaction induced settlements. As described by McCann and Dix (2007), continuous induced settlement (CIS) technology involves a GPS and computerised based monitoring system, which collects the settlement measurements in a 2D plan during the impact rolling. This method adopts accelerometers mounted on the 3-sided module frame and the developers claim that it overcomes the difficulties associated with conventional methods and facilitates identifying localised highly compressible locations that are often available in sites with heterogeneous materials. However, given that the system is proprietary, extremely limited information is publicly available that describes the relationship between the CIS measurements and in situ density, nor how these measurements are influenced by soil type and ground conditions. The CIS system is essentially used to identify soft softs, as in proof rolling.

In situ density testing

In order to assess density improvement with depth, in situ density tests are conducted. These test methods can be categorised into two groups: destructive and non-destructive tests, depending on whether the testing procedure disturbs the ground or not. In addition, these methods either provide the in situ density directly or indirectly via correlation relationships.

Destructive test methods

The most commonly used in situ density test methods involve the use of penetrometers, which provide indirect measurements. The most frequent penetrometer techniques are the cone penetrometer test (CPT), the dynamic cone penetrometer (DCP), the Perth penetrometer test (PPT) and the standard penetrometer test (SPT). Penetration resistance data obtained from these test methods, with respect to before and after a number of roller passes, can be considered as a qualitative measure of the efficacy of impact roller (Auzins and Southcott, 1999).

The CPT in usually conducted in accordance with AS 1289.6.5.1 (Standards Association of Australia, 1999) and the DCP in accordance with AS 1289.6.3.2 (Standards Association of Australia, 1997), and these are described in detail in Chapter 4. The CPT provides a continuous profile of soil strength in terms of cone tip resistance (q_c) and sleeve friction (f_s), which together provided an indication of soil type by means of the friction ratio (R_f). Furthermore, the specific energy of penetration, in terms of the improvement index for densification (I_d), that is computed from the CPT profiles, has been found to be useful in the evaluation of the influence depth of RDC (Avalle and Carter, 2005) as discussed in §2.3.5. The CPT has been successfully implemented many times for the verification of RDC (Kelly, 2000; Avalle and Young, 2004; Avalle and Carter, 2005; Avalle, 2007b; Scott and Jaksa, 2014).

The DCP and PPT are the most commonly used in situ test methods, which provides an indication of soil strength in terms of rate of penetration (blows/mm). Both tests are similar in operation, however, the DCP incorporates a small conical tip at the end of

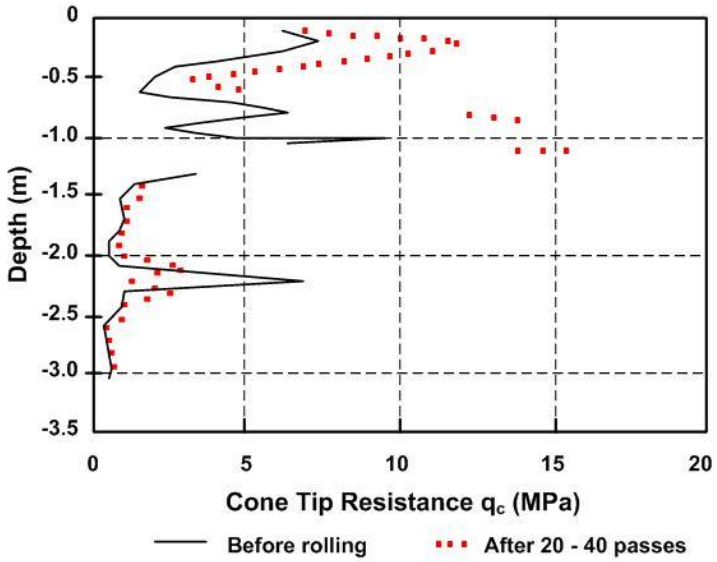
the testing rod, whereas the PPT employs a flat tip. However, both methods are very frequently implemented in RDC projects to quantify the improvement of soil strength with increased roller passes (Avalle and Carter, 2005; Avalle, 2007b; Scott and Suto, 2007; Jaksa et al., 2012).

The SPT is usually performed in accordance with the AS 1289.6.3.1 (Standards Association of Australia, 2004b) and involves driving a split sample tube into the ground at the base of a borehole via a series of blows from a 63.5 kg hammer dropping from a height of 760 mm. The total number of blows needed to penetrate the last 300 mm of 450 mm is reported as the SPT-*N* value. Although, the SPT method is incapable of producing a continuous profile of the soil like the CPT, a sample of soil is available for inspection. In addition, equipment and procedural irregularities lead to significant uncertainties with the measured *N*-value, which is also the case for the DCP/PPT.

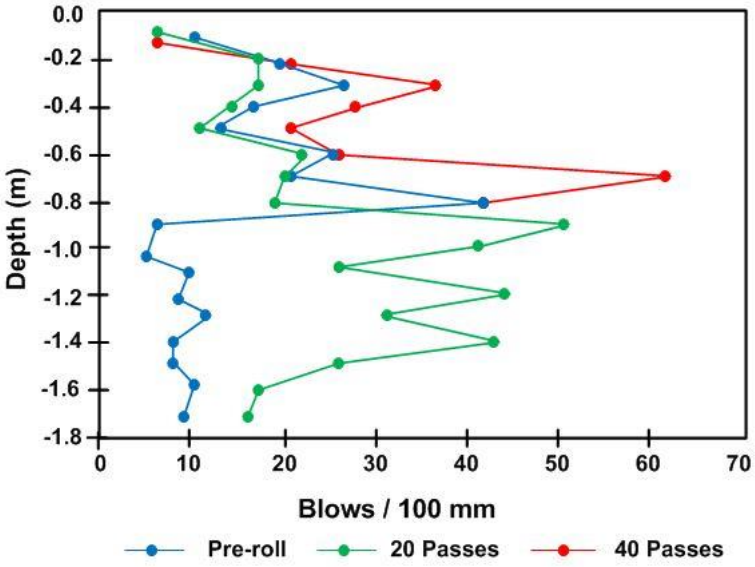
In addition, each of these penetrometer tests (i.e. CPT, DCP, PPT and SPT) are vulnerable to disturbance and damage by the hard and large particles in the ground and, thus, the verification of ground improvement by RDC using these techniques is limited within such heterogeneous materials. For example, Avalle and Grounds (2004) utilised the CPT and DCP in a trial at the Adelaide Airport project, where the heterogeneous fill, that included rock and concrete fragments, resulted in relief drilling and a loss of the continuous CPT data [Figure 2.18(a)]. In addition, the required DCP test depth was compromised by refusal on hard particles [Figure 2.18(b)]. Avalle and McKenzie (2005), in another study, also experienced penetration refusal of the SPT in a field trial conducted in a site consisting of granular and heterogeneous refuse fill.

However, all the above mentioned penetrometer tests provide indirect measurements of in situ density at a specific location. However, it is desirable to correlate these with some additional parameters to obtain the actual in situ density values. Several in situ density tests that provide direct measurements are sometimes used in practice to ascertain the degree of compaction on site, such as the sand replacement method and nuclear density test. Sand replacement tests, performed in accordance with the AS 1289.5.3.1 (Standards Association of Australia, 2004a), were undertaken in a field trial

by Jaksa et al. (2012) to investigate the degree of compaction with respect to different numbers of impact roller passes. However, they experienced difficulties in conducting a sufficient number of tests within the available time, as this test is particularly time consuming and can only be carried out at the ground surface or within a trench.



(a)



(b)

Figure 2.18 Effects of large soil particles on penetrometer testings: (a) relief drilling of CPT; and (b) limited test depth of DCP [obtained from Avalle and Grounds (2004)].

Over the last two decades, the nuclear density gauge has become popular and accepted among the available field density tests. The nuclear density test is performed in accordance with the AS 1289.5.8.1 (Standards Association of Australia, 2007) and involves measuring the in situ density at discrete locations within the upper 300 mm layer of soil using a probe fitted with a radioactive substance. Thus, it is considered well suited for the verification of thin layer compaction, which is usually undertaken by conventional compaction methods (Scott et al., 2012). Nonetheless, when nuclear density tests are adopted for the verification of in-depth compaction of RDC, it is often necessary to excavate the compacted material down to the targeted bench level (Scott et al., 2012). Scott and Suto (2007) reported that the nuclear density method is time consuming and also difficult to operate in a mixed soil, especially in the presence of large particles.

Non-destructive test methods

The non-destructive tests that are utilised to verify the ground improvement by RDC involve geophysical techniques. There are 3 main types of geophysical methods that have been adopted in RDC projects and they all involve surface waves and are the: (i) spectral analysis of surface wave (SASW) method; (ii) multichannel analysis of surface wave (MASW) method; and continuous surface wave system (CSWS). Geophysical test methods incorporate a series of geophones that capture the energy waves passing through the ground and enable the measurement of ground accelerations, which are related to the seismic shear wave velocity at a given depth. The resulting dispersion curves can be readily converted to velocity-depth profiles, which are indicative of stiffness-depth profiles and thus, facilitates an on-site immediate assessment of the degree of ground improvement (Bouazza and Avalle, 2006). In many ways, the geophysical methods are found to be economical and rapid approaches, since a single survey can cover a wider area compared to the traditional methods that investigate one location at a time (Avalle and McKenzie, 2005). In addition, geophysical tests are found to be of value in sites that contain mixed soils and large oversized materials (Scott and Jaksa, 2012). However, in the same study, the authors described the difficulties in implementing geophysical testing in the presence of noise disturbances (Scott and Jaksa, 2012). Some of the documented

implementation of geophysical methods for the verification of RDC includes: CSWS by Avalle and McKenzie (2005) and Bouazza and Avalle (2006); MASW by Scott and Suto (2007); and SASW by Scott and Jaksa (2012).

In addition to the above methods, there are several other additional techniques that can also be incorporated in RDC field trials depending on the specific application. In this regard, Avalle (2004b) presented a comprehensive list of test methods, which can be implemented for the purpose of monitoring and verification of RDC. However, these methods are not discussed further here, as they are beyond the scope of the present study.

2.4 SUMMARY

In this section it has been shown that the RDC is an effective and efficient technique for ground improvement, with many other advantages over traditional soil compaction methods in current use. Furthermore, RDC has proven to be suitable for many compaction applications. As described in detail, to date, RDC has been studied experimentally through a number of field-based case studies in order to investigate the effectiveness and zone of influence. In addition, a significant amount of data has been gathered through an extensive number of RDC projects conducted worldwide. However, there still exists a lack of knowledge for a *priori* estimation of the effectiveness of RDC in different soil profiles. Indeed, the development of a reliable, theoretical model for prior estimation of the effectiveness of RDC is complex due to the various site-specific factors that affect the ground compaction process. However, in this research, a reliable predictive tool will be developed based on artificial intelligence techniques to investigate and quantify the effectiveness of RDC in different ground conditions. These techniques are treated in the next chapter.

INTENTIONALLY BLANK

Chapter 3.

Literature Review:
Artificial Intelligence Techniques

3.1 INTRODUCTION

Machine learning is a sub discipline of artificial intelligence (AI) inspired by biological learning and involves the approaches that adaptively learn and improve by experience (Mitchell, 1997). As such, these methods do not require the prior understanding of underlying relationships in developing models and instead, they are capable of extracting knowledge of complex patterns and a variety of discriminants from machine readable data (Aminian et al., 2013). Machine learning approaches, thus, may considered to be alternatives for traditional approaches for solving real world problems (Alavi and Sadrossadat, 2016). Over the last decade, an extensive number of applications of machine learning techniques have been used to model geotechnical engineering problems associated with non-linearity and these applications are discussed later in this chapter.

With this in mind, this research investigates the feasibility of using two machine learning techniques, i.e. artificial neural networks (ANNs) and genetic programming (GP), to develop predictive models that forecast the performance of RDC. ANNs can be considered as the most widely used AI technique (Aminian et al., 2013), by which its architecture attempts to simulate the functionality of the biological neuron system, i.e. the brain (Shahin et al., 2008). The other approach that is adopted in the present study, GP, makes use of the principle of Darwinian natural evolution (Koza, 1992). In comparison, GP and its variants can be considered as fairly novel among other modelling approaches of prediction and forecasting problems in the field of geotechnical engineering (Alavi et al., 2013). However, as emphasised in this chapter, both of these AI techniques are different and have each been shown to possess superior capability, when compared with traditional statistical modelling when dealing with geotechnical engineering systems.

This chapter provides an overview of ANNs and GP, which are the focus of the numerical models employed in this study, followed by the details of their structure and operation. The classification of different ANN types and GP variants are described and the process of model development is also presented. Finally, the applications of both

these AI techniques to geotechnical engineering problems are discussed and assessed for their applicability in this arena.

3.2 ARTIFICIAL NEURAL NETWORKS (ANNS)

ANNs is one of the most popular AI techniques, which has gained wide acceptance in predictions and forecasting applications in many disciplines including engineering, medicine and finance. The ANN concept has become apparent from natural neural networks (NNNs), where it makes use of the knowledge of the functionality of a biological neuron system that includes the human brain and nervous system and mimics their assessment capabilities. ANNs learn from previous examples and obtain knowledge from data and, as such, the functional relationships among variables are captured when they are presented with a set of inputs and their corresponding outputs.

ANNs belong to the statistically-based set of approaches of forecasting models that determine relationships from historical data sets, whereas its counterparts, i.e. physically-based approaches, attempt to model directly the underlying physical relationships (Maier and Dandy, 1996). In addition, ANNs have a relatively less demanding requirement for the amount of data in comparison to the physically-based systems, which often require an adequate amount of data to calibrate the models (Maier and Dandy, 1996). ANNs are of particular value in geotechnical engineering applications as they do not require the statistical distribution of the input dataset to be known (Burke, 1991), which is the case for most geotechnical data as often they are not normally distributed and the transformation to normality is difficult. Moreover, ANNs have been found to perform well even with noisy or incomplete data (Maren et al., 1990; Burke, 1991; Tang et al., 1991).

When used for forecasting purposes, ANNs employ a similar philosophy to that of traditional statistically-based approaches in that both the methods use a historical data set to estimate the non-linearity between variables (Shahin et al., 2008). In other words, both these methods capture the relationships between variables by continuously adjusting either the connection weights, in the case of ANNs, or model parameters, in

the case of traditional statistical approaches (Maier and Dandy, 1996; Shahin et al., 2000). However, ANNs are one of the data-driven techniques that do not require *a priori* knowledge of the relationships between variables (Lachtermacher and Fuller, 1994). In spite of that, they utilise data to approximate both the optimal model structure as well as the unknown model parameters (Shahin, 2010). Thus, ANNs are well suited for modelling complex problems, where non-linear patterns among the variables exist with undetermined relationships (Maier and Dandy, 1996). In contrast, most of the conventional statistical methods are model driven and often require the structure of the model, i.e. order of the model, to be established prior to parameter estimation (Shahin et al., 2000). Another advantage that ANNs possess over traditional statistical modelling, is their flexibility of implementation. A simple modification to the transfer function or the number of nodes in a hidden layer can vary greatly the model complexity in ANNs (Maier and Dandy, 2000; Shahin et al., 2000). ANNs have become increasingly popular over the years despite the lack of governing rules in the modelling approach in comparison to the standard statistical modelling (Maier and Dandy, 2000). Additionally, a number of research studies have shown that ANNs overcome these limitations and outperform the traditional methods, as will be discussed later in this chapter. In the following subsections, the details of ANN structure are provided along with a treatment of their functionality and operation.

3.2.1 Structure and Operation

The structure and operation of ANNs are briefly described herein, which have already been widely discussed in the literature [e.g. Erzin et al. (2009)]. The primary structural component of an ANN is a simple processing element, which is also termed as node or a neuron. These nodes are usually arranged in layers and structured into a highly interconnected network. Essentially, the nodes in each layer are fully or partially interconnected with the nodes in other layers, whereas each connection is comprised of a particular adaptation coefficient, generally known as a weight (Maier, 1995).

A typical node in the ANN structure, receives inputs from a single or several other nodes via the weighted connections. The summation of the weighted inputs at a

particular node is then combined with a threshold level. This combination then passes through a non-linear transfer function to produce the output of that particular node. This procedure is continued from layer to layer, where the output of each node becomes the input to another node/s. The threshold and transfer function are explained in detail in Chapter 8. This process for node *j* is shown in Figure 3.1 and summarised in the following equations.

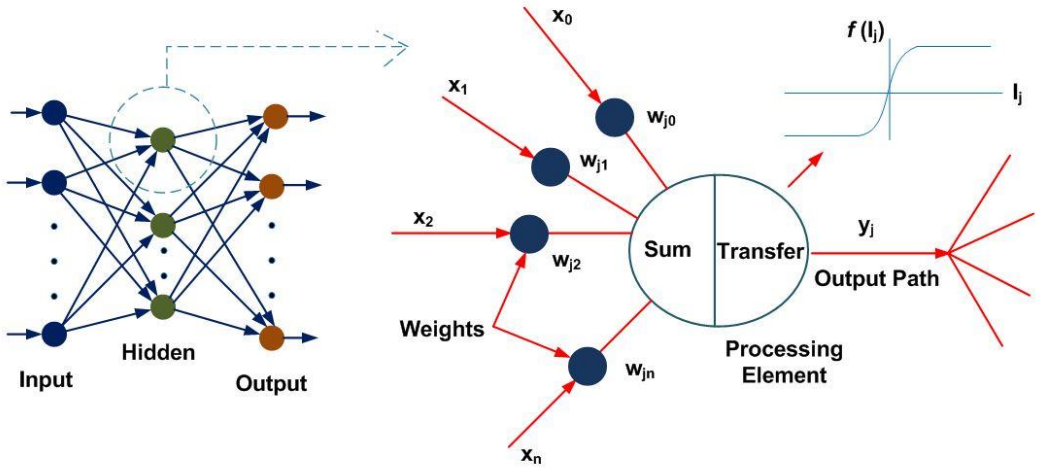


Figure 3.1 Operation of a neural network node [modified from Maier (1995)].

Summation:

$$I_j = \sum_{i=0}^{n} w_{ji}\, x_i + \theta_j \qquad (3.1)$$

Transfer:

$$y_j = f(I_j) \qquad (3.2)$$

Where:

	I_j	is the activation level of node *j*;
	w_{ji}	is the connection weight between nodes *i* and *j*;
	x_i	is the input from node *i*, *i*=0, 1,...*n*;
	θ_j	is the threshold for node *j*;
	y_j	is the output of node *j*; and
	$f(.)$	is the transfer function.

Data propagation in a neural network begins with the presentation of an input stimulus at the input layer. The data are then operated by nodes, as described above, and conveyed through the connections until an output stimulus is produced at the output layer. The network predicted output and the actual (desired, measured) output are then compared and a certain error measure is calculated. Given that, the connection weights of the network are adjusted in accordance with the specified learning algorithm of the network. With the presentation of input stimuli to the network, the connection weights are repeatedly updated (Maier, 1995) and this iterative process is called 'training' or 'learning'. However, the aim of the training phase is to find the optimal set of connection weights that enables the network to produce the output similar, or near similar, to the desired output giving the smallest possible error. Hence, for ANNs to capture the input-output mapping appropriately, it is essential to provide an adequate amount of training samples that represent a wide range of situations relevant to the particular problem domain (Najjar and Basheer, 1996).

3.2.2 Classifications of ANN Model Architecture

ANN model architecture can be classified in to two main categories, based on the direction of the information flow through the network: feed-forward networks and recurrent networks (Shahin et al., 2008). The feed-forward networks permit only forward connections, where information propagates from the input layer to the output layer. Conversely, nodes in a recurrent network may have connections with the previous layer, the succeeding layer, the same layer or even to the nodes themselves (Warner and Misra, 1996). Thus, recurrent networks allow information to propagate through the network both in the forward and backward directions.

However, feed-forward networks are the most common model architecture used in forecasting applications (Maier et al., 2010). Despite the potential benefits of recurrent networks, there are several reasons for the popularity of feed-forward networks. As highlighted by Maier and Dandy (2000), some of them are: (i) feed-forward networks are found to outperform recurrent networks [e.g. Khotanzad et al. (1997)]; (ii) the processing speed is high compared to other currently available models (Masters,

1993); and (iii) there is no clear practical advantage for recurrent networks over the feed-forward networks (Hochreiter and Schmidhuber, 1997). Based on these facts, and the success of feed-forward networks in previous similar applications in geotechnical engineering, this research also adopts feed-forward networks. As illustrated in Figure 3.2, currently, there can be found several different forms of feed-forward model architectures.

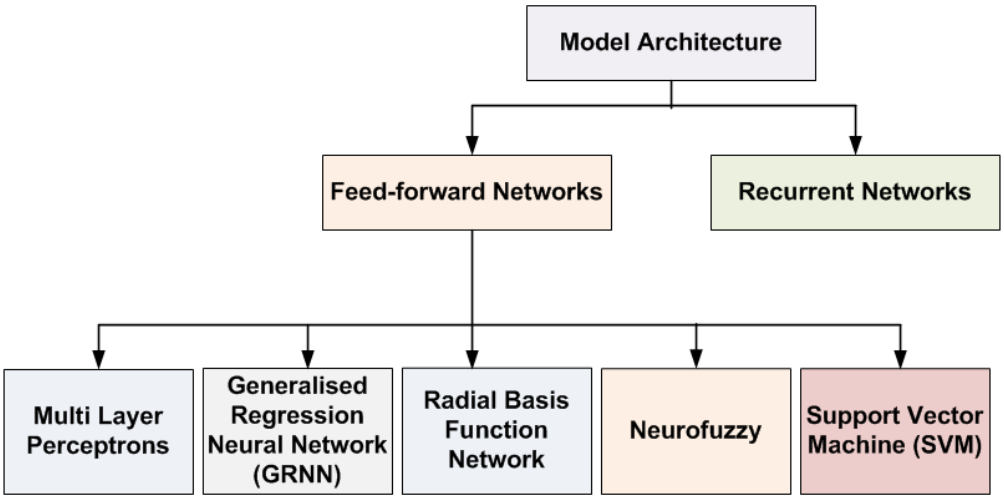


Figure 3.2 Different ANN model architectures.

3.2.3 Multi-Layer Perceptrons

Among other different forms of feed-forward network architectures, the fully inter-connected multi-layer perceptrons (MLPs) are the most common form used in prediction and forecasting applications (Maier and Dandy, 2000; Maier et al., 2010). A number of publications are available in the literature, which have discussed the topology and algorithm details of MLPs (Hecht-Nielsen, 1988; Fausett, 1994; Ripley, 1994). A typical multi-layered structure consists of an input layer, an output layer and a single or several hidden layer(s) between them. MLPs are capable of capturing the complexity and non-linearity of the system being modelled as they use non-linear activation functions at the hidden and output layers (Maier et al., 2010).

It has been demonstrated that single, hidden layer networks with adequate connection weights are capable of approximating any continuous function (Cybenko, 1989; Hornik et al., 1989). As such, there several geotechnical engineering applications that have yielded reliable predictions using the simplest form of MLP that consists of 3 layers, including one hidden layer comprised of an adequate number of hidden nodes (Goh, 1994; Banimahd et al., 2005; Sinha and Wang, 2008; Günaydın, 2009; Kuo et al., 2009; Isik and Ozden, 2013). However, several applications of MLPs in geotechnical engineering have found difficulties in approximating the exact functional relationship between the variables and a lack of robustness in the presence of a single hidden layer, and in these situations multi-hidden layer networks need to be adopted [e.g. Pooya Nejad et al. (2009)]. Nevertheless, it is emphasised that the selection of an optimal network structure is highly problem-dependent.

3.2.4 Error Back-propagation Algorithm

Once the model architecture is decided, the network can be optimised using a suitable optimisation algorithm. The model optimisation involves finding a global solution for a highly non-linear optimisation problem (White, 1989). The error back-propagation method (Rumelhart et al., 1986) is by far the most widely used algorithm for optimising feed-forward ANNs and has been successfully implemented in many geotechnical engineering applications (Goh, 1994; Najjar et al., 1996; Sinha and Wang, 2008; Günaydın, 2009; Kuo et al., 2009; Pooya Nejad et al., 2009). With the adequate provision of internal parameters, the back-propagation algorithm, which is based on the first order gradient descend rule, has the capability of escaping local minima (Maier and Dandy, 1998). As such, the selection of internal parameters has been shown to have a significant influence on the performance of the back-propagation algorithm (Dai and MacBeth, 1997). Therefore, in order to optimise performance, the majority of applications that employ MLPs trained with back-propagation algorithm adopt a trial-and-error approach, whereby the different combinations of internal parameters, including momentum term, learning rate, error and transfer functions, are considered (Maier and Dandy, 2000). However, the slow convergence rate can be considered as a drawback of the back-propagation algorithm. Nevertheless, according

to Breiman (1994), if the optimisation speed is not an issue, there is no other reason to avoid the back-propagation algorithm for optimising feed-forward ANNs.

As described earlier, during the model learning phase, a certain error between the actual and model predicted outputs is being reduced. Accordingly, the set of connection weights of the network is adjusted into an optimal setting that enables the network, with a given functional form, to effectively map the preferred input-output relationship (Maier et al., 2010). The mean squared error (MSE) is the global error function that is frequently used (Warner and Misra, 1996), which has certain benefits including: ease of calculation, penalises larger errors, as well as the simplicity of subsequent derivatives (Masters, 1993). However, when the back-propagation error algorithm is adopted for model optimisation, the error between the actual and model predicted output is propagated backward so that connection weights are optimised layer-by-layer in the backward direction during the training phase. As outlined by Maier and Dandy (1998), the basic steps of the training process of MLPs with the back-propagation algorithm are as follows:

- Initially, small, arbitrary values are assigned to the connection weights;
- A training sample is presented to the network and accordingly, the network output is produced at the output layer;
- The global error is calculated with respect to the actual output and network predicted output. For example, the error function, for node *j*, can be estimated by:

$$E = \frac{1}{2}\sum(d_j - y_j)^2 \qquad (3.3)$$

Where: E is the global error function;
d_j is the desired (actual or measured) output; and
y_j is the predicted output of the network.

- The connection weights (w_{ji}) are adjusted as per the gradient descent rule;

$$\Delta w_{ji}(t+1) = -\eta \frac{\partial E}{\partial w_{ji}} + \mu \Delta w_{ji}(t) \quad (3.4)$$

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t+1) \quad (3.5)$$

Where: $w_{ji}(t)$ is the value of the connection weight at t^{th} iteration;
η is the learning rate; and
μ is the momentum term.

- The next input/output combination is presented and above steps are repeated.

This process is continued until a certain stopping criterion is met. For an instance, training can be ceased either after a certain fixed number of training samples are presented to the network or when the global error diminishes to an acceptable value.

3.2.5 Process Involved in ANN Model Development

In order to achieve better model performance, ANN model development needs to be carried out in a systematic manner. Therefore, special consideration needs to be given to the selection of the model inputs and outputs, and other essential steps including data division, pre-processing, determination of model architecture, model optimisation, stopping criteria and model validation. These major steps in ANN model development as discussed by Maier et al. (2010) are illustrated in the flowchart below (Figure 3.3). These steps are described in detail in later chapters.

3.2.6 Limitations of ANN Modelling

Despite the success of ANN applications in geotechnical engineering, several limitations have been observed. As highlighted by Shahin et al. (2009), some of the key concerns that require further attention, when implementing ANNs for prediction applications are the: necessity of ensuring model robustness, increased model transparency, ability to extract knowledge from the trained ANNs, ability to extrapolate data, and the uncertainty in the model predictions.

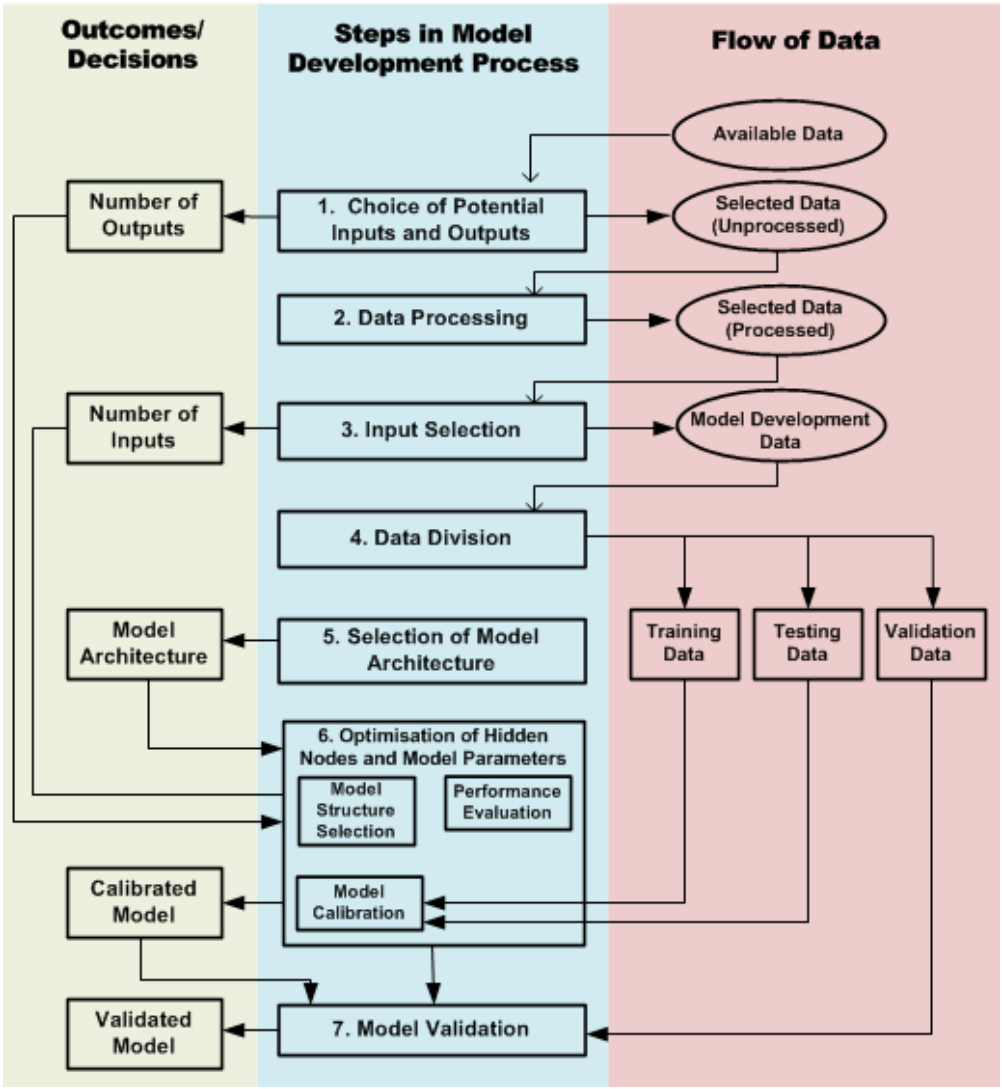


Figure 3.3 ANN model development process as presented by Maier et al. (2010).

Essentially, ANNs require the network structure and the parameters to be predetermined, i.e. number of hidden layers, hidden nodes, learning rate, momentum term, which usually entails the implementation of somewhat ad-hoc, trial-and-error methods. In addition, it is important to ensure the robustness of the model predictions. Model robustness involves the evaluation of the generalisation ability of the ANN model and confirms the validity and the accuracy of model predictions. In an attempt to predict the settlement of shallow foundations using ANNs, Shahin et al. (2005b) observed that good network performance with respect to the training and validation data does not guarantee that the model has the capability to generalise over the range

of the data used in ANN model calibration. For this reason, they proposed a method of sensitivity analysis that tests the robustness of the predictive ability of the ANNs by examining how consistent the model predictions are with the underlying physical behaviour of the system. In addition, they recommended the examination of the optimal set of connection weights by using methods, such as that suggested by Garson (1991), that quantifies the relative importance of the input variables used in the ANN models. Such methods are also adopted in the present study and the details are given in later chapters.

Moreover, a common criticism levelled at ANNs is their lack of transparency and knowledge extraction. This is because the set of connection weights, where the ANN knowledge is stored, is associated with difficulties in interpretation. Nevertheless, some researchers (Kuo et al., 2009; Shahin, 2010) expressed the resulting optimal ANNs in the form of a numerical equation, that can be easily utilised in practice, but in all these cases, the networks are relatively simple, i.e. single hidden layer networks. However, it is difficult to produce straightforward and practical prediction equations, especially when the network includes more than one hidden layer and/or many hidden layer nodes. A limitation of ANNs that inhibits their applicability is the lack of transparency of ANN models, in that they often fail to explain explicitly the underlying physical processes associated with the phenomenon under investigation. In addition to these limitations, the uncertainty associated with ANN model predictions are rarely quantified (Shahin et al., 2009). However, several researchers including Goh et al. (2005) have emphasised that implementing Bayesian techniques overcomes model uncertainty and provides an assessment of the confidence of model predictions.

In addition to ANNs, the present study investigates the applicability of another, relatively new AI technique, genetic programming (GP), that is reported to overcome many of the shortcomings associated with ANNs and other conventional forecasting applications.

3.3 GENETIC PROGRAMMING (GP)

Genetic programming (GP) is one of the subsets of evolutionary algorithms and was first introduced by Koza (1992). This supervised machine learning technique is based on the principles of Darwinian theory and mimics the aspects of genetics and natural selection. In recent years, GP has emerged to be a promising approach for non-linear modelling of different aspects in civil engineering.

Similar to ANNs, GP can be considered as another alternative approach to conventional methods, because of its ability to approximate any linear/non-linear relationship among a set of observed input and output data in the absence of prior knowledge of the underlying mechanisms of the system. However, ANNs require the network structure to be known in advance, but conversely, GP is capable of self-parameterisation and builds the structure itself without any user involvement (Mehr et al., 2014). In general, GP is considered as an extension to genetic algorithms (GAs). As a result, GP utilises the majority of the genetic operators used in GAs, with slight modifications (Alavi et al., 2013). Moreover, both GA and GP approaches are similar in that they both involve a multi-directional, simultaneous search for an optimal solution, from a pool of many potential solutions, which is collectively known as a *population*. The fact that these methods operate from a population enables them to escape local minima in the error surface and are thus able to find optimal/near optimal solutions (Selle and Muttil, 2011). However, GAs and GP have differences in their solution representation. GAs are often recognised by individuals represented as fixed-length binary strings (Holland, 1975), which require post-processing prior to execution. Conversely, in GP, the individuals are the computer programs whose size, shape and complexity is dynamically varied during evolution (Koza, 1992) and are usually executable without post-processing. In addition, GAs and other conventional statistical methods are merely involved in determining the optimal values of a pre-specified form of a function. Conversely, GP evolves the program structure of the approximation model along with the values of its parameter setting (Torres et al., 2009; Mousavi et al., 2011; Alavi et al., 2013).

3.3.1 Classifications of Genetic Programming

As mentioned above, GP involves a population of randomly created individuals, each of which represent a possible solution to the system being modelled. As far as GP is concerned, the individuals or the solutions are represented by computer programs. These individuals are comprised of two basic elements, i.e. functions and terminals. The function set can be composed of arithmetic functions (+, –, ×, /), mathematical functions (*sin*, *cos*, *ln*), logical expressions (IF or THEN), Boolean logic operators (AND, OR, NOT), iterative functions (DO, CONTINUE, UNTIL) and/or other user-defined functions (Sette and Boullart, 2001). The terminal set typically comprises input variables attached to the problem domain and pre-specified or randomly generated numeric constants.

In the traditional GP approach, which is also known as *tree-based genetic programming* (TGP), the computer programs (individuals) have a symbolic representation of a rooted tree-like structure. The tree architecture is composed of several links and nodes, where the internal nodes are called *functional nodes*, as they accommodate the functions, whilst external nodes are the *terminals* that hold the input variables or constants (Koza, 1992). In more advanced levels, several subtrees are hierarchically grouped under a root node to form the architecture of the TGP individual.

Although all the GP approaches follow a similar theoretical foundation with respect to the automatic evolution of computer programs, there are differences in the representation of individuals. As such, 3 distinguishing forms of representation in GP individuals (computer programs) can be found. Besides the traditional tree-based GP approach, these can be either a linear or graphical representation (Banzhaf et al., 1998; Poli et al., 2007). In linear-based GP variants, there is a clear difference between the genotype and phenotype of an individual (Alavi and Gandomi, 2012). Moreover, the individuals have a linear string representation, which is decoded and expressed like nonlinear entities (Oltean and Grosan, 2003). In the recent past, several linear-based variants of GP have been utilised in civil engineering applications, i.e. linear genetic programming (LGP), multi-expression programming (MEP), Cartesian genetic

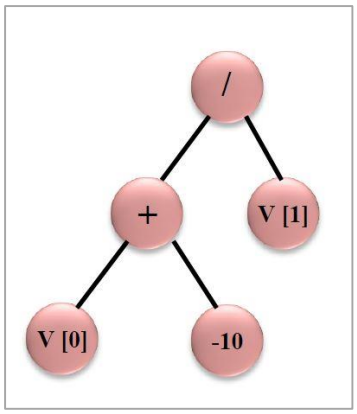
programming (CGP), gene-expression programming (GEP) and grammatical evolution (GE) (Oltean and Grosan, 2003). In the present study, the LGP approach is adopted. The major benefit of LGP utilisation is the ability of implementing native binary machine code that may be immediately executable by the processor, which in turn speeds up the process significantly. Moreover, several numerical studies have shown that LGP outperforms several other linear-based GP variants [e.g. Oltean and Grosan (2003)] and is also capable of outperforming the traditional tree-based approach [e.g. Brameier and Banzhaf (2007)]. These are described in detail in the next section.

3.3.2 Linear Genetic Programming

In LGP, the evolved programs are represented as a sequence of instructions, either from an imperative language (e.g. C, C++ or Java) (Brameier and Banzhaf, 2001; Brameier and Banzhaf, 2007) or from a machine language (Nordin, 1994). In contrast to the rigidly determined, tree-structured, data flow of TGP, LGP has a more general, specially directed graphical structure in the functional level resulting from multiple usages of register contents (Brameier and Banzhaf, 2007; Alavi et al., 2013; Gandomi et al., 2014), as can be seen in Figure 3.4. Moreover, the existence of ineffective code segments, which are also referred to as *introns* in LGP makes them different from their traditional tree-based counterparts. As such, these structurally ineffective codes denote the instructions, which manipulate the registers that have no influence on the output calculation (Gandomi et al., 2010a). Although these ineffective code segments coexist with the effective code, they are not connected to the data flow, unlike in TGP, where the structural introns do not exist because all the program components have a connection with the root node (Brameier and Banzhaf, 2007). However, the structural introns in LGP can be detected efficiently and completely due to its imperative program structure (Francone and Deschaine, 2004; Alavi et al., 2013).

There is a special variant of LGP, known as *automatic induction of machine code by genetic programming* (AIMGP), where the individuals are represented and manipulated as native binary machine code (Nordin, 1994; Banzhaf et al., 1998).

During fitness evaluation in GP, the programs are executed multiple times or at least once, which is considered to be the most time-critical step in evolutionary algorithms (Brameier and Banzhaf, 2007). In this context, program execution implies the interpretation of internal program representation. However, in AIMGP, the individuals are directly executable by the processor and this avoids the use of an expensive interpreter (Francone and Deschaine, 2004; Brameier and Banzhaf, 2007). As a result, AIMGP is found to be significantly faster and more memory efficient when compared with other interpreting GP variants (Nordin, 1994; Brameier and Banzhaf, 2001). Given these advantages, AIMGP is also utilised in this study.




```
float DiscipulusCFunction (float v[])
{
 f[0]=0;
 double Input000=v [0];
 double Input001=v [1];

 L0:   f[0]+= v[0];
 L1:   f[0]+=-10;
 L2:   f[0]/=v[1];

 return f[0];
}
```

(a) (b)

Figure 3.4 Comparison of the GP structures: (a) TGP; and (b) LGP [modified from Alavi and Gandomi (2012)].

3.3.3 Steps in Genetic Programming

As described earlier, each individual in a population that is made up of functions and terminals represents a possible solution for the problem being considered. However, these individuals in a population compete with each other, such that the fittest individuals survive and are subsequently modified and subjected to several genetic operations to form a new population. This process is continued, such that the individuals are eventually evolved through a series of generations to perform well in the given environment.

However, in this process, evolutionary algorithms like GP, require less user input because they are capable of solving problems automatically, starting with a high-level statement of what needs to be done (Poli et al., 2008). Therefore, as described by Koza and Poli (2005), it is necessary to implement certain, well defined preparatory steps, through which the user communicates with the high-level statement of the problem of the GP algorithm. The major preparatory steps (Koza and Poli, 2005) are:

1. Specification of the set of terminals;
2. Specification of the set of primitive functions;
3. Definition of the fitness measure that estimates the fitness of individuals in a population;
4. Allocation of certain parameters for controlling the GP run (e.g. population size, maximum size of program, probabilities of genetic operations); and
5. Specification of the termination criterion and the method for designating the results of the run.

After implementing the above preparatory steps, a GP run is executed through a well-defined systematic approach. Therefore, as described by Brameier and Banzhaf (2007), the basic execution steps of the LGP evolutionary algorithm are as follows:

1. Initialising a population of randomly generated programs and evaluating their fitness;
2. Performing two fitness tournaments with randomly selected programs from the population and selecting the winning programs;
3. Making temporary copies of the two winning programs;
4. Transforming the two winning programs into offsprings subjected to genetic operations, i.e. crossover and mutation with certain probabilities;
5. Replacing the two tournament losing programs with the temporary copies of the winning programs; and
6. Repeating Steps 2 to 5 until the termination or convergence criteria are satisfied.

The above specified, preparatory steps and execution steps of the LGP algorithm are described in greater detail later in Chpaters 6, 7 and 8.

3.4 APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS AND GENETIC PROGRAMMING IN COMPACTION OF THE GROUND

AI techniques, especially ANNs and GP, have shown much better success in modelling complex problems than traditional approaches due to its superior predictive ability. As a result, these techniques have shown increased use in many areas of geotechnical engineering for predicting soil properties and behaviour. In this section, several geotechnical engineering applications are examined and discussed briefly. However, since the scope of the present study is to apply ANNs and GP techniques to predict the ground improvement due to rolling dynamic compaction (RDC), the applications discussed herein are limited to aspects related to soil compaction. Therefore, in the following section, the applications of ANNs and GP for the prediction of compaction characteristics and the permeability of compacted soils are described in detail, whilst several other applications, such as predictions of soil properties, settlements and structural stability are examined briefly. It is, however, beyond the scope of this work to review every AI application available in the literature in detail.

3.4.1 Predictions of Compaction Characteristics of Soil

The maximum dry density (MDD) and optimum moisture content (OMC) are the key parameters used to represent compaction characteristics of soil. These parameters have attracted much research activity in relation to the application of AI techniques. Najjar et al. (1996) developed ANNs trained using the back-propagation algorithm to predict both MDD and OMC. Two separate models were developed incorporating two datasets from previous research, i.e. synthetic soil (Wang and Huang, 1984) and natural soil (Jeng and Strohm, 1976). The parameters that are considered to be most influential to compaction, such as specific gravity, consistency parameters and soil composition

percentages, were selected as the inputs for the networks that enable multiple outputs, i.e. MDD and OMC. The authors developed several ANNs with a number of different combinations of input variables and numbers of hidden nodes. When the models were examined with the associated mean absolute relative error (*MARE*), lower values were observed with an increasing number of hidden nodes and also with many input variables. However, having considered the practical application, the models with a smaller number of nodes were selected and recommended, although the lowest *MARE* was observed with complex networks, i.e. those with a greater number of hidden nodes. Their work further compared the resulting ANN models with the available regression equations, and found that ANNs produce better predictions than the regression equations. In addition, they highlighted the benefits of utilising ANNs over correlation equations. In the ANN modelling, they used a separate set of testing data that was never utilised during the model calibration phase, whilst in previous work (Wang and Huang, 1984), all data sets were used to obtain the correlation equations.

Another research study was carried out by Günaydın (2009) to predict the compaction parameters of soil. The author employed feed-forward MLPs, trained using the back propagation algorithm. The number of input variables, number of hidden nodes and transfer function were varied in the networks that adopted multiple outputs, i.e. MDD and OMC. In selecting the optimum network, several statistical parameters, i.e. coefficient of determination (R^2), standard deviation (σ), standard error (*SE*) and mean value were evaluated. In addition, the compaction parameters were also determined using two other statistical approaches, simple regression analysis (SRA) and multivariate linear regression (MLR). All of these 3 approaches demonstrated strong correlations with the estimations. However, with SRA, the regression equations incorporated only a single independent variable at a time, whilst in MLR, it became possible to use several independent variables simultaneously in developing the linear equations for predicting the compaction parameters. Conversely, in ANN modelling, highly non-linear relationships were obtained with the weights and biases relevant to the optimal network, as given by Günaydın (2009).

Naderi et al. (2012) conducted research, which was more or less an extension of the above described study by Günaydın (2009), where the compaction characteristics were estimated based on the soil classification properties. In this study, GP was employed for the non-linear modelling and, in addition, the MLR approach was used to develop linear numerical correlations. In this research, the database reported by Günaydın (2009) was used and data were divided into statistically consistent subsets, such that 80% of the data were used for model calibration, while the remainder was allocated for validation. The evaluation criteria employed in this study included the correlation coefficient (*R*), root mean square error (*RMSE*) and mean absolute percentage error (*MAPE*) and the results are presented in Table 3.1.

Table 3.1 Comparison of compaction model parameters.

Reference	**Method**	***R***		***RMSE***		***MAPE***	
		OMC	**MDD**	**OMC**	**MDD**	**OMC**	**MDD**
Günaydın (2009)	SRA	0.82	0.76	2.29	0.85	0.13	0.04
	ANN	0.89	0.84	–	–	–	–
Naderi et al. (2012)	MLR	0.88	0.87	1.62	0.60	0.08	0.03
	GP	0.94	0.94	1.19	0.41	0.06	0.02

As it was evident from the comparison of the results that the ANN models slightly outperform both the SRA and MLR approaches with respect to the OMC predictions, but conversely, with the MDD predictions, the MLR approach was found to have superior performance over ANN, with respect to correlation value. However, this could not be confirmed further due to the lack of results. Nonetheless, the statistical analysis indicated that GP based formulae are capable of producing the highest correlation and the lowest regression errors and, in the sensitivity analysis, the model predictions have also conformed with the expected physical behaviour. Given that, the authors (Naderi et al., 2012) concluded that the GP-based formulae presented in their study are accurate and reliable for the estimation of compaction characteristics to be employed in practical situations, where time and cost are major concerns.

Sulewska (2010a) explored the possibility of using ANNs to predict the compaction characteristics (MDD and/or OMC) of coarse-grained soil by using the parameters used in soil grain size distribution. In this study, MLPs were developed using the test results obtained from the laboratory tests of 6 soil groups and both the one and two output situations were considered (either MDD or OMC were predicted separately or simultaneously). The networks were optimised for the number of input nodes by considering different combinations of input variables, while several network architectures, with different numbers of hidden layer nodes, were also considered for each scenario. In addition, the input variables that had no influence on the outputs were identified as they received zero connection weights. With regards to the single output networks, the prediction accuracy of the optimal ANN model (R^2 = 0.73 and 0.89, for MDD and OMC respectively) was observed to be slightly higher than that obtained from the statistical analysis (R^2 = 0.64 and 0.85, respectively). As a result, the study emphasised the applicability of ANNs for analysing relationships in geotechnics to solve practical engineering problems. The models were proposed to be used for quick estimations of compaction parameters in lieu of time-consuming laboratory tests.

Another similar neural network model was introduced by Abdel-Rahman (2008) to predict the compaction characteristics for coarse-grained soil. Separate networks were developed to predict OMC and MDD, and found to have good prediction accuracy (R = 0.99 for MDD and 0.89 for OMC). From the results of the ANN models, empirical equations were derived to estimate the compaction parameters.

In another study, Sulewska (2010b) examined the density index, which is used as an index of the mechanical properties of coarse-grained soils. In estimating the density index, this study utilised statistical and ANN methods for the predictions of maximum and minimum dry density on the basis of grain size and distribution parameters of coarse-grained soils. The resulting optimal ANN models yielded a higher prediction accuracy than that obtained from the statistical linear regression analysis, when compared the performance measures in terms of relative error (*RE*) and *R*.

Basheer (2001) used neural networks to model compaction curves for natural fine-grained soils empirically by using soil index properties and compaction energy. The developed models were found to yield highly accurate predictions, as indicated by the higher correlation coefficient value, R^2 = 0.98. Two other research studies by Sinha and Wang (2008) and Al-saffar et al. (2013) also developed ANN models for the prediction of compaction parameters. Their results were also found to be encouraging and the non-linear ANN approach again resulted in better predictions than those from other correlation models.

Recently, Noor and Singh (2012) developed predictive models for the estimation of compaction parameters (MDD and OMC) by means of GP. The models were developed correlating the compaction characteristics with the index and physical properties, i.e. plastic limit, liquid limit, plasticity index and specific gravity of fine-grained soils. The developed GP model results were compared with the actual test results, as well as with the available models in the literature. They found that the predictive model proposed in their study yielded better predications than those obtained from the other available models based on statistical regression. In addition, Noor and Singh (2012) emphasised a benefit of utilising GP for non-linear modelling as it incorporates several input parameters, unlike most of the empirical correlations existing in the literature, that use only a single input variable, i.e. either the liquid limit or plastic limit, for the prediction of the compaction characteristics of soils.

However, most of the above described prediction models used common soil index properties and incorporated homogenous data sets, either fine- or coarse-grained soil, in estimating the compaction parameters and thus, their applicability to practical situations is very limited. To overcome this limitation to some extent, recently, Isik and Ozden (2013) presented ANN models to estimate the compaction parameters for both coarse- and fine-grained soils. In this study, a MLP using the back propagation training algorithm was used. The basic soil index parameters were used as the input parameters of the ANNs, along with the transition fine content ratio (TFR), to predict MDD and OMC in separate models. The resulting model predictions yielded very good accuracy (R = 0.951 for MDD and 0.952 for OMC) when tested against the

independent validation dataset. Furthermore, in this study, a sensitivity analysis as proposed by Shahin (2010), was carried out to evaluate the generalisation ability of the optimum network and the results were found to be in a good agreement with the expected physical behaviour of compacted soil and thus, confirmed the reliability of the model predictions.

3.4.2 Predicting the Permeability in Compacted Soil

The compaction of soils also affects other soil parameters. Therefore, the prediction of engineering properties of compacted soils, particularly the permeability or hydraulic conductivity of compacted soils, have been examined by means of AI techniques by several researchers.

Najjar and Basheer (1996) introduced computational neural network models (CNNs) based on the back-propagation algorithm for the prediction of the permeability of compacted clay liners. Two separate CNNs (CNN1 and CNN2) were developed with differing numbers of input parameters that account for both the soils' physical properties and their compaction characteristics. However, later, Boroumand and Baziar (2005), addressed the issues involved in their earlier models and proposed another ANN for the prediction of the permeability of clay liners. As such, they sought to reduce the number of input variables and, at the same time, increase prediction performance. The resulting prediction performance, in terms of R^2 and *MARE*, of these models are summarised in Table 3.2.

Table 3.2 Comparison of prediction performance of different models.

Reference	Model	No. of nodes (In-Hidden-Out)	Training		Testing	
			R^2	*MARE*	R^2	*MARE*
Najjar and Basheer (1996)	CNN1	11-3-1	0.923	1.31%	0.281	4.90%
	CNN2	5-3-1	–	1.19%	–	4.07%
Boroumand and Baziar (2005)	Fine6	5-4-1	0.791	–	0.544	–

As can be seen, CNN1 shows a significant difference between the correlation factors for the training and testing sets. The testing set measures the generalisation ability of the network, as the data were not incorporated in the model calibration, which is referred to as a *validation set* by most other researchers, and this terminology will be adopted in the present study, as outlined in later chapters. Thus, the optimality of the developed models is usually judged based on the testing set data. As a consequence, the validity of the CNN1 model is highly questionable. However, the inconsistency may be due to the poor calibration of the model, causing over fitting, or sometimes the testing and training data sets may not be statistically consistent to represent the same population. In addition, as can be seen in the table, CNN2 used only 5 input nodes, whilst in CNN1, 11 input variables were used. However, some of the input variables in CNN1 were found to be interdependent, so that they were subsequently omitted in the study by Boroumand and Baziar (2005), which included a smaller number of input variables in their model (Fine6). In their case, the inputs related to compaction properties were replaced by the dry density of the soil after compaction. However, comparing the performance of the Fine6 model with that of CNN1, a slight reduction in R^2 for the training set was observed, whilst it was increased nearly two times for the testing dataset. In addition, the results obtained from CNN2 were also compared with the regression equation, which was derived for the same input variables as adopted by Benson et al. (1994), and they found that CNN2 yielded superior performance and hence the effectiveness of the neural network predictions was confirmed. Another successful application of ANNs for the prediction of the hydraulic conductivity of clay liners is reported by Das and Basudhar (2007).

Erzin et al. (2009) employed ANNs for the prediction of the hydraulic conductivity of compacted fine-grained soils. In this study, feed-forward MLPs were trained using the back propagation algorithm. The ANN model results were compared with experimental results and found to have a good agreement (R^2 was closed to unity), whilst multiple regression analysis yielded poor prediction accuracy (R^2 = 0.30 to 0.55). The research indicated that the developed ANN models were successful in modelling the non-linear relationships and therefore, recommended for future applications.

3.4.3 Other Applications

Besides the fact that ANNs have been successfully applied to many geotechnical engineering problems, the use of other AI techniques, including GP, is relatively limited. For example, the GP approach was applied by Johari et al. (2006) to the prediction of soil-water characteristics curves, which are necessary to describe unsaturated soil behaviour. The proposed GP simulations were found to be very promising, as they outperformed the currently available conventional methods. In addition to the traditional GP approach used by Johari et al. (2006), other sophisticated GP variants have also been applied to several geotechnical engineering problems. For example, Heshmati et al. (2008) used LGP to model soil classification, whilst Alavi et al. (2010) used the MEP approach for the same purpose.

Taskiran (2010) explored the applicability of AI methods in the form of ANNs and gene expression programming (GEP) to predict the California bearing ratio (CBR) of fine-grained soil. Feed-forward type MLPs, trained using the back propagation algorithm, and GEP, which is an algorithm based on the combination of GA and GP, were employed. As with ANNs, the GEP approach was basically used for the development of a mathematical function that best fits a set of presented data. However, in this study, different numbers of inputs and different parameters (number of genes, head size) were used in selecting the optimal GEP model. When the performance statistics of the predictions from the developed ANN and GEP models were compared, it was found that both model predictions were in good agreement with the actual results ($R^2 > 0.8$). This study concluded that both the ANN and GEP approaches were capable of capturing the relationship between the basic soil properties and CBR and thus, were recommended as useful tools for preliminary forecasting of CBR values, which might replace time consuming and costly soil testing.

In Table 3.3, several other applications of GP variants are summarised, as well as the highlights of these research studies.

Table 3.3 Various AI applications in geotechnical engineering.

Application	Reference	Method	Comments
Liquefaction resistance of Sandy soil	Alavi and Gandomi (2012)	TGP, LGP, MEP	• GP models outperformed conventional (MLR) methods. • LGP and MEP models were superior to TGP.
Liquefaction induced lateral displacements	Javadi et al. (2006)	TGP	• TGP outperformed the traditional MLR model. • Design equations were developed based on optimal TGP model and overcame the shortcomings of ANN method.
Estimation of ultimate bearing capacity of shallow foundations resting on rock masses	Alavi and Sadrossadat (2016)	LGP	• LGP models showed satisfactory performance, as indicated by $R > 0.9$ for all data subsets. • Derived LGP model outperformed traditional methods. • The LGP made it possible to incorporate the effects of both qualitative and quantitative parameters.
Settlement of shallow foundations on cohesionless soil	Rezania and Javadi (2007)	TGP	• TGP model had superior performance to the traditional empirical methods. • TGP outperformed ANN-based formula.
Formulation of soil classification	Heshmati et al. (2008)	LGP	• LGP model predictions had higher accuracy ($R > 0.99$) and slightly outperformed the ANN. • Numerical equations were proposed.
Formulation of soil classification	Alavi et al. (2010)	MEP	• MEP model predictions had higher accuracy ($R > 0.99$) and slightly outperformed the ANN. • Quite short and very simple formulae were proposed based on the optimal MEP models.

3.5 SUMMARY

In this chapter, background knowledge of two AI techniques, in the form of ANNs and GP, has been provided, both of which mimic aspects of biological neuron systems and natural genetic evolution, respectively. It is evident from the detailed description of their function and behaviour, that these techniques are capable of extracting knowledge of complex patterns and the non-linearity of the submitted data. This is particularly valuable in studies associated with spatial variability of soil properties. Thus, it has been confirmed that both ANNs and GP facilitate a promising approach for quantifying and estimating the effectiveness of RDC in different ground conditions. It has also been demonstrated that both these AI techniques have the potential to provide accurate predictions and are relevant to a wide range of geotechnical engineering applications, including those related to soil compaction. In the vast majority of these applications of ANNs and GP, the performance statistics demonstrated that these techniques possess obvious superiority over traditional statistical modelling when dealing with complex, non-linear functional relationships. In the remaining chapters of this thesis, ANN and GP models are developed in order to predict the amount of ground improvement resulting from the use of RDC.

INTENTIONALLY BLANK

Chapter 4.

Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Neural Networks and Cone Penetration Test Data

INTENTIONALLY BLANK

Statement of Authorship

Title of Paper	Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Neural Networks and Cone Penetration Test Data.
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Neural Networks and Cone Penetration Test Data. Soils and Foundations, Submitted for Review.

Principal Author

Name of Principal Author (Candidate)	R. A. T. M. Ranasinghe		
Contribution to the Paper	Data collection and pre-processing, Model development based on Artificial Neural Networks (ANNs), Analysis of results and interpretation, and Preparation of the manuscript.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	20/02/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mark B. Jaksa		
Contribution to the Paper	Overall percentage – 10% Initial concept, research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Yien Lik Kuo		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Fereydoon Pooya Nejad		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

ABSTRACT

Rolling Dynamic Compaction (RDC), which is a ground improvement technique involving non-circular modules drawn behind a tractor, has provided the construction industry with an improved ground compaction capability, especially with respect to a greater influence depth and a higher speed of compaction, resulting in increased productivity. However, to date, there is no reliable method to predict the effectiveness of RDC in a range of ground conditions. This paper presents a novel and unique predictive tool developed by means of artificial neural networks (ANNs) that permits *a priori* prediction of density improvement resulting from a range of ground improvement projects that employed 4-sided RDC modules; commercially known as 'impact rollers.' The strong coefficient of correlation (i.e. $R > 0.86$) and the parametric behaviour achieved in this study indicate that the model is successful in providing reliable predictions of the effectiveness of RDC in various ground conditions.

Keywords: *Rolling dynamic compaction, artificial neural networks, cone penetration test ground improvement*

4.1 INTRODUCTION

Compaction is the method of densification of soil by means of mechanically applied energy. A rapid volume reduction takes place during the compaction process due to pore air expulsion, which results in particle rearrangement and sometimes crushing. A number of methods are widely used in ground compaction, such as vibration, impact, kneading and static pressure. These different types of compaction techniques are essentially subdivided into two categories; static and dynamic compaction. Static compaction is the application of a downward force on the ground surface by the self-weight of the equipment, such as circular rollers, which usually employ drums, pad foots and pneumatic multi-tires. Dynamic compaction methods, on the other hand, apply a kinetically-driven downward force, in addition to the equipment's self-weight. Dynamic compaction makes use of heavy tamping, vibratory drums and plates, rammers, vibroflotation and rolling dynamic compaction (Hausmann, 1990).

Rolling dynamic compaction (RDC) has increasingly become popular over the past few decades in the global construction industry and provides an alternative to the traditional approaches of soil compaction (Pinard, 1999). RDC was originally developed by Aubrey Berrangé in South Africa in the late 1940s, but its value was not fully realised until the mid-1980s. It involves towing heavy (6-12 tonnes) non-circular modules (3-, 4- and 5-sided), which rotate about their corners and fall to impact the ground (Avalle, 2004d). When the impact roller module is drawn forward and begins to rotate about its corners due to friction, a series of high amplitude impact blows are delivered onto the ground at a low frequency of 90 to 130 blows per minute (Pinard and Ookeditse, 1988). Thus, the compactive effort is derived from the energy of the mass falling from a corner to the adjacent compacting face of the polygonal shaped mass. This results in deep compaction and a greater influence depth – more than 1 m beneath the ground surface and sometimes as deep as 3 m in some soils (Avalle and Carter, 2005) – compared to the conventional static and vibratory compaction (Clegg and Berrangé, 1971; Clifford, 1976, 1978b), which generally influences depths less than 0.5 m below the ground. As a result, thicker lifts, in excess of 0.5 m can be employed as compared to traditional compaction lifts of approximately 0.15 m (Avalle, 2004d, 2006). Furthermore, RDC is particularly efficient when employed in

large and open sites, as it traverses the ground at speeds of 9–12 km/h compared to traditional vibratory roller, which travels at 4 km/h (Pinard, 1999). As a consequence, RDC has been applied: (1) to the in situ densification of existing fills, such as on brownfield sites, landfills, earth embankments and sub grade proof-rolling (Avalle, 2004a); (2) in the agricultural sector (Avalle, 2004d), mainly for the improvement of existing water storages, channels and embankments; and (3) in the mining industry for the construction of tailing dams (Avalle, 2006), rock demolition in open cut mine waste tips, compaction of the capping over waste rocks (Scott and Jaksa, 2012), compaction of bulk earthworks of mine spoil materials and to induce fracturing on surface layers in rock quarries in lieu of drilling and blasting (Avalle, 2006).

To date a number of field and case studies have assessed the efficacy of RDC. As a result, RDC is often adopted based on experience from previous work undertaken in similar soils and site conditions. In most cases, in order to determine the optimal number of RDC passes to achieve the design specifications, a field trial is undertaken, where measurements of various soil characteristics are obtained before and after compaction.

This study aims to develop a robust predictive model to forecast the performance of RDC by means of the artificial intelligence (AI) technique known as artificial neural networks (ANNs). It is intended that the model will provide additional, *a priori* information to supplement field trials undertaken on site prior to ground improvement. Attention is focused on the 4-sided, 8 tonne 'impact roller' (BH-1300). ANN models have been developed based on the cone penetration test (CPT) data, which are collected from previous ground improvement projects in Australia that employed RDC. Design equations are developed based on model parameters, and a parametric study is carried out to assess the robustness of the model. It is important to note that no such predictive model exists for RDC, neither empirical, theoretical nor numerical.

4.2 DATABASE AND DATA ANALYSIS

The data used in this study have been obtained from the results of several field trials undertaken by Broons SA Hire, an Australian company marketing a range of ground improvement technologies, including RDC. The database is comprised of in situ strength data in the form of cone penetration test (CPT) results with respect to the number of roller passes. CPT measurements provide continuous soil profiles that express variations of soil strength in terms of cone tip resistance (q_c) and sleeve friction (f_s). The CPT is often used as a profiling tool employing the friction ratio, which is the ratio, expressed as a percentage, of sleeve friction to cone tip resistance, measured at the same depth. The literature contains several soil classification charts based on the friction ratio as a function of cone tip resistance (Robertson, 1990).

The CPT (Standards Association of Australia, 1999) has been widely used for monitoring and evaluating the effectiveness of deep compaction methods because of the continuous, reliable and repeatable nature of the measurements (Lunne et al., 1997). Avalle and Carter (2005) investigated the depth of influence of RDC in sandy soils using CPT profiles of prior to and after impact rolling, where a noticeable improvement was evident between depths of approximately 0.5 – 3 m below the ground surface. Moreover, research by Kelly (2000) presented the results from CPTs where a significant depth of influence of RDC was evident to depths of 4 m in natural sand and un-compacted/un-controlled variable clay fill, and in the reclaimed sand deposit to the depths of 5 m below the surface. Another recent case study conducted by Scott and Jaksa (2014) reported on the verification of RDC, both vertically and laterally, in a sand fill having quantified the differences in cone tip resistance between a series of closely-spaced CPT locations.

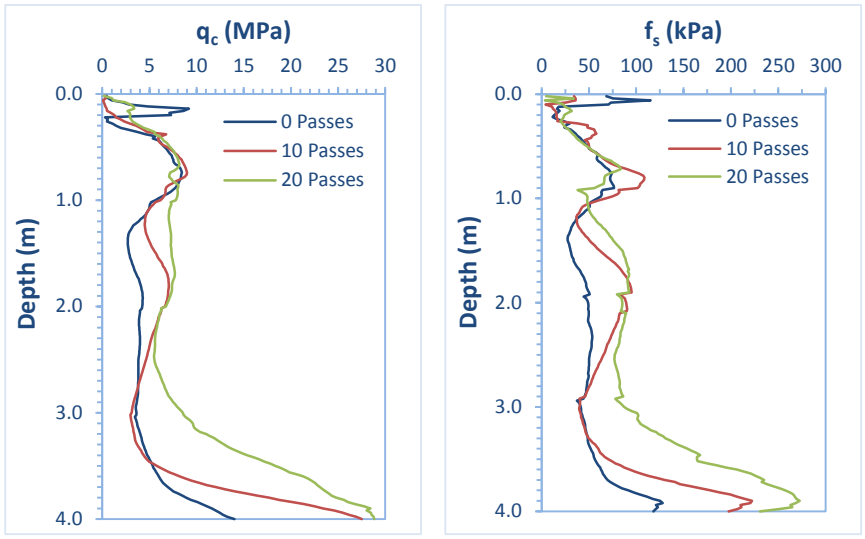
The database used for ANN modelling is summarised in Table 4.1 and the data were acquired from Broons' records of previous projects. Figure 4.1 shows typical plots of cone tip resistance and sleeve friction measurements which are obtained at essentially the same location prior to (0 passes) and after (10 passes and 20 passes) of RDC. As can be seen, there is a noticeable variation in cone tip resistance and sleeve friction measurements with respect to the number of roller passes. The differences between

these individual measurements quantify the variation of soil strength and density, as well as a minor degree of spatial variation, resulting from RDC at the test location.

Table 4.1 Summary of CPT plots.

No	Project name	Max. depth (m)	No. of CPT locations	Soil type	No. of roller passes at the time of test
1	Port Botany	9.8	64	Sandy/silty sandy fill overlying marine deposits of sand and peaty clay deposits	0 10 20 30 40
2	Potts Hill	8.8	19	Shale fill overlying residual soil and sandstone bedrock	0 10 20 30 40
3	Outer Harbor	3.8	8	Uncontrolled fill overlying St. Kilda formation soils	0 24
4	Banksmeadow	6.2	3	Silty sandy filling overlying natural sands	0 5 20
5	Cairns	4.9	9	Uncontrolled fill underlain by silty clay/clayey silt/sandy silt	0 20 30 40

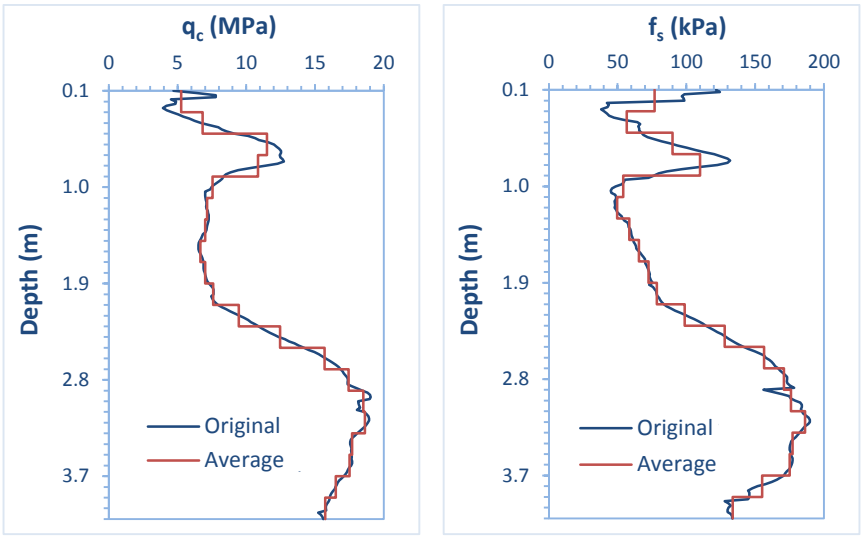
However, the CPT measurements are influenced by large, hard particles present in the compaction material and may not be representative of the subsurface condition, particularly at sites with uncontrolled fill. As a result, some localised peaks are visible in the CPT measurements, predominantly in the cone tip resistance measurements. However, in the data pre-processing, these anomalies are filtered from the dataset, otherwise the derived ANN models will be vulnerable to learning these random and unrepresentative irregularities. In that case, a number of CPT plots from adjacent CPT soundings are superimposed to highlight the localised peaks present due to the spatial variability of the compacted material.



(a) (b)

Figure 4.1 Variation with number of roller passes of CPT parameters: (a) cone tip resistance; and (b) sleeve friction.

In this study, CPT measurements are considered for the depth range of 0.1 m to 4 m. It is well established in RDC, that a significant reduction in the soil strength in the upper 0.1 m is always evident due to the surface disturbance caused by the RDC module (Avalle, 2006). For this reason, the near-surface CPT records (< 0.1 m) have been removed from the dataset. However, below this depth (0.1 m onwards), cone tip resistance is expected to increase with increasing number of roller passes, with a steady decrease down to the depth of influence, which is dependent on the subsurface material. It has been demonstrated by several researchers that RDC influences the ground to depths of 3 m and beyond depending on the soil type and ground conditions (Kelly, 2000; Avalle and Carter, 2005; Jaksa et al., 2012). Hence, CPT measurements are considered to a depth of 4 m for each of the RDC project sites considered. Although the CPT readings from the RDC projects are available at 20 mm depth intervals, the values are averaged over 0.2 m depth intervals for each CPT sounding when incorporated in the ANN modelling in order to obtain a compromise between model parsimony and predictive accuracy. Figure 4.2 shows an example of this arithmetic averaging process with respect to cone tip resistance and sleeve friction.



(a) (b)

Figure 4.2 Example arithmetic average plots of: (a) cone tip resistance; and (b) sleeve friction measurements.

4.3 DEVELOPMENT OF ARTIFICIAL NEURAL NETWORKS (ANNS)

In this study, CPT-based ANNs are developed for the prediction of the effectiveness of RDC. ANN modelling is carried out using the PC-based software, *NEUFRAME* version 4.0 (Neusciences, 2000). The process adopted in the ANN model development is well established in the literature (Maier et al., 2010) and involves the determination of model input(s)/output(s), data division, selection of model architecture, optimisation of network structure, model validation and performance evaluation.

Prior to model development, the database is divided into two subsets:

a. Modelling dataset – This is used to train and validate the ANN models and consists of 1,755 records from 91 CPT soundings in total. This dataset is further divided into three subsets: training, testing and validation. Their statistics and applicability are discussed later in this paper.

b. Verification dataset – Further verification of the developed ANN model is carried out by introducing a new unseen data set, which is not a part of the modelling stage in any capacity. The dataset comprises several CPT soundings randomly chosen from each of the RDC projects included in Table 4.1 and accounts for different numbers of roller passes. It is important to note that this particular dataset functions the same as the validation subset, but differs in that it contains the entire data record of each CPT sounding over the full depth (from 0.1 m to 4 m) and is not mixed with several other CPT soundings. A total of 222 records are used in the verification dataset taken from 12 CPT soundings.

In the present study, appropriate model inputs and outputs are defined based on the prior knowledge of the fundamental factors that influence ground density improvement by means of soil compaction. It is understood that the degree of soil compaction depends upon key factors including: the inherent physical properties of the soil, such as initial density, moisture content, soil type; and the amount of energy imparted to the ground. Therefore, 4 parameters, the depth of measurement (D), cone tip resistance (q_{ci}) and sleeve friction (f_{si}) prior to compaction, and the number of roller passes (P), are selected as potential input variables for the ANN models. In order to predict the level of ground improvement, the models include a single output variable, the cone tip resistance after compaction (q_{cf}). The inclusion of sleeve friction, together with cone tip resistance, provides an indirect and useful representation of soil type in the ANN models through the friction ratio. The number of roller passes effectively expresses the amount of energy imparted to the soil, as the ANN model is unique to a specific RDC module; in this case the 4-sided, 8 tonne impact roller. The statistics of model variables are graphically represented in Figure 4.3 and summarised in Table 4.2.

Since the widely adopted cross validation technique Stone (1974) is used as the stopping criterion in the ANN model development, it requires the dataset to be divided into three subsets: training, testing and validation, as mentioned above. The training set is used to train and calibrate the model and it is with this data subset that the model's connection weights are optimised. While the training progresses, model performance

is assessed periodically with respect to the testing set. When the testing error begins to increase, even though the error obtained using the training set might continue to decrease, training is terminated to avoid overfitting and hence preserve its generalisation ability. After model calibration, the validation data are used to validate the performance of the model using data unseen during model development.

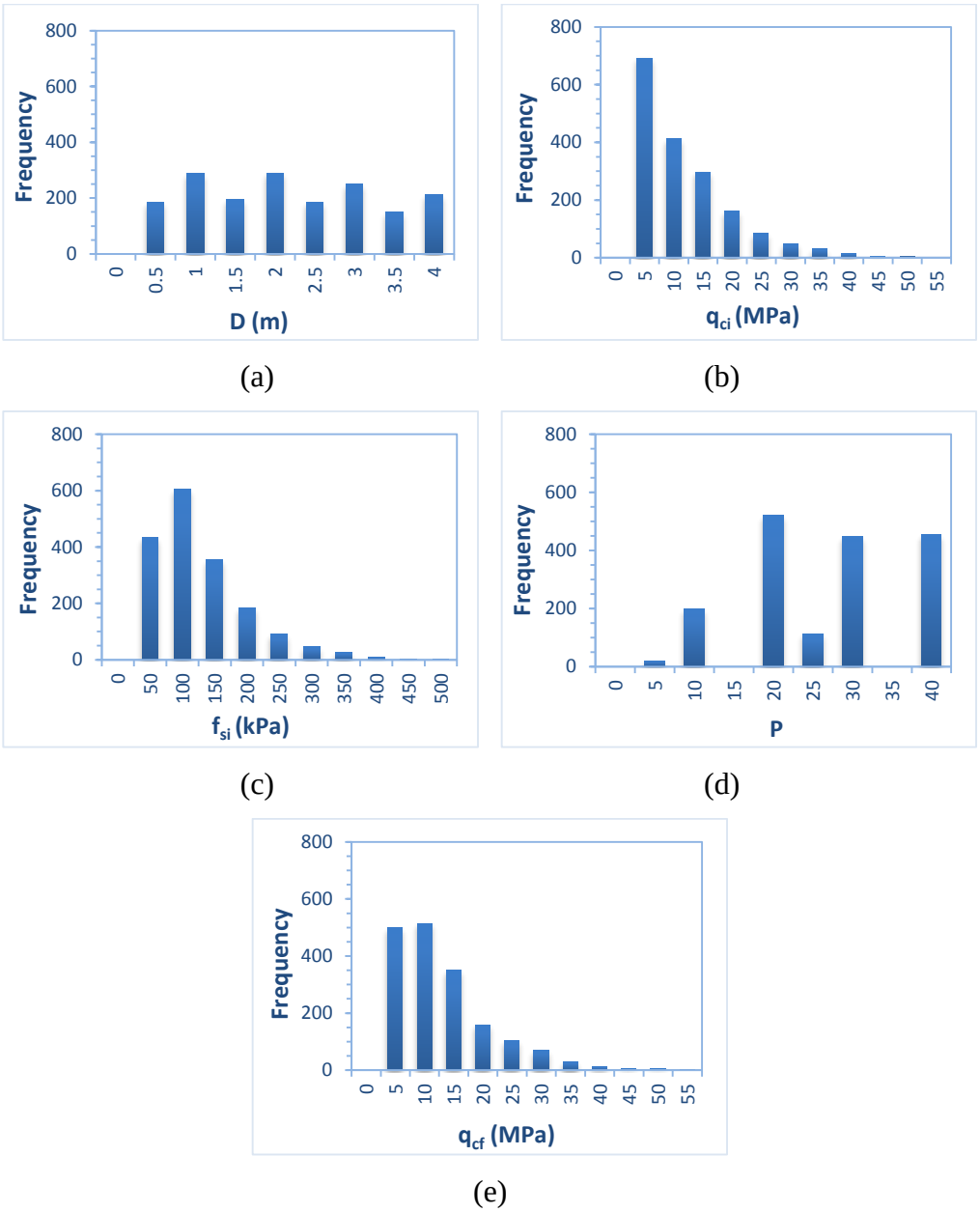


Figure 4.3 Histograms of the data used in the ANN model development: (a) D; (b) q_{ci}; (c) f_{si}; (d) P; and (e) q_{cf}.

Table 4.2 Statistical properties of the data used in the ANN model development.

Model variable	Minimum	Maximum	Mean	Standard deviation
Depth, $\boldsymbol{D}$ (m)	0.20	4.00	1.98	1.12
Cone tip resistance prior to compaction, $\boldsymbol{q_{ci}}$ (MPa)	0.19	50.65	9.34	8.16
Sleeve friction prior to compaction, $\boldsymbol{f_{si}}$ (kPa)	1.67	473.86	102.78	71.37
No. of Roller Passes, $\boldsymbol{P}$	5.00	40.00	26.69	9.97
Cone tip resistance after compaction, $\boldsymbol{q_{cf}}$ (MPa)	0.17	50.36	10.44	8.29

However, as described in the literature (Tokar and Johnson, 1999; Shahin et al., 2004) the method used to divide the data into their subsets may adversely affect ANN model performance and thus in this study data division is carried out using self-organizing maps (SOMs) (Bowden et al., 2002). The SOM method is beneficial since it involves dividing the dataset in such a way that the subsets are statistically consistent and effectively represent the same population. The statistical properties considered in this study include the mean, standard deviation, minimum, maximum and range. ANNs are considered as an interpolation technique and models are expected to perform well when they do not extrapolate beyond the set limits of the training set (Flood and Kartam, 1994; Minns and Hall, 1996). Therefore, it is essential during data division to ensure that the training set contains all the possible patterns included in the dataset so that the final ANN model is as general as possible. The data division is carried out in such a way that the training set contains 80% of the data and the remaining 20% is used for validation purposes. The training set is further divided into two subsets; 80% for training and 20% for testing. However, when using the SOM method there is no absolute rule when selecting the most favourable map size and, for that reason, several map sizes (e.g. 5 × 5, 10 × 10) are examined and the map size which ensures the maximum number of clusters is considered to be optimal. The datasets of each RDC project mentioned in Table 4.1 are individually subjected to SOM data division, rather than introducing the entire dataset into a single SOM. This ensures an even distribution

of variables that represent the site specific characteristics among the three subsets. The selected optimal map sizes for the Port Botany and Potts Hill projects are 20 × 20 and 10 × 10, respectively, while 5 × 5 is optimal for the Outer Harbor, Banksmeadow and Cairns projects. Thereafter, the individually divided datasets are combined to form the three major subsets: training, testing and validation, and their statistics are shown in Table 4.3.

Table 4.3 ANN input-output summary statistics for the training, testing and validation data.

Statistical parameters	**Model variables**				
	***D* (m)**	**q_{ci} (MPa)**	**f_{si} (kPa)**	***P***	**q_{cf} (MPa)**
Mean					
Training	1.95	9.33	103.36	26.59	10.42
Testing	2.03	9.32	99.23	27.21	10.50
Validation	2.03	9.39	103.54	26.62	10.43
Standard Deviation					
Training	1.11	8.23	71.76	9.94	8.30
Testing	1.14	8.37	71.29	9.64	8.66
Validation	1.14	7.83	70.37	10.30	8.03
Minimum					
Training	0.20	0.19	1.67	5.00	0.17
Testing	0.20	0.30	8.70	5.00	0.29
Validation	0.20	0.32	7.08	5.00	0.39
Maximum					
Training	4.00	50.65	473.86	40.00	50.36
Testing	4.00	47.39	441.04	40.00	45.12
Validation	4.00	47.94	470.29	40.00	46.20
Range					
Training	3.80	50.46	472.19	35.00	50.19
Testing	3.80	47.09	432.34	35.00	44.83
Validation	3.80	47.63	463.21	35.00	45.81

Once the available data are divided into the three subsets, data pre-processing is carried out using the min-max normalisation method. In data normalisation, all model variables are scaled into a single range that is commensurate with the limits of the activation function used in the output layer (Minns and Hall, 1996). This can expedite the model training rate and ensures that all the variables receive equal attention during the model training phase. For the ANN modelling undertaken in the present work, the logistic transfer function is used in the output layer and therefore the model variables are scaled between 0.1 and 0.9. Following the model calibration phase, it is necessary to de-normalise the network output by reverse scaling. In this work, multi-layer perceptron (MLP) models are developed with the use of the error back-propagation method. The feed forward type MLP is the most common network architecture used for prediction and forecasting applications (Maier et al., 2010) whereas, the error back-propagation method (Rumelhart et al., 1986) is by far the most widely used algorithm for optimising feed forward ANNs. A comprehensive description of the MLPs trained with the error back-propagation algorithm is beyond the scope of the paper but is well documented in the literature [e.g. Fausett (1994)].

The selected MLP network architecture is comprised of three layers: the input layer, one hidden layer and the output layer. It has been demonstrated that one hidden layer can approximate any continuous function by providing a sufficient number of connection weights (Cybenko, 1989; Hornik et al., 1989). The number of nodes in the input and output layers represent the number of model inputs and outputs and thus the models consist of 4 nodes in the input layer: depth (D), cone tip resistance (q_{ci}) and sleeve friction (f_{si}) prior to compaction and the number of RDC module passes (P). Since the models have a single output variable, i.e. cone tip resistance at depth (D) after the P module passes (q_{cf}), the output layer consists of a single output node. Optimisation of the number of hidden nodes is a crucial aspect of ANN model development and it is essential to achieve a structure that is neither too complex nor too simple but adequately captures the nuances contained in the training data. Therefore, to identify the optimal network architecture/topology a stepwise trial-and-error procedure is used, as is usual practice in ANN model development (Maier et al., 2010). On this regard, several ANN models are trained, starting from the smallest

possible network involving a single hidden node and successively increasing the number of hidden nodes to a maximum of 9. As suggested by Caudill (1988), *2I+1* is the upper limit of hidden nodes for a network to map any continuous function, with *I* being the number of input nodes, and accordingly 9 nodes are considered to be the maximum number of hidden nodes required for the models.

In order to obtain the optimal model, ANN parameters, such as learning rate, momentum and transfer function, are sequentially varied. ANN models are initially trained with the default software parameters (i.e. learning rate = 0.2, momentum term = 0.8 and the sigmoidal transfer function) are used for both the hidden and output layers. After determining the best topology, the network with the optimal number of hidden nodes is subjected to different combinations of learning rates and momentum terms. In addition, as the backpropagation algorithm is based on the steepest descent method, the obtained network results may be sensitive to the initial weight conditions (Maier and Dandy, 2000). Therefore, the selected ANN model is retrained several times after randomising the initial weight allocations to ensure that model training does not cease at sub-optimal levels.

Upon the completion of ANN model calibration, the networks undergo model validation using the validation dataset. As mentioned earlier, this dataset is not used in the training process in any capacity and therefore, it is able to provide a rigorous assessment of the network's predictive and generalisation ability. The criteria used to evaluate the performance of the trained network include the root mean square error (*RMSE*), mean absolute error (*MAE*) and coefficient of correlation (*R*). The prediction accuracy of a well-trained ANN model is represented by the smaller error values and an *R* value close to unity (Smith, 1993).

4.4 RESULTS AND DISCUSSION

This section presents the results of the ANN optimisation process and assessment of the predictive ability and robustness of the optimal model.

4.4.1 Results of ANN Model Optimisation

The performance statistics of the developed models in terms of *RMSE*, *MAE* and *R* value are summarised in Table 4.4, where the model architecture, i.e. the number of hidden nodes, is varied. These performance measures are compared in order to select the optimum model topology which yields the best predictions.

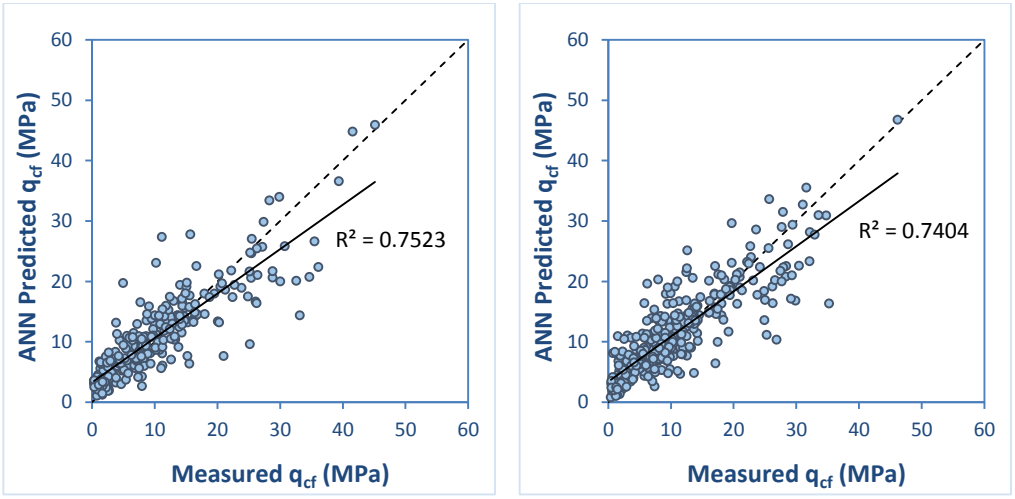
Table 4.4 Performance results of ANN models with different numbers of hidden nodes.

Hidden nodes	Training set			Testing set			Validation set		
	R	***RMSE* (MPa)**	***MAE* (MPa)**	***R***	***RMSE* (MPa)**	***MAE* (MPa)**	***R***	***RMSE* (MPa)**	***MAE* (MPa)**
1	0.844	4.59	3.11	0.860	4.61	3.11	0.850	4.38	3.07
2	0.861	4.27	2.87	0.860	4.48	2.99	0.859	4.14	2.82
3	0.861	4.26	2.87	0.861	4.45	2.98	0.859	4.13	2.82
4	**0.866**	**4.19**	**2.89**	**0.867**	**4.33**	**3.03**	**0.860**	**4.16**	**2.93**
5	0.872	4.13	2.89	0.868	4.34	3.09	0.853	4.29	3.06
6	0.872	4.14	2.90	0.869	4.33	3.07	0.853	4.29	3.07
7	0.865	4.21	2.91	0.866	4.35	3.04	0.860	4.17	2.94
8	0.866	4.17	2.85	0.867	4.32	2.97	0.860	4.13	2.87
9	0.862	4.23	2.84	0.863	4.41	2.95	0.859	4.12	2.79

As can be observed, good consistency is obtained with respect to the *RMSE*, *MAE* and *R* values among the three data sets: training, testing and validation. This implies that the model performance with respect to each of these three subsets is very similar, which in turn, indicates the given subsets represent the same population and that the SOM method for data division is effective. In order to select the optimal ANN model, as mentioned previously, a compromise between predictive accuracy and model parsimony is sought. As a consequence, the model with 4 hidden nodes, as shown in

bold in Table 4.4, is deemed to be optimal due to its comparatively low prediction error (*RMSE* = 4.16 MPa and *MAE* = 2.93 MPa for the validation data subset). However, additionally, the multiple hidden layer (i.e. 2 and 3 hidden layers) models are also investigated but, a significant improvement in model performance over the selected optimal single layer model is not experienced. Therefore, in the interest of selecting a more parsimonious model, a single hidden layer, 4 hidden node network is considered to be optimal, which also facilitates the development of a tractable and useable form of the ANN model, as will be discussed later. In order to refine further the model with 4 hidden layer nodes, the learning rate and the momentum terms are varied, as summarised in Table 4.5. From this analysis, it is observed that the model with a learning rate of 0.2 and momentum of 0.8 performs best.

Figure 4.4 presents a plot of predicted and measured cone tip resistance after compaction (q_{cf}) with respect to the data in the testing and validation sets, in which the solid lines represent the linear trend line fitted to the predicted values and the dotted line indicates perfect prediction. It can be clearly seen that there is minimal scatter, the solid and dotted lines are in relatively close agreement, and hence the model performs well.



(a) (b)

Figure 4.4 Actual versus predicted q_{cf} for the optimum ANN model with respect to: (a) testing set data; and (b) validation set data.

Table 4.5 Effect of varying momentum terms and learning rates on the optimum model.

Learning rate	Momentum term	Training set			Testing set			Validation set		
		R	***RMSE*** **(MPa)**	***MAE*** **(MPa)**	***R***	***RMSE*** **(MPa)**	***MAE*** **(MPa)**	***R***	***RMSE*** **(MPa)**	***MAE*** **(MPa)**
0.2	0.1	0.861	4.22	2.88	0.862	4.39	2.99	0.860	4.10	2.85
0.2	0.2	0.861	4.22	2.90	0.862	4.39	3.01	0.860	4.11	2.87
0.2	0.4	0.862	4.21	2.88	0.862	4.39	3.00	0.860	4.11	2.87
0.2	0.5	0.862	4.26	2.86	0.861	4.47	2.98	0.859	4.13	2.82
0.2	0.6	0.862	4.20	2.85	0.863	4.37	2.98	0.861	4.10	2.84
0.2	0.7	0.862	4.20	2.85	0.861	4.39	2.99	0.858	4.12	2.84
0.2	**0.8**	**0.866**	**4.19**	**2.89**	**0.867**	**4.33**	**3.03**	**0.860**	**4.16**	**2.93**
0.2	0.9	0.868	4.26	3.04	0.868	4.40	3.20	0.860	4.28	3.13
0.05	0.6	0.860	4.23	2.88	0.861	4.40	2.99	0.860	4.10	2.85
0.1	0.6	0.865	4.17	2.84	0.865	4.35	2.98	0.861	4.11	2.85
0.3	0.6	0.871	4.13	2.79	0.866	4.39	2.96	0.858	4.16	2.83
0.4	0.6	0.864	4.18	2.84	0.865	4.35	2.97	0.861	4.10	2.83
0.5	0.6	0.871	4.07	2.76	0.865	4.35	2.98	0.856	4.14	2.87
0.6	0.6	0.873	4.14	2.79	0.864	4.48	2.98	0.855	4.24	2.87
0.7	0.6	0.872	4.50	3.11	0.866	4.80	3.28	0.852	4.59	3.23
0.8	0.6	0.873	4.59	3.53	0.865	4.79	3.74	0.854	4.78	3.78
0.05	0.8	0.865	4.17	2.84	0.865	4.34	2.98	0.861	4.11	2.84
0.1	0.8	0.862	4.21	2.87	0.862	4.39	2.99	0.860	4.10	2.85
0.3	0.8	0.867	4.35	2.92	0.866	4.57	3.04	0.861	4.27	2.92
0.4	0.8	0.870	5.30	4.26	0.864	5.38	4.44	0.849	5.49	4.54
0.5	0.8	0.873	4.29	3.13	0.865	4.52	3.34	0.855	4.45	3.35
0.6	0.8	0.870	5.54	4.47	0.870	5.54	4.59	0.854	5.76	4.85
0.7	0.8	0.874	4.05	2.73	0.864	4.38	2.91	0.856	4.16	2.81
0.8	0.8	0.875	4.28	3.12	0.868	4.49	3.32	0.858	4.43	3.33

4.4.2 Further Verification of Selected Optimum ANN Model

The capabilities of the optimal ANN model are further evaluated by introducing a series of unseen complete CPTs to the model. Details of this dataset were discussed earlier and the performance of the model with respect to these additional CPTs is summarised in Table 4.6 and Figure 4.5. As is evident, the model performs very favourably.

Table 4.6 Model performance on the verification data set.

CPT location	*R*	*RMSE* (MPa)	*MAE* (MPa)
Port Botany - 30	0.955	3.63	3.39
Port Botany - 11	0.840	3.65	2.86
Port Botany - 45	0.971	6.06	5.30
Port Botany - 35	0.723	3.75	2.51
Potts Hill - 37/44	0.422	3.15	2.12
Potts Hill - 27/54	0.436	2.23	1.93
Potts Hill - 24/57	0.530	2.88	2.17
Outer Harbor - EFC 5	0.838	3.15	2.84
Banksmeadow - C3	0.141	2.49	1.90
Cairns-CPT 2	0.613	4.26	2.95
Cairns-CPT 5	0.794	2.83	2.31
Cairns-CPT 8	0.943	2.55	2.27

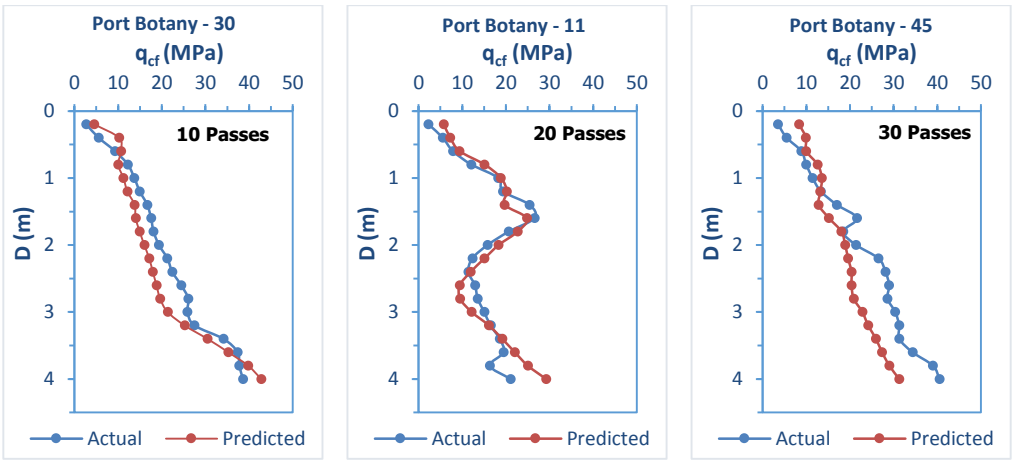


Figure 4.5 Plots of actual and model predicted CPT results (continued overleaf).

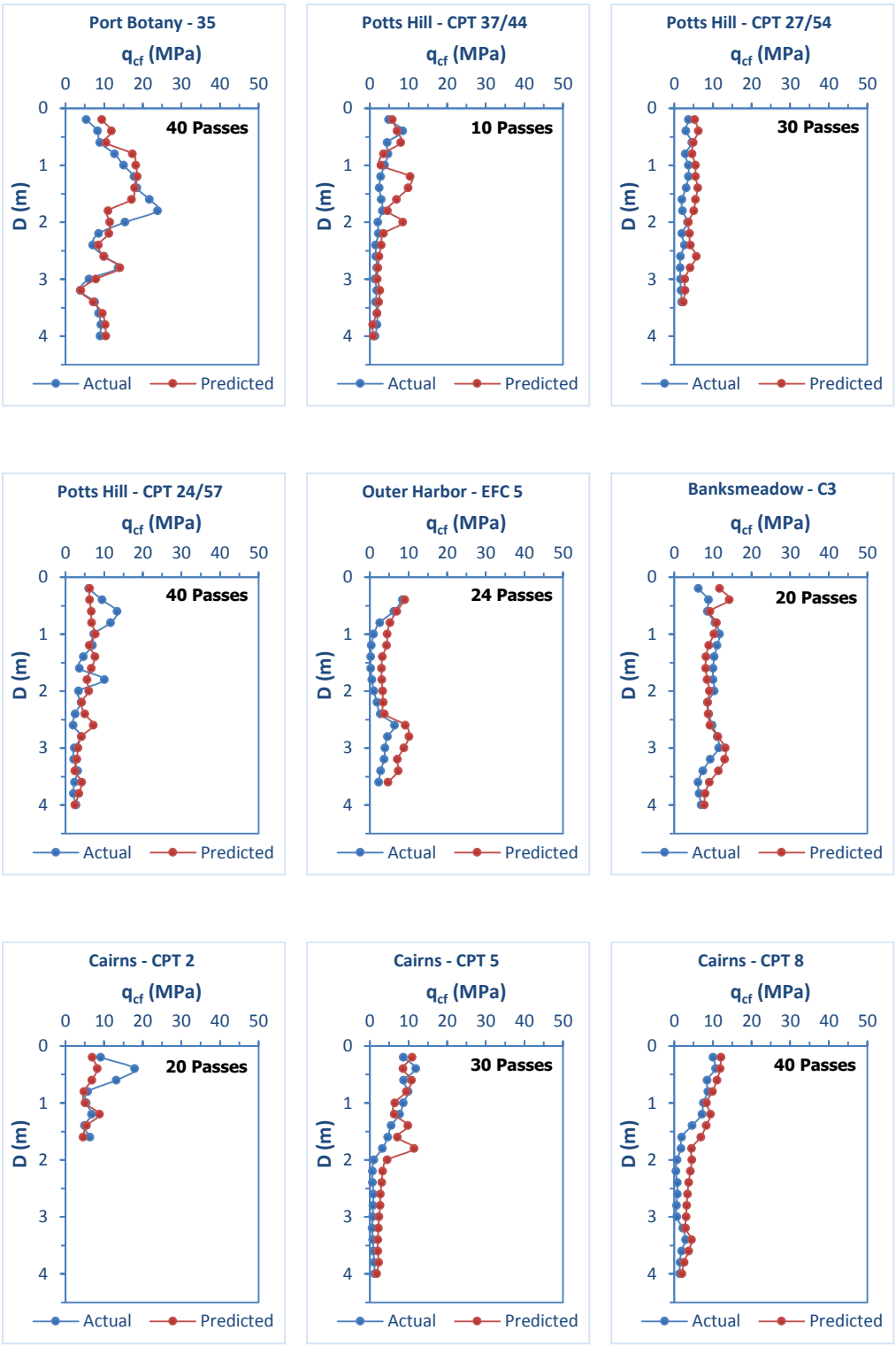


Figure 4.5 Plots of actual and model predicted CPT results (continued from previous page).

4.4.3 Robustness of the Optimum ANN Model

Finally, the validity and the accuracy of the optimal model are tested by examining how well the model predicts outputs that are consistent with the underlying physical behaviour of the system. Therefore, to assess the generalisation ability and the robustness of the selected optimal ANN model, a parametric study is carried out as recommended by Shahin et al. (2005a). This involves investigating the response of the model output using a synthetic input data set, where the input variables are varied one at a time, while all other input variables remain at a pre-specified value. For this study, the output, q_{cf}, is examined while the input variables are varied in turn as follows: q_{ci} = 2, 5, 8, 15, 20 MPa, f_{si} = 50, 100, 150, 200 kPa and P = 10, 20, 30, 40. However, as mentioned earlier, ANNs perform best when they are used to interpolate (Flood and Kartam, 1994). Thus, the generated, hypothetical input variables are specified so that they lie within the ranges of the data used in the ANN model development. The resulting model predictions are presented in Figures 4.6 – 4.9.

It is observed that the model predictions from the parametric study are in a good agreement with the expected behaviour of RDC compaction. For instance, as illustrated in Figure 4.6(a), there is a consistent increase in q_{cf} as the number of roller passes, P, increase from 10 to 40, while q_{ci} and f_{si} remain constant at 2 MPa and 50 kPa, respectively. These model predictions clearly demonstrate the improvement in strength resulting from the increasing number of roller passes at the same given location. Similarly, as shown in Figures 4.6(b) – (d), increasing trends of q_{cf} with increasing P are also observed at different f_{si} values varying between 100 and 200 kPa while q_{ci} remains constant at 2 MPa. In addition, Figures 4.6 – 4.8 indicate that q_{cf} is less affected by f_{si}, resulting in only modest changes in the predicted q_{cf}. Nevertheless, it is evident from the parametric study that the distinct non-linear relationship between q_{cf} and P has been appropriately captured by the developed ANN model.

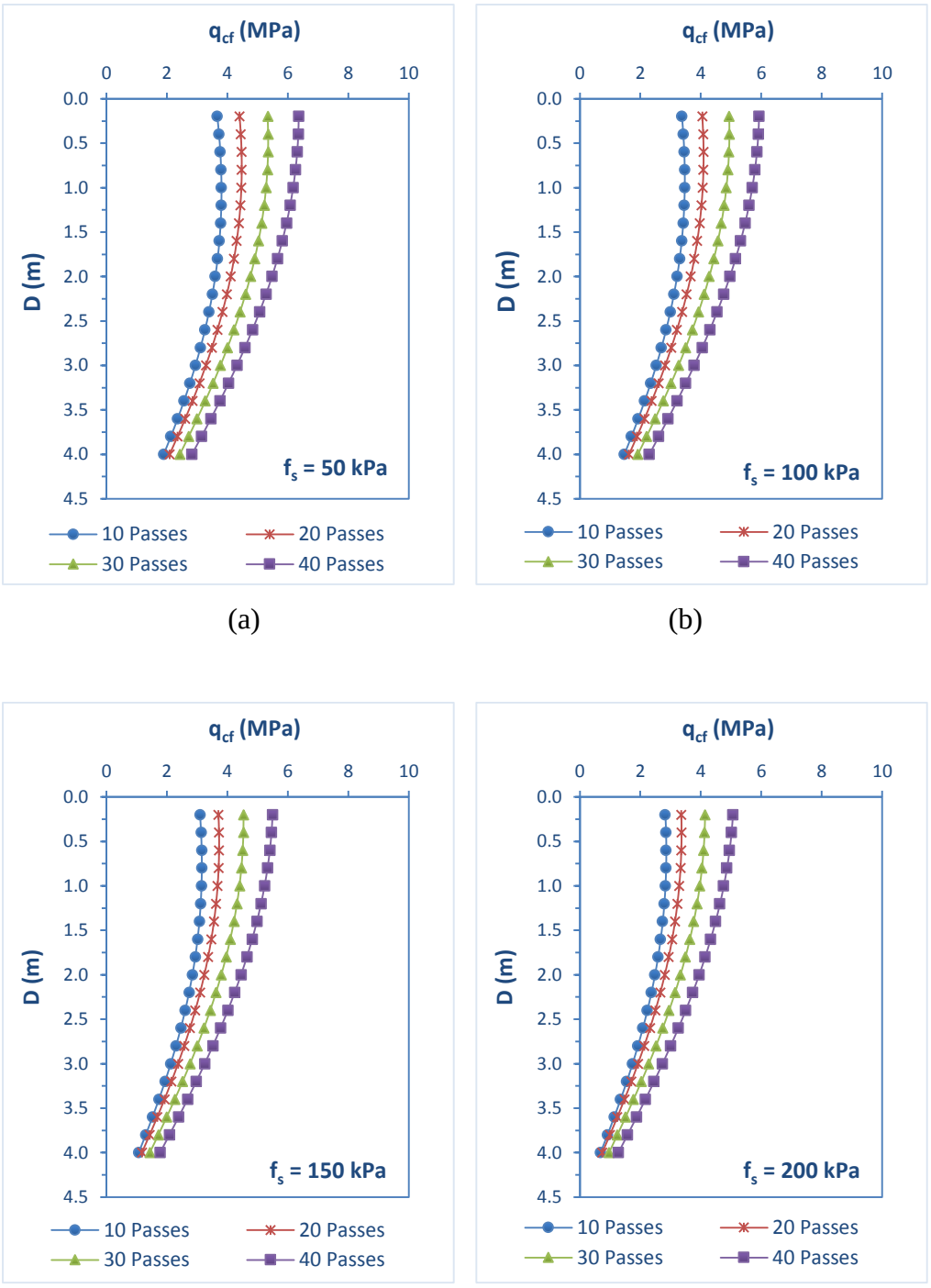


Figure 4.6 Variation of q_{cf} with different number of roller passes at q_{ci} = 2 MPa and: (a) f_{si} = 50 kPa; (b) f_{si} = 100 kPa; (c) f_{si} = 150 kPa; and (d) f_{si} = 200 kPa.

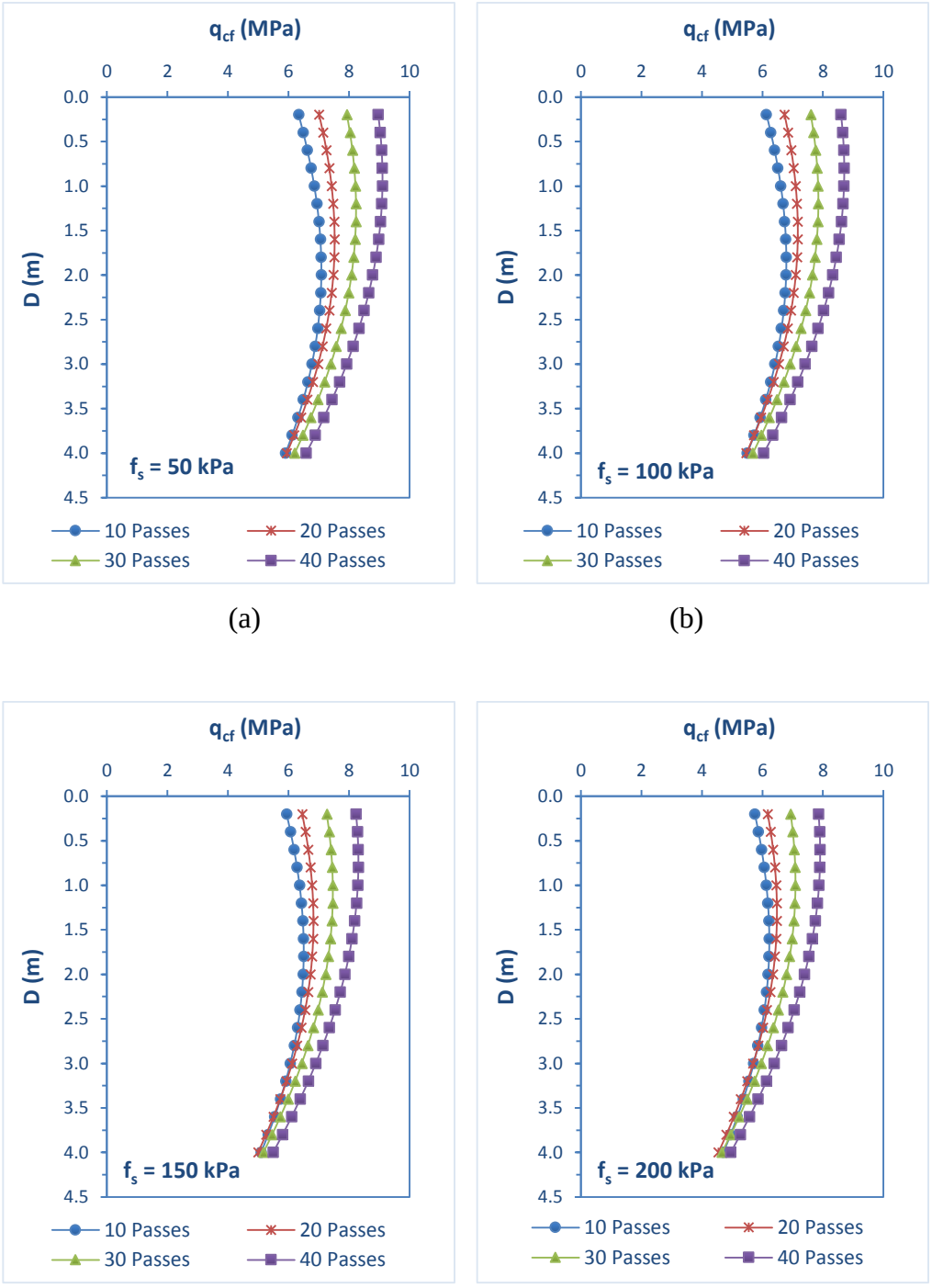


Figure 4.7 Variation of q_{cf} with different number of roller passes at q_{ci} = 5 MPa and: (a) f_{si} = 50 kPa; (b) f_{si} = 100 kPa; (c) f_{si} = 150 kPa; and (d) f_{si} = 200 kPa.

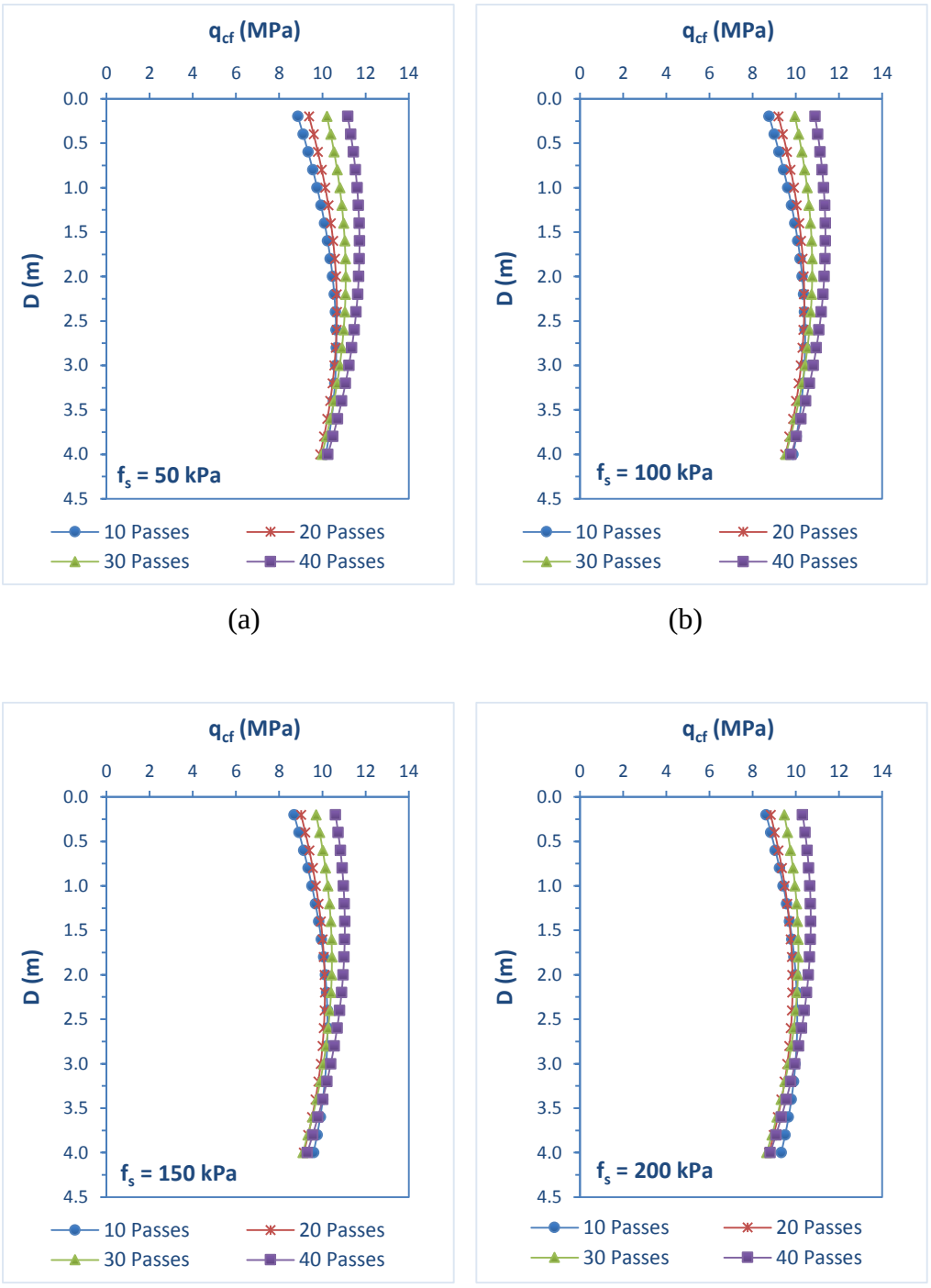


Figure 4.8 Variation of q_{cf} with different number of roller passes at q_{ci} = 8 MPa and: (a) f_{si} = 50 kPa; (b) f_{si} = 100 kPa; (c) f_{si} = 150 kPa; and (d) f_{si} = 200 kPa.

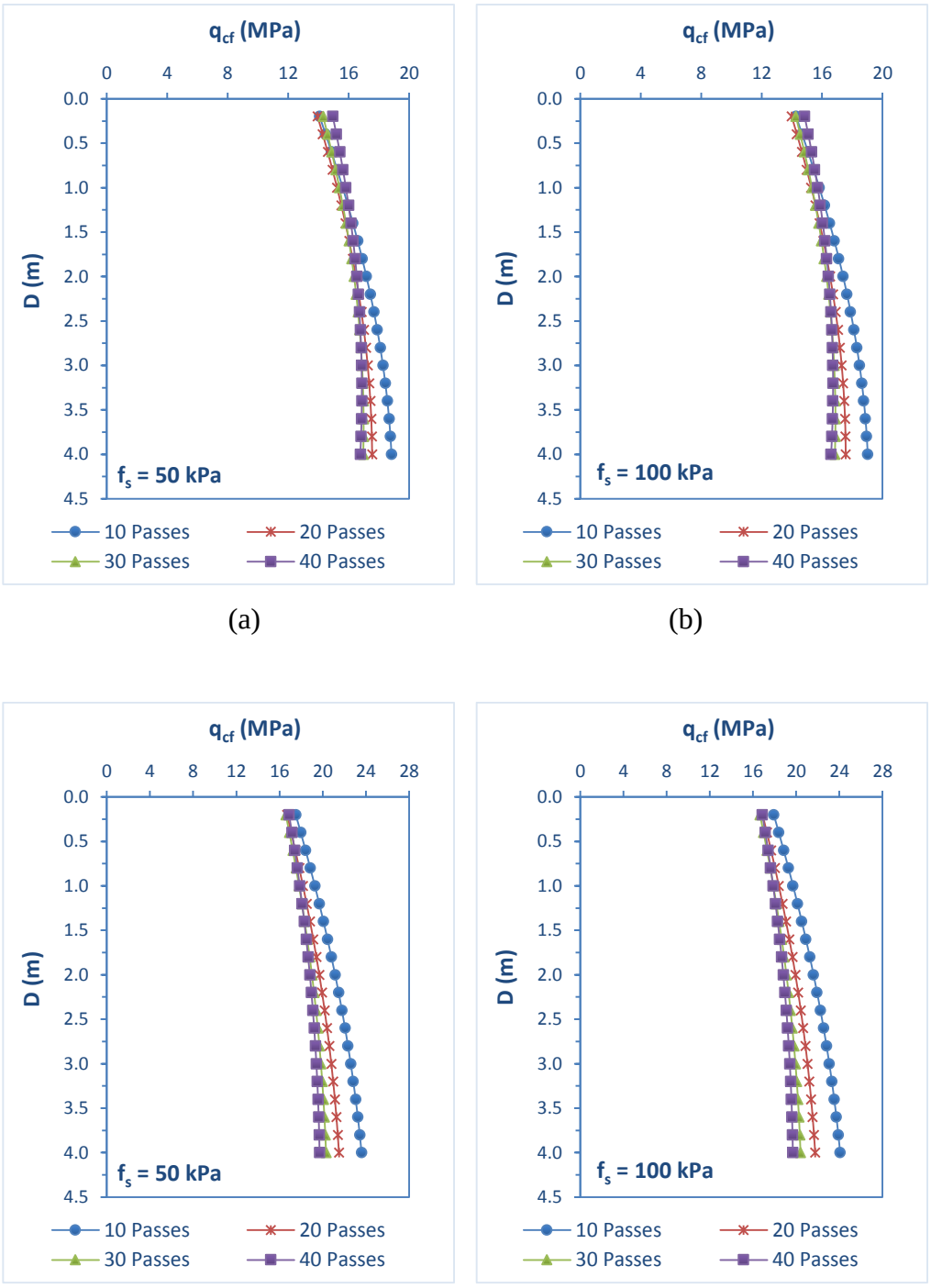


Figure 4.9 Variation of q_{cf} with different number of roller passes at (a) q_{ci} = 15 MPa and f_{si} = 50 kPa, (b) q_{ci} = 15 MPa and f_{si} = 100 kPa, (c) q_{ci} = 20 MPa and f_{si} = 50 kPa, (d) q_{ci} = 20 MPa and f_{si} = 100 kPa.

It is noted, from Figures 4.6 – 4.8, that the predictions of q_{cf} yield appropriate trends with increasing numbers of roller passes and depth. In contrast, however, Figure 4.9 displays less satisfactory predictions when the q_{ci} values are relatively large, e.g. 15 and 20 MPa, when the ground is initially quite dense. This may be attributed to the fact that the ANN model has been calibrated successfully mostly for lower values of q_{ci}. Figure 4.10 shows the histograms of both the q_{ci} and q_{cf} records included in the training dataset and which have been used in the model calibration. The dotted vertical lines indicate the mean values of q_{ci} and q_{cf} values in the training subset, which are 9.33 and 10.42, respectively. As indicated, the histograms for the q_c records contain a higher proportion in the lower range with respect to the mean values. Moreover, the data distributions are skewed to the left-hand side and exhibit sharp peaks rather than being normally distributed around the mean. Hence, as a result of this, it is likely that the model predictions are less satisfactory with regards to soil with high q_{ci} values because of the paucity of such data available from previous RDC projects. This is likely due to the fact that RDC is normally applied to ground which is initially loose and not dense. Hence, it is suggested that the optimal ANN model can be used with confidence when the q_{ci} values are below 10 MPa.

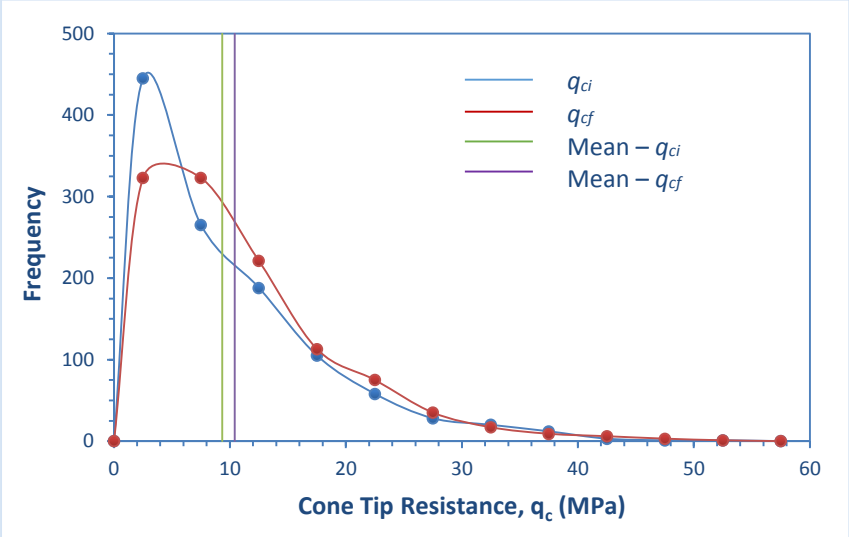


Figure 4.10 Data distribution of cone tip resistance values in training data set.

4.4.4 MLP-based Equation

Having finalised the optimal MLP model, in order to facilitate its adoption in practice, it is disseminated as a series of simplified equations. The optimal model structure is shown in Figure 4.11 and the associated weights and biases are presented in Table 4.7. These weights are utilised in the development of the equation resulting from the ANN model.

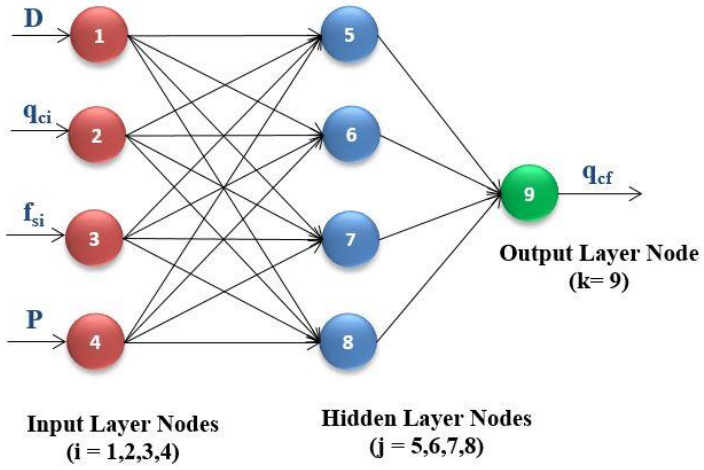


Figure 4.11 The structure of the optimum four hidden nodes network.

Table 4.7 Weights and threshold levels of the optimum ANN model.

Hidden nodes	**Weight from node *i* in the input layer to node *j* in the hidden layer (w_{ji})**				**Hidden layer threshold (θ_j)**
	***i* =1**	***i* =2**	***i* =3**	***i* =4**	
***j* = 5**	1.190	–8.585	0.754	–0.128	–1.200
***j* = 6**	–1.677	–1.592	0.310	–1.300	0.224
***j* = 7**	–0.709	–1.474	0.157	–0.369	–0.232
***j* = 8**	0.123	3.546	1.211	–2.277	–3.149
Output nodes	**Weight from node *j* in the hidden layer to node *k* in the output layer (w_{kj})**				**Output layer threshold (θ_k)**
	***j* = 5**	***j* = 6**	***j* = 7**	***j* = 8**	
***k* = 9**	–6.172	–1.819	–0.916	3.820	–0.113

The equation, which relates the input and output variables (Das et al., 2011) can be written as follow:

$$y_{k=9} = f_{sig}\left\{\theta_k + \sum_{j=5}^{8}\left[w_{kj} \times f_{sig}\left(\theta_j + \sum_{i=1}^{4}\left(w_{ji}\ x_i\right)\right)\right]\right\} \tag{4.1}$$

Where, y_k is the single output variable, θ_k is the threshold value of the k^{th} output node in the output layer and w_{kj} is the connection weight between the j^{th} node in the hidden layer and the k^{th} node in the output layer. Similarly, θ_j is the threshold value of the j^{th} hidden node and w_{ji} is the connection weight between the i^{th} input node and the j^{th} hidden node. The parameter x_i is the i^{th} input variable and f_{sig} is the sigmoid transfer function.

Equation 4.1 can be simplified as follows:

$$y_{k=9} = \frac{1}{1+e^{-\left[\theta_k + \sum_{j=5}^{8}\left(w_{kj}\ T_j\right)\right]}} \tag{4.2}$$

and

$$T_{j=5,\ldots,8} = \frac{1}{1+e^{-\left[\theta_j + \sum_{i=1}^{4}\left(w_{ji}\ x_i\right)\right]}} \tag{4.3}$$

The variables x_1, x_2, x_3 and x_4 represent the depth below the ground surface (m), q_{ci} (MPa), f_{si} (kPa) and the number of roller passes, P, respectively. However, it should be noted, as mentioned above, that the input variables are required to be scaled down prior to them being used in Equations (4.2) and (4.3). Therefore, the actual input variables ($x_{unscaled}$) are scaled between 0.1 and 0.9 using Equation (4.4) in accordance with the given extremes of the training dataset which has been shown in Table 4.3.

$$x_{scaled} = a + \frac{(x_{unscaled} - A)(b-a)}{(B-A)} \tag{4.4}$$

Where, A and B are the minimum and maximum values of the unscaled dataset, respectively. Similarly, the minimum and maximum values of the scaled dataset are denoted by *a* and *b*, which are equal to 0.1 and 0.9 in this study.

With reference to Figure 4.11 and according to the connection weights in Table 4.7, the mathematical equation for the optimal ANN model containing 4 hidden nodes can be re-written as follows:

$$q_{cf} = \frac{62.738}{1 + \exp(6.172\,T_5 + 1.819\,T_6 + 0.916T_7 - 3.82\,T_8 + 0.113)} - 6.104$$

And,

$$T_5 = [1 + \exp(-0.251\,x_1 + 0.137x_2 - 0.002x_3 + 0.003x_4 + 1.889)]^{-1}$$
$$T_6 = [1 + \exp(0.354x_1 + 0.025x_2 - 0.001x_3 + 0.03x_4 - 0.023)]^{-1}$$
$$T_7 = [1 + \exp(0.15x_1 + 0.024x_2 - 0.0003x_3 + 0.008x_4 + 0.395)]^{-1}$$
$$T_8 = [1 + \exp(-0.026\,x_1 - 0.057x_2 - 0.002\,x_3 + 0.052x_4 + 2.647)]^{-1} \quad (4.5)$$

4.5 SUMMARY AND CONCLUSIONS

This paper presents a new and unique model to predict the performance of rolling dynamic compaction (RDC) based on the artificial intelligence technique known as artificial neural networks (ANNs). The model is developed using a database consisting of cone penetration test (CPT) results obtained from several ground improvement projects which employed the 4-sided, 8 tonne 'impact roller'. The model utilises 4 input variables, including depth (m), cone tip resistance prior to compaction (MPa), sleeve friction prior to compaction (kPa) and the number of roller passes to obtain the cone tip resistance after compaction (MPa), as the sole output.

The resulting optimal ANN model yields relatively accurate predictions with a coefficient of correlation (*R*) of 0.86 and a root mean square error (*RMSE*) of 4.16 MPa, when validated against a set of unseen data. The performance of the optimal

ANN model has been further verified by introducing a series of complete and unseen CPT soundings. The resulting model predictions are in very good agreement with the actual CPT records. The robustness of the optimal model is further investigated by conducting a parametric study and it is observed that the predicted model outputs agree well with the underlying physical behaviour of the system. It is concluded that the selected optimal network is robust and reliable for values of initial cone resistance (q_{ci}) less than or equal to 10 MPa. In order to disseminate the model and to facilitate its use in practice, it is expressed as a series of tractable equations which can be incorporated into a spreadsheet or calculated by hand. As with all ANN models, they perform best when interpolating within the data ranges used in the model's development. These are given in the body of the paper.

The model is intended to provide initial predictions for planning purposes and not to replace field trials, which will yield more accurate and site-specific conclusions. As with all ANN models, its accuracy can be improved by incorporating data from additional RDC-related projects.

Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP120101761). The authors wish to acknowledge Mr. Stuart Bowes at Broons Hire (SA) Pty Ltd for his kind assistance and continuing support, especially in providing access to the in situ test results upon which the numerical models are based. The authors are also grateful to Mr. Brendan Scott for his contribution to this work.

Chapter 5.

Application of Artificial Neural Networks for Predicting the Impact of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results

INTENTIONALLY BLANK

Statement of Authorship

Title of Paper	Application of Artificial Neural Networks for Predicting the Impact of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results.
Publication Status	☐ Published ☑ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Application of Artificial Neural Networks for Predicting the Impact of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results. Journal of Rock Mechanics and Geotechnical Engineering, In Press.

Principal Author

Name of Principal Author (Candidate)	R. A. T. M. Ranasinghe		
Contribution to the Paper	Data collection and pre-processing, Model development based on Artificial Neural Networks (ANNs), Analysis of results and interpretation, and Preparation of the manuscript.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	20/02/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

iv. the candidate's stated contribution to the publication is accurate (as detailed above);

v. permission is granted for the candidate in include the publication in the thesis; and

vi. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mark B. Jaksa		
Contribution to the Paper	Overall percentage – 10% Initial concept, research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Yien Lik Kuo		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Fereydoon Pooya Nejad		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

ABSTRACT

Rolling dynamic compaction (RDC), which involves the towing of a non-circular module, is now widespread and accepted among many other soil compaction methods. However, to date, there is no accurate method for reliable prediction of the densification of soil and the extent of ground improvement by means of RDC. This study presents the application of artificial neural networks (ANNs) for a *priori* prediction of the effectiveness of RDC. The models are trained with in situ dynamic cone penetration (DCP) test data obtained from previous civil projects associated with the 4-sided 'impact roller'. The predictions from the ANN models are in good agreement with the measured field data, as indicated by the model correlation coefficient of approximately 0.8. It is concluded that the ANN models developed in this study can be successfully employed to provide more accurate prediction of the performance of the RDC on a range of soil types.

Keywords: *Rolling dynamic compaction, ground improvement, artificial neural network, dynamic cone penetration test*

5.1 INTRODUCTION

Soil compaction is one of the major activities in geotechnical engineering applications. Among many other soil compaction methods, rolling dynamic compaction (RDC) is now becoming more widespread and accepted internationally. The RDC technology emerged with the first full-sized impact roller from South Africa for the purpose of improving sites underlain by collapsible sands in 1955 (Avalle, 2004d). Over the years, the RDC concept has been refined with updated and improved mechanisms. Since the mid-1980s, impact rollers have been commercially available and are now adopted internationally using module designs incorporating 3, 4 and 5 sides.

The 4-sided impact roller module consists of a steel shell filled with concrete to produce a heavy, solid mass (6–12 tonnes), which is towed within its frame by a four-wheeled tractor (Figure 5.1). When the impact roller traverses the ground, the module rotates eccentrically about its corners and derives its energy from three sources: (1) potential energy from the static self-weight of the module; (2) additional potential energy from being lifted about its corners; and (3) kinetic energy developed from being drawn along the ground at a speed of 9–12 km/h. As a result, the impact roller is capable of imparting a greater amount of compactive effort to the soil, which often leads to a deeper influence depth; i.e. in excess of 3 m below the ground surface in some soils (Avalle, 2006; Jaksa et al., 2012), which is much deeper than the 0.3 m to 0.5 m generally achieved using traditional vibratory and static rollers (Clegg and Berrangé, 1971; Clifford, 1976, 1978b). Furthermore, it is able to compact thicker lifts, in excess of 500 mm, which is considerably greater than the usual layer thicknesses of between 200 mm and 500 mm (Avalle, 2006) and can also operate with larger particle sizes.

Moreover, RDC is more efficient since the module traverses the ground at a higher speed; about 9–12 km/h compared with traditional vibratory rollers which operate at around 4 km/h (Pinard, 1999). This creates approximately two module impacts over the ground each second (Avalle, 2004d). Thus, the faster operating speed and deeper compactive effort make this method very effective for bulk earthworks. In addition, it also appears that prudent use of RDC can provide significant cost savings in the civil

construction sector. Due to these inherent characteristics of RDC, modern ground improvement specifications often replace or provide an alternative to traditional compaction equipment. It has been demonstrated to be successful in many applications worldwide, particularly in civil and mining projects, pavement rehabilitation and in the agricultural sector (Avalle, 2004d, 2006; Jaksa et al., 2012).

Figure 5.1 The 4-sided 'impact roller' and tractor.

To date, a significant amount of data has been gathered from RDC projects through an extensive number of field and case studies in a variety of ground conditions. However, these data have yet to be examined holistically and there currently exists no method, theoretical or empirical, for determining the improvement in in situ density of the ground at depth as a result of RDC using dynamic cone penetrometer (DCP) test data. The complex nature of the operation of the 4-sided impact roller, as well as the consequent behaviour of the ground, has meant that the development of an accurate theoretical model remains elusive. However, recent work by the authors (Ranasinghe et al., 2016c) in relation to RDC, as well as by others in the broader geotechnical engineering context (Shahin and Jaksa, 2006; Günaydın, 2009; Kuo et al., 2009; Pooya Nejad et al., 2009; Isik and Ozden, 2013), have demonstrated that artificial intelligence (AI) techniques, such as artificial neural networks (ANNs), show great promise in this regard.

In a recent and separate study by the authors (Ranasinghe et al., 2016c), ANNs have been applied to predict the effectiveness of RDC using cone penetration test (CPT) data in relation to the 4-sided impact roller. The model, based on a multi-layer perceptron (MLP), incorporates 4 input parameters, the depth of measurement (D), the CPT cone tip resistance (q_{ci}) and sleeve friction (f_{si}) prior to compaction, and the number of roller passes (P). The model predicts a single output variable, namely the cone tip resistance (q_{cf}) at depth, D, after the application of P roller passes. The ANN model architecture, hence, consists of 4 input nodes, a single output node, and the optimal model incorporates a single hidden layer with 4 hidden nodes. The authors also translated the ANN model into a tractable equation, which was shown to yield reliable predictions with respect to the validation data set.

This paper aims to develop an accurate tool for predicting the performance of RDC in a range of ground conditions. Specifically, the tool is based on ANNs using dynamic cone penetrometer (DCP) test data (Standards Association of Australia, 1997) obtained from a range of projects associated with the Broons BH-1300, 8 tonne, 4-sided impact roller, shown previously in Figure 5.1. Whilst the DCP is a less reliable test than the CPT, it is nevertheless used widely in geotechnical engineering practice and a model which provides reliable predictions of RDC performance based on DCP data is likely to be extremely valuable to industry.

5.2 ANN MODEL DEVELOPMENT

In recent years, ANNs have been extensively used in modelling a wide range of engineering problems associated with non-linearity and have demonstrated extremely reliable predictive capability. Unlike statistical modelling, ANNs is a data-driven approach and hence does not require prior knowledge of the underlying relationships of the variables (Shahin et al., 2002). Moreover, these non-linear parametric models are capable of approximating any continuous input-output relationship (Onoda, 1995). A comprehensive description of ANN theory, structure and operation is beyond the scope of the paper, but is readily available in the literature (Hecht-Nielsen, 1989; Fausett, 1994; Ripley, 1994; Shahin, 2016).

In this study, the ANN models for predicting the effectiveness of RDC are developed using the PC-based software *NEUFRAME* version 4.0 (2000). As mentioned above, the data used for ANN model calibration and validation incorporate DCP test results obtained from several ground improvement projects using the Broons BH-1300, 4-sided impact roller, which has a static mass of 8 tonnes. The data used in this study are summarised in Table 5.1.

Table 5.1 Summary of the database of DCP records.

No.	Project	No. of DCP soundings	Soil type		No. of roller passes
			Primary	Secondary	
1	Arndell Park	23	Clay	Silt	0, 5, 10, 20, 25, 30
2	Banyo	2	Clay	Silt	4, 8, 16
3	Banksmeadow	10	Sand	None	0, 10, 20
4	Ferguson	7	Clay	Silt	5, 10, 15
5	Kununurra	5	Sand	None	0, 5, 10, 20, 25, 30, 40, 50, 60
6	Monarto	6	Sand	Gravel	0, 5, 10, 30
7	Outer Harbor	9	Clay Sand	Silt Gravel	0, 6, 12, 18, 24
8	Pelican Point	8	Clay	Silt	0, 6, 12, 18
9	Penrith	39	Sand	Clay	0, 2, 4, 6, 10, 20
10	Potts Hill	4	Clay	Silt	0, 10, 20, 30, 40
11	Revesby	4	Clay Sand Sand	Silt Clay None	0, 5, 10, 15
12	Whyalla	12	Sand	Gravel	0, 8, 16

It is important to note that the DCP data are obtained at effectively the same location prior to RDC (i.e. 0 pass) and after several passes of the module (e.g. 10, 20 passes), since it is essential to include both pre- and post-compaction conditions in the ANN model simulations. In total, the database contains 2,048 DCP records from 12 projects.

ANN model development is carried out using the process outlined by (Maier et al., 2010), including determination of appropriate model inputs/outputs, data division, selection of model architecture, model optimisation, validation and measures of performance. This methodology is briefly discussed and contextualised below.

5.2.1 Selection of Appropriate Model Inputs and Outputs

The most common approach for the selection of data inputs in geotechnical engineering is based on the prior knowledge of the system in question and this is also adopted in the present study. Therefore, the input/output variables of the ANN models are chosen in such a manner that they address the main factors that influence RDC behaviour. It is identified that the degree of soil compaction depends upon a number of key parameters, including: the geotechnical properties at the time of compaction, such as ground density, moisture content, soil type; and the amount of energy imparted to the ground during compaction.

As mentioned previously, in this study the ANN model is based on DCP test results collected from a range of ground improvement projects involving the 4-sided impact roller. The DCP (Standards Association of Australia, 1997) is one of the most commonly used in situ test methods available, which provides an indication of soil strength in terms of rate of penetration (blows/mm). In this study, the average DCP blow count per 300 mm is used as a measure of the average density improvement with depth as a result of RDC.

Moisture content is not routinely measured in ground improvement projects in practice. Nevertheless, moisture content is considered to be implicitly included in the DCP data, as the number of blows per 300 mm is affected by moisture content. In

addition, whilst the natural ground is often characterised as part of site investigations associated with earthworks projects, soil characterisation during the process of filling and compacting is not. However, in order to include the soil type in the ANN model, a generalised soil type is defined at each DCP location by adopting primary (dominant) and secondary soil types. The ground improvement projects included in the database can each be characterised into one of 4 distinct soil types: (i) Sand–Clay; (ii) Clay–Silt; (iii) Sand–None and (iv) Sand–Gravel. As *NEUFRAME* requires the allocation of one input node for every text parameter, therefore in this model, the soil type variable represents 4 input nodes.

Hence, in summary, the ANN prediction models developed in this study each have a total of 5 input variables consisting of 8 nodes, together representing: (1) Soil type: (a) Sand–Clay, (b) Clay–Silt, (c) Sand–None, (d) Sand–Gravel; (2) Average depth below the ground surface, D (m); (3) Initial number of roller passes; (4) Initial DCP count (blows/300 mm); and (5) Final number of roller passes. The single output variable is the Final DCP count (blows/300 mm) at depth D after compaction.

5.2.2 Data Division and Pre-processing

In this study, the commonly adopted cross-validation technique (Stone, 1974) is used as the stopping criterion, which requires the entire data set to be divided into 3 subsets: (1) a training set, (2) a testing set, and (3) a validation set. The training set contains 80% of the data (1,629 records), whereas the remaining 20% (419 records) is allocated to the validation set. The training set is further subdivided into the training and testing sets in the proportion of 80% (1,310 records) and 20% (319 records), respectively. The application of these 3 individual subsets is discussed later.

The distribution of data among the 3 subsets may have a significant impact on model performance (Shahin et al., 2004). Therefore, it is necessary to divide the data into 3 subsets in such a way that they represent the same statistical population exhibiting similar statistical properties (Masters, 1993). The statistical properties considered in this study include the mean, standard deviation, minimum, maximum and range. The

present study uses the method of self-organizing maps (SOMs) (Bowden et al., 2002), a detailed explanation of which is given by Kohonen (1982). However, the determination of the optimal map size is an iterative process as there is no absolute rule to select the most favourable map size and thus several map sizes (e.g. 10×10, 20×20, 30×30) are investigated. Once the clusters are generated, samples are randomly selected from each cluster and assigned to each of the 3 data subsets.

Prior to model calibration, data are pre-processed in the form of scaling which ensures that each model variable receives equal attention during model training. Therefore, the output variables are scaled so that they are commensurate with the limits of the sigmoid transfer function that is used in the output layer. Although scaling of the input variables is not necessarily important, as recommended by Masters (1993), in this study, they are also subjected to scaling similar to that for the output variable. In such a way, all the variables are scaled into the selected range of 0.1 to 0.9 by using Equation 5.1. However, subsequent to model training, the model outputs undergo reverse scaling.

$$x_{scaled} = a + \frac{(x_{unscaled} - A)(b-a)}{(B-A)} \tag{5.1}$$

Where, A and B are the minimum and maximum values of the unscaled dataset, respectively, and similarly, a and b are the minimum and maximum values of the scaled data set.

5.2.3 Determination of Network Architecture

The determination of network architecture includes the selection of model geometry and the manner in which information flows through the network. Among many other different types of network architectures, the fully inter-connected, feed forward type, multi-layer perceptrons (MLPs) are the most common form used in prediction and forecasting applications (Maier and Dandy, 2000). To date, feed forward networks have been successfully applied to many and varied geotechnical engineering problems (Günaydın, 2009; Kuo et al., 2009; Pooya Nejad et al., 2009).

Network geometry requires the determination of the number of hidden layers and the number of nodes incorporated in each layer. The simplest form of MLP, which is used in this study, consists of 3 layers, including a single hidden layer between the input and output layers. It has been shown that single, hidden layer networks with sufficient connection weights are capable of approximating any continuous function (Cybenko, 1989; Hornik et al., 1989). The ability to use non-linear activation functions in the hidden and output layers allows the MLP to capture the complexity and non-linearity of the system in question.

The number of nodes in the input and output layers is restricted by the number of model input and output variables. As mentioned above, this model consists of 8 nodes in the input layer and a single node in the output layer. Selection of the optimal number of hidden layer nodes is again an iterative process. If too few nodes are adopted, the predictive performance to the model is compromised, whereas, if too many nodes are used, the model may be overfitted and thus lack the ability to generalise. The stepwise approach (Shahin et al., 2002) is adopted in this study to obtain the optimal architecture where several ANN models are trained, starting from the simplest form with a single hidden layer node model and successively increasing the number of nodes to 11. According to Caudill (1988), the upper limit of hidden nodes which are needed to map any continuous function for a network with I input nodes is equal to $(2I + 1)$.

5.2.4 Model Optimisation

In this study, model optimisation, which involves evaluating the optimum weight combination for the ANN, is carried out using the back-propagation algorithm (Rumelhart et al., 1986). It is the most widely used optimisation algorithm in feed-forward neural networks and has been successfully implemented in many geotechnical engineering applications (Shahin and Jaksa, 2006; Günaydın, 2009; Pooya Nejad et al., 2009). The back-propagation algorithm is based on the first-order gradient descent rule and has the capability of escaping local minima having appropriately defined the ANNs' internal parameters (Maier and Dandy, 1998). The approach adopted in this study involves the models, consisting of each trial number of hidden nodes, first being

trained with the default parameter values (i.e. learning rate = 0.2, momentum term = 0.8) assigned to a random initial weight configuration. The models are then retrained with different combinations of learning rates and momentum terms and the network performance is assessed with respect to the validation set. However, the networks are vulnerable to being trapped in local minima if training is initiated from an unfavourable position in the weight space (Shahin et al., 2003). Therefore, the selected network with optimal parameters is retrained several times and allowed to randomise the initial weight configuration to ensure that model training does not cease at a sub-optimal level.

5.2.5 Stopping Criteria

The stopping criterion is used to determine when to cease the ANN model training phase. Since overfitting is a possibility during model training, the cross validation technique is used which, as discussed earlier, requires data division into 3 subsets: training, testing and validation. The training data are used in the model training phase where the connection weights are estimated. The models are considered to achieve the optimal generalisation ability when the error measure, with respect to the testing set, is a minimum, having ensured that the training and testing sets are representative of the same statistical population. Although the testing set error shows a reduction at the beginning, it starts to increase when overfitting occurs. Therefore, the optimal network is obtained at the onset of the increase in test data error, assuming the error surface converges at the global minimum. However, model training is continued for some time, even after the testing error starts to increase initially, to ensure that the model is not trapped in local minima (Maier and Dandy, 2000).

5.2.6 Model Validation and Performance Measures

Once the model has been optimised, the network is validated against the independent validation set, which provides a rigorous check of the model's generalisation capability. The network is expected to generate non-linear relationships between the input and output variables rather than simply memorising the patterns that are

contained in the training data (Shahin et al., 2002). Since the model is assessed with respect to an unseen data set, the results are significant for the evaluation of network performance.

The measures used in this study in evaluating the networks' predictive performance are the often used root mean squared error (*RMSE*), mean absolute error (*MAE*) and coefficient of correlation (*R*). When using the *RMSE*, larger errors receive much greater attention than smaller errors (Hecht-Nielsen, 1989), whereas *MAE* provides information on the magnitude of the error. The coefficient of correlation is used to determine the goodness of fit and it describes the relative correlation between the predicted and actual results. The guide proposed by Smith (1993) is used as follows:

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

5.3 RESULTS AND DISCUSSION

In the following subsections, the results of data division and model optimisation are presented followed by the behaviour of the optimal network when assessed for robustness using a parametric study.

5.3.1 Results of Data Division

The SOM size of 25×25 is found to be optimal. The statistics of the 3 subsets are presented in Table 5.2. As expected, in general, the statistics are in a good agreement, apart from slight inconsistencies that result from the appearance of singular and rare events in the data, which cannot be replicated in all 3 subsets. It is accepted that ANNs are best used to interpolate within the limits of the data included in the ANN model development process and are best not used for extrapolation.

Table 5.2 ANN input and output statistics.

Model variables and data sets	Statistical parameters				
	Mean	*SD	Minimum	Maximum	Range
Average depth (m)					
Training set	0.81	0.51	0.15	1.95	1.80
Testing set	0.82	0.51	0.15	1.95	1.80
Validation set	0.83	0.52	0.15	1.95	1.80
Initial number of roller passes					
Training set	7.69	10.61	0.00	50.00	50.00
Testing set	7.65	10.44	0.00	50.00	50.00
Validation set	8.71	10.93	0.00	50.00	50.00
Initial DCP count (blows/300 mm)					
Training set	16.57	10.86	3.00	65.00	62.00
Testing set	15.88	10.64	3.00	59.00	56.00
Validation set	16.31	10.20	3.00	61.00	58.00
Final number of roller passes					
Training set	21.14	16.25	2.00	60.00	58.00
Testing set	21.16	16.49	2.00	60.00	58.00
Validation set	21.08	16.11	2.00	60.00	58.00
Final DCP count (blows/300 mm)					
Training set	18.30	11.29	2.00	84.00	82.00
Testing set	17.80	10.81	2.00	73.00	71.00
Validation set	17.93	11.47	3.00	75.00	72.00

*SD – Standard Deviation

5.3.2 Results of the Optimal ANN Model

In selecting the optimal model, several models with a single hidden layer consisting of different numbers of hidden nodes are compared with respect to *R*, *RMSE* and *MAE*. However, with the parallel aim of parsimony, a model with a smaller number of hidden nodes that performs well, with respect to the validation set and with a consistent performance with the training and testing data is considered to be optimal. From this perspective, it is observed that the model with 4 hidden nodes yields the best performance with respect to the single hidden layer ANNs.

With the intention of improving prediction accuracy, networks are examined with an additional hidden layer. Similar to the single hidden layer model optimisation, several models with different numbers of nodes in the 2 hidden layers are trained and validated. Consequently, the model with 4 and 6 hidden nodes in the first and second hidden layers respectively, is deemed to be optimal among the 2 hidden layer ANNs. The performance statistics of the selected optimal networks for single and two hidden layer networks are summarised in Table 5.3.

Table 5.3 Performance statistics of the optimal networks with single and two hidden layers.

Data set	***RMSE*** **(blows/300 mm)**	***MAE*** **(blows/300 mm)**	***R***
Single hidden layer model			
Training	6.45	4.88	0.85
Testing	6.52	4.74	0.83
Validation	7.54	5.59	0.79
Two hidden layer model			
Training	5.72	3.97	0.86
Testing	5.67	3.88	0.85
Validation	6.81	4.85	0.81

The optimal single hidden layer model is compared with the optimal two hidden layer model in terms of model accuracy and model parsimony. It is evident that the prediction accuracy of the two hidden layer model is only marginally better than that of the network with a single hidden layer, given the error difference with respect to the validation set: *RMSE* = 0.73, *MAE* = 0.74; and with the difference in correlation: *R* = 0.02. Given that the two hidden layer model sacrifices model parsimony for only marginal improvement in performance, it is decided to proceed with the single hidden layer model. This is advantageous, as will be discussed later, as this model facilitates the development of a simple numerical equation which expresses the relationship between the model inputs and output.

As produced by the optimal, single hidden layer ANN model, the plot of predicted versus measured DCP count with respect to the data in the testing and validation sets is shown in Figure 5.2, where the solid line indicates equality. According to the guide proposed by Smith (1993), it can be concluded that there exists very good correlation between the model predictions and the measured values of the Final DCP count. However, it is expected that the random errors associated with the input data, as a result of testing uncertainties [operator, procedural, equipment] (Orchant et al., 1988) have adversely affected model performance.

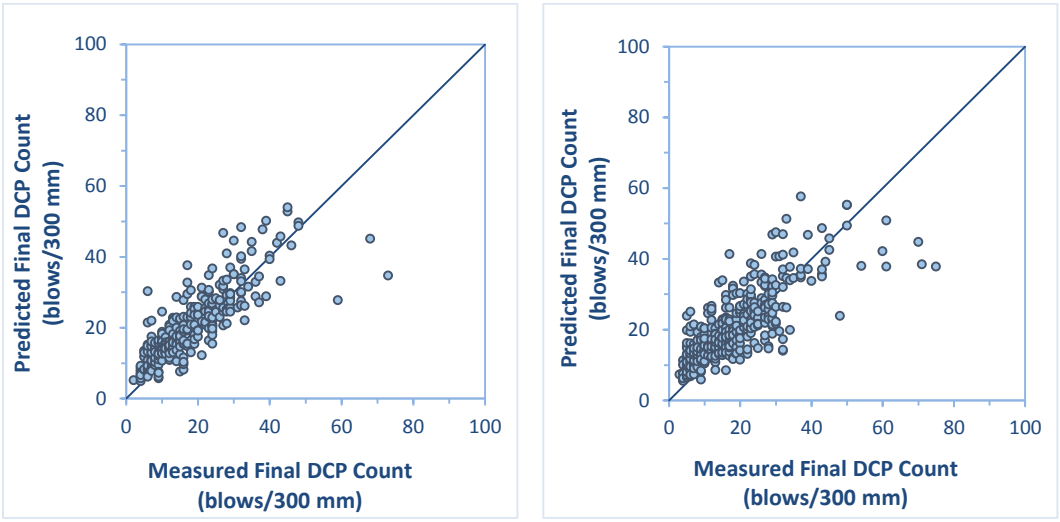
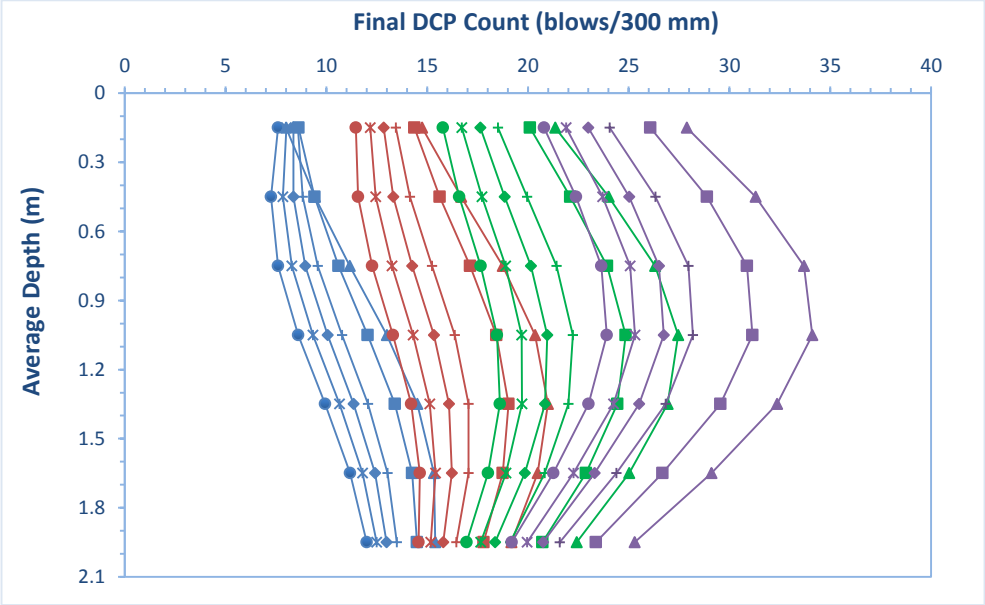


Figure 5.2 Measured versus predicted final DCP count (blows/300 mm) for the optimal ANN model with respect to: (a) testing set; and (b) validation set.

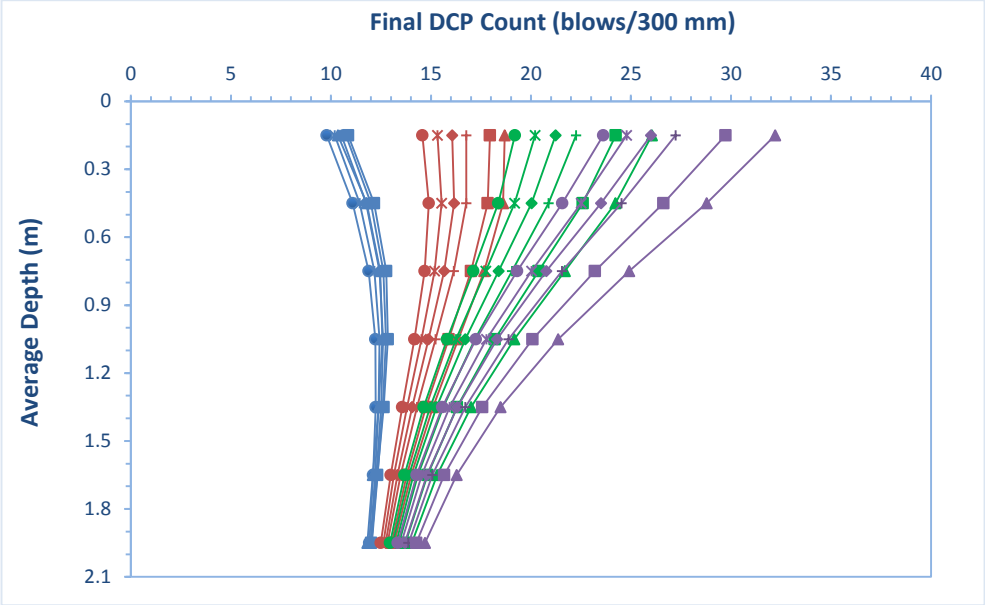
5.3.3 Robustness of the Optimal ANN Model

It is essential to conduct a parametric study in order to further confirm the validity, accuracy and generalisation ability of the optimal model. It is crucial that the model behaviour conforms to the known underlying physical behaviour of the system. Therefore, the network's generalisation ability is investigated with respect to a set of synthetic input data generated within the limits of the training data set. Each input variable is varied in succession, with all other input variables remaining constant at a pre-specified value.

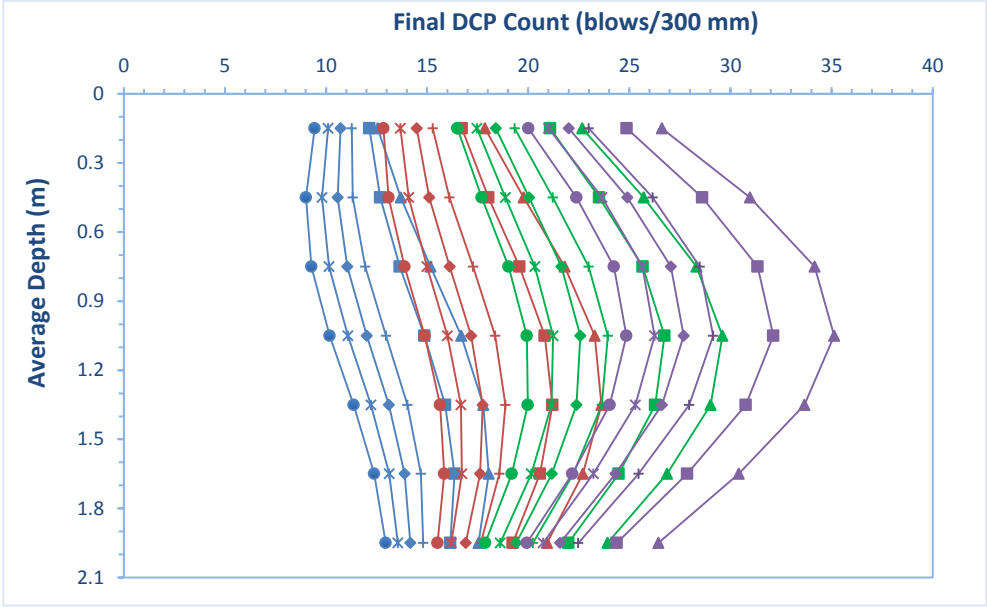
The post-compaction condition of the ground, represented by the Final DCP count, is predicted from the optimal ANN model for a given initial DCP count (i.e. 5, 10, 15, 20 blows/300 mm) in each of the different soil types (i.e. Sand–Clay, Clay–Silt, Sand–None and Sand–Gravel) for several, different numbers of roller passes (i.e. 5, 10, 15, 20, 30, 40 passes). The resulting model predictions are presented in Figure 5.3.



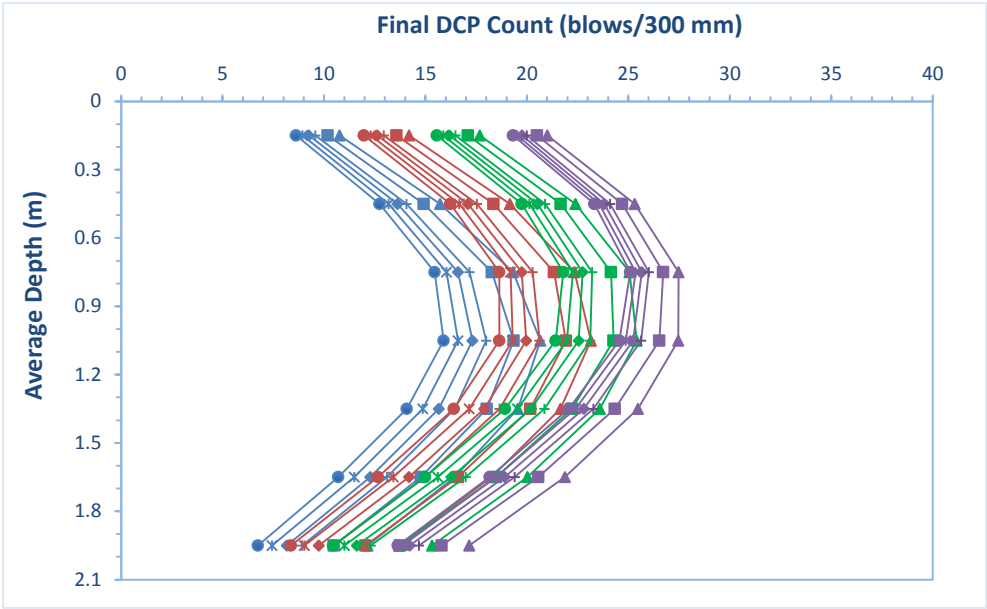
(a)



(b)



(c)



(d)

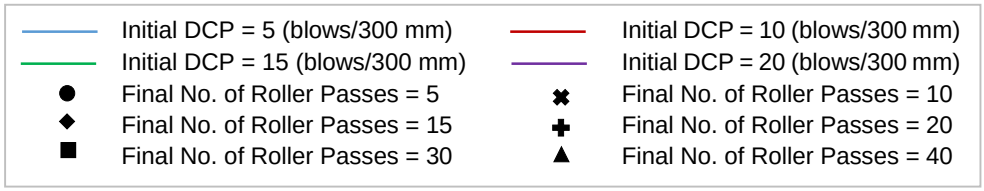


Figure 5.3 Variation of final DCP count with respect to initial DCP count and final number of roller passes in: (a) Sand–Clay; (b) Clay–Silt; (c) Sand–None; and (d) Sand–Gravel.

It is evident that the final DCP count increases with increasing numbers of roller passes, for a given initial DCP in each soil type, which confirms that the ground is significantly improved with RDC. As such, the graphs verify that the optimal ANN model predictions agree well with the expected behaviour based on the impact of RDC. In addition, there are no irregularities in behaviour, with respect to each of these variables. As a result, it is concluded that the optimal ANN model is robust when predicting the effectiveness of RDC and can be used with confidence.

Furthermore, the final DCP count is analysed over average depths between 0.45 m and 1.95 m for each soil type as a function of the number of roller passes, and the results are summarised in Figures 5.4(a) to (f). It is noted that upper 300 mm soil layer is disturbed by the action of RDC module and therefore, for this analysis, model predictions at the average depth of 0.15 m are neglected. However, in all cases, it can be seen that the final DCP count increases as the number of roller passes grows. It can be further observed that the coarse-grained soils undergo greater compaction when fine particles are present in the material. For example, in Figures 5.4(a) to (f), it is evident that, for a given initial DCP count, the final DCP count reaches higher values as the number of roller passes increases in the Sand–None and Sand–Clay soils as compared with Sand–Gravel. In addition, the final DCP count curves exhibit a higher gradient with respect to Sand–None and Sand–Clay soils than that to the Sand–Gravel. This suggests that, when sand is mixed with some fine particles, the compaction characteristics are improved when compared with sand mixed with gravel. This is consistent with conventional wisdom that some fine particles added to coarse-grained materials enhance the soil's compaction characteristics. In contrast, it can be seen that the fine-grained soils are more difficult to compact when compared with coarse-grained materials, as indicated by the relatively lower values of final DCP count for the Clay–Silt soil when compared with the Sand–None and Sand–Clay materials. Again, this is consistent with conventional wisdom.

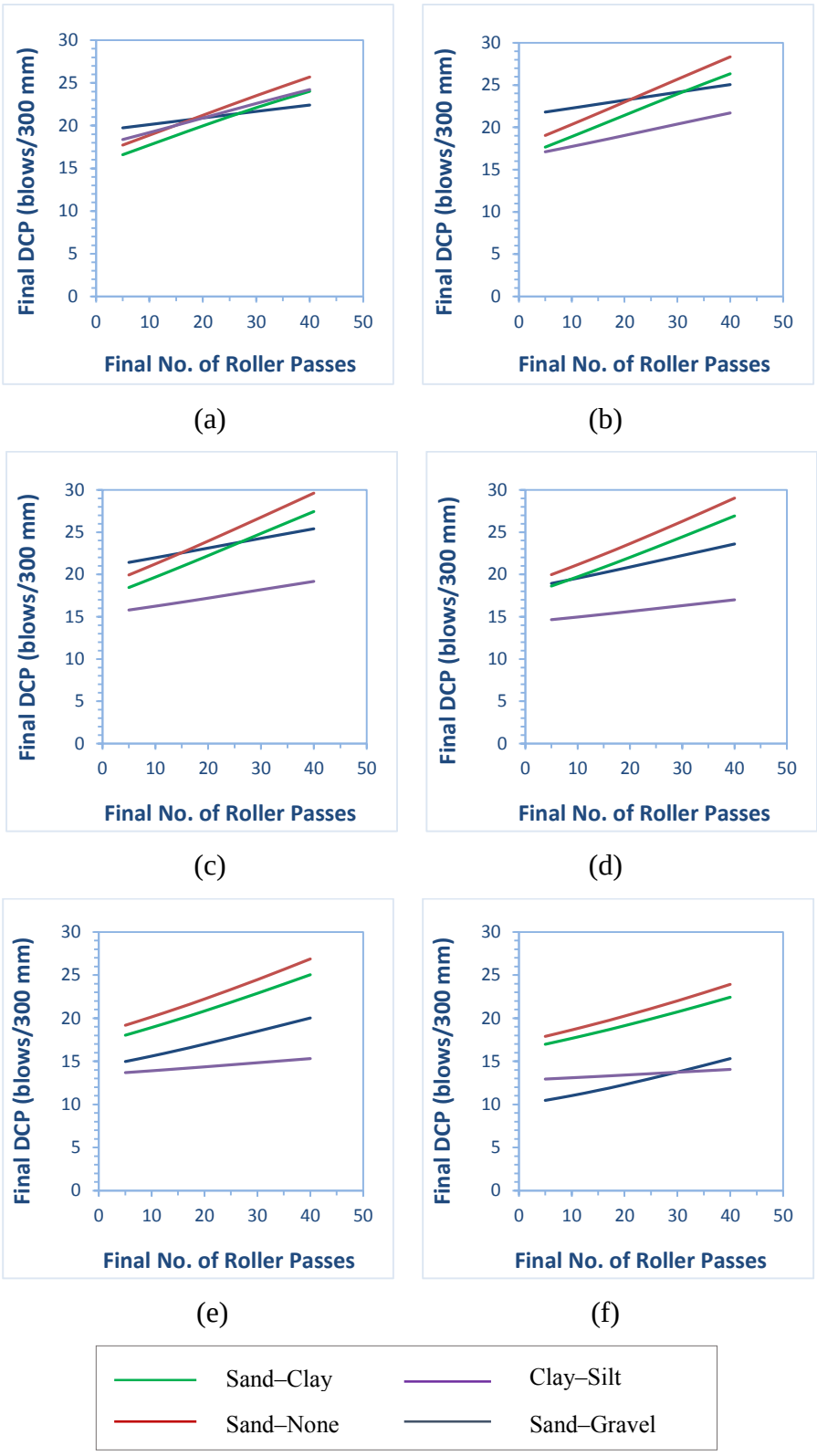


Figure 5.4 Variation of final DCP count with final number of roller passes when initial DCP count = 15 and initial passes = 0 in different soil types at depth of: (a) 0.45 m; (b) 0.75 m; (c) 1.05 m; (d) 1.35 m; (e) 1.65 m; and (f) 1.95 m.

5.3.4 MLP-based Numerical Equation

In order to facilitate the dissemination and deployment of the optimal MLP model, a relatively simple equation is developed to predict the level of ground improvement derived from RDC. The optimal model structure is shown in Figure 5.5 and the associated weights and biases are presented in Table 5.4.

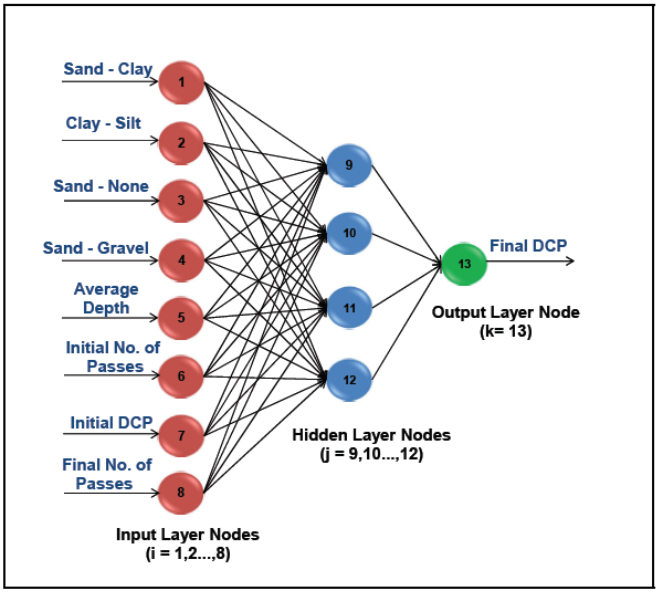


Figure 5.5 The structure of the optimal MLP model.

The numerical equation, which relates the input and output variables, can be written as:

$$y_{k=13} = f_{sig}\left\{\theta_k + \sum_{j=9}^{12}\left[w_{kj} \times f_{sig}\left(\theta_j + \sum_{i=1}^{8}\left(w_{ji}\ x_i\right)\right)\right]\right\} \qquad (5.2)$$

Where, y_k is the single output variable, i.e. the Final DCP count (blows/300 mm) at average depth D below the ground, θ_k is the threshold value of the k^{th} output node in the output layer and w_{kj} is the connection weight between the j^{th} node in the hidden layer and the k^{th} node in the output layer. Similarly, θ_j is the threshold value of the j^{th} hidden node and w_{ji} is the connection weight between the i^{th} input node and the j^{th} hidden node, x_i is the i^{th} input variable and f_{sig} is the sigmoid transfer function.

Table 5.4 Weights and threshold levels for the optimal ANN model.

Input layer nodes	**Weight from node *i* in the input layer to node *j* in the hidden layer (w_{ji})**			
	***j* = 9**	***j* = 10**	***j* = 11**	***j* = 12**
***i* = 1**	–3.128	–2.291	0.082	1.486
***i* = 2**	–5.257	–2.225	1.678	–0.743
***i* = 3**	1.216	–3.206	–0.014	1.482
***i* = 4**	–0.973	7.350	–1.869	–3.939
***i* = 5**	–2.230	–2.218	2.932	–3.939
***i* = 6**	–2.481	1.760	1.961	–1.450
***i* = 7**	13.630	–12.704	–1.286	–4.115
***i* = 8**	0.419	3.431	–0.888	–0.164
Hidden layer threshold $(\boldsymbol{\theta_j})$	–7.963	0.366	–0.908	1.055
Hidden layer nodes	**Weight from node *j* in the hidden layer to node *k* in the output layer (w_{kj})**			
	***k* = 13**			
***j* = 9**	–2.113			
***j* = 10**	–2.307			
***j* = 11**	–3.725			
***j* = 12**	–3.163			
Output layer threshold $(\boldsymbol{\theta_k})$	2.269			

Equation 5.2 can be further simplified as follows:

$$y_{k=13} = \frac{1}{1+e^{-\left[\theta_k + \sum_{j=9}^{12}\left(w_{kj}\ T_j\right)\right]}} \tag{5.3}$$

and,

$$T_{j=9,\ldots,12} = \frac{1}{1+e^{-\left[\theta_j + \sum_{i=1}^{8}\left(w_{ji}\ x_i\right)\right]}} \tag{5.4}$$

The variables x_1, x_2, x_3 and x_4 represent the soil types Sand–Clay, Clay–Silt, Sand–None and Sand–Gravel, respectively. These 4 input variables use the binary

representation, where the units 1 and 0 are simply used to indicate their presence or absence, respectively. For instance, when the numerical equation (Equation 5.2) is used for the soil type Sand–Clay, the following is applied: x_1= 1, x_2 = 0, x_3 = 0 and x_4 = 0. The remaining input variables, given by x_5, x_6, x_7 and x_8 represent the average depth, *D* (m), the initial of number roller passes, the initial DCP count (blows/300 mm) and the final number of roller passes, respectively.

However, it is noted that the input and output variables are required to be scaled down before using the above equations, as mentioned earlier. Therefore, the input variables are scaled between 0.1 and 0.9, by means of Equation 5.1, according to the data extremes incorporated in the training set (Table 5.2), and scaled values are substituted into Equations 5.3 and 5.4. In addition, the connection weights (w_{ji} and w_{kj}), as well as the threshold levels (θ_j and θ_k), are substituted into Equations 5.3 and 5.4 using the values given in Table 5.4. As a consequence, the mathematical relationship for the optimal ANN model incorporating 4 hidden nodes is simplified as follows:

$$Final\ DCP = \frac{102.5}{1+ exp\ (2.113\ T_9 + 2.307\ T_{10} + 3.725\ T_{11} + 3.163\ T_{12} - 2.269)} - 8.25 \qquad (5.5)$$

Where,

$$T_9 = [1 + exp(3.128x_1 + 5.257x_2 - 1.216x_3 + 0.973x_4 + 0.99x_5 + 0.04x_6 - 0.177x_7 - 0.006x_8 + 7.424)]^{-1}$$

$$T_{10} = [1 + \exp(2.291x_1 + 2.225x_2 + 3.206x_3 - 7.35x_4 + 0.985x_5 - 0.028x_6 + 0.165x_7 - 0.048x_8 + 0.059)]^{-1}$$

$$T_{11} = [1 + exp(-0.082x_1 - 1.678x_2 + 0.014x_3 + 1.869x_4 - 1.302x_5 - 0.031x_6 + 0.017x_7 + 0.012x_8 + 0.757)]^{-1}$$

$$T_{12} = [1 + exp(-1.486x_1 + 0.743x_2 - 1.482x_3 + 0.301x_4 + 1.749x_5 + 0.023x_6 + 0.053x_7 + 0.002x_8 - 0.517)]^{-1}$$

5.3.5 Sensitivity Analysis – Selection of Important Input Parameters

The relative importance of the factors that are significant to ground improvement predictions is identified by carrying out a sensitivity analysis of the selected optimal network. Garson's algorithm (Garson, 1991) is used in this study, which partitions the network's connection weights to determine the relative importance of each input variable. This method has been used by many researchers (Shahin et al., 2002; Pooya Nejad et al., 2009). The sensitivity analysis is repeated 4 times with the connection weights obtained from the optimal ANN model trained with 4 different initial random weight configurations. The average of the relative importance is adopted to rank the input variables and the results are summarised in Table 5.5. As one would expect, the input variables of soil type and initial DCP count are found to be the most important. The relative importance of the input variables appears to be highly sensitive to the initial starting position in the weight space, however the ranks of the input variables are found to be consistent with each trial.

Table 5.5 Sensitivity analysis of the relative importance of ANN input variables.

Input variable	Relative importance (%)				Average	Rank
	Trial 1	Trial 2	Trial 3	Trial 4		
Soil type	35.55	47.38	35.40	31.90	37.56	1
Average depth, *D*	17.52	11.81	18.73	17.86	16.48	3
Initial no. of roller passes	10.59	9.71	9.89	11.33	10.38	4
Initial DCP count	31.16	24.09	31.13	25.96	28.09	2
Final no. of roller passes	5.17	7.01	4.85	12.94	7.49	5

5.4 SUMMARY AND CONCLUSIONS

The work presented in this paper investigates the effectiveness of rolling dynamic compaction (RDC) on different soil types and seeks to establish a predictive tool by means of the often applied artificial intelligence technique, artificial neural networks

(ANNs). The ANN models incorporate a database of ground density data involving dynamic cone penetrometer (DCP) test results associated with RDC using the 4-sided, 8 tonne 'impact roller'. ANNs in the form of multi-layer perceptrons are trained with the back-propagation algorithm, where the model input variables are: soil type, average depth, *D* (m), initial number of roller passes, initial DCP count (blows/300 mm) and the final number of roller passes, with the sole output being the final DCP count (blows/300 mm) at depth *D* after compaction. It is found that the selected optimal model, with a single hidden layer incorporating 4 nodes, is capable of effectively capturing the density development with respect to the number of impact roller passes and the associated subsurface conditions. The resulting optimal ANN model demonstrates very good accuracy, with a coefficient of correlation (*R*) of 0.79, root mean square error (*RMSE*) of 7.54 and mean absolute error (*MAE*) of 5.59, when validated against a set of unseen data. In addition, a parametric study is carried out to assess the generalisation ability and robustness of the optimal model, where the results emphasise that the model's responses agree well with the expected physical relationships among the parameters. Therefore, the model is recommended as a reliable tool to predict ground improvement as a result of RDC.

A sensitivity analysis is also carried out where the relative importance of the parameters affecting ground improvement is investigated. It is identified that the soil type and the initial DCP count (blows/300 mm) are the dominant parameters. Subsequently, based on the optimal model characteristics, a simplified numerical equation that defines the functional form of the relationship between the model inputs and output is formulated to assist with hand calculations in practice.

Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP120101761). The authors wish to acknowledge Mr. Stuart Bowes at Broons Hire (SA) Pty Ltd for his kind assistance and continuing support, especially in providing access to the in situ test results upon which the numerical models are based. The authors are also grateful to Mr. Brendan Scott for his contribution to this work.

INTENTIONALLY BLANK

Chapter 6.

Predicting the Effectiveness of Rolling Dynamic Compaction Using Genetic Programming and Cone Penetration Test Data

INTENTIONALLY BLANK

Statement of Authorship

Title of Paper	A Genetic Programming Approach for Predicting the Effectiveness of Rolling Dynamic Compaction Using Cone Penetration Test Data.
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). A Genetic Programming Approach for Predicting the Effectiveness of Rolling Dynamic Compaction Using Cone Penetration Test Data. ICE-Ground Improvement, Under Review.

Principal Author

Name of Principal Author (Candidate)	R. A. T. M. Ranasinghe		
Contribution to the Paper	Data collection and pre-processing, Model development based on Genetic Programming (GP), Analysis of results and interpretation, and Preparation of the manuscript.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	20/02/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

vii. the candidate's stated contribution to the publication is accurate (as detailed above);

viii. permission is granted for the candidate in include the publication in the thesis; and

ix. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mark B. Jaksa		
Contribution to the Paper	Overall percentage – 10% Initial concept, research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Yien Lik Kuo		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Fereydoon Pooya Nejad		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

ABSTRACT

Rolling dynamic compaction (RDC) is a soil compaction method which involves impacting the ground with a non-circular roller. This technique is currently in widespread use internationally and has proven to be suitable for many compaction applications with improved capabilities over traditional compaction equipment. However, there is nevertheless a lack of knowledge for a *priori* estimation of the effectiveness of RDC on different soil profiles. To this end, the aim of this paper is to develop a reliable predictive tool based on a machine learning approach, linear genetic programming (LGP). The models are developed from a database of cone penetration test (CPT) based case histories. It is shown that the developed LGP-based correlations yield accurate predictions for unseen data and, in addition, the results of a parametric study demonstrate its generalisation capabilities. Furthermore, the selected optimal LGP-based model is found to yield superior performance when compared with an artificial neural network (ANN) model recently developed by the authors. It is concluded that the LGP-based model developed in this study is capable of providing reliable predictions of the effectiveness of RDC in various ground conditions.

Keywords: *Rolling dynamic compaction, linear genetic programming, cone penetration test*

6.1 INTRODUCTION

Rolling dynamic compaction (RDC) is now a well-established method of ground improvement where soil densification is achieved by means of high-energy impact blows. RDC employs heavy (6 to 12 tonnes), non-circular modules (3-, 4- and 5-sided), which rotate about their corners as it drawn forward towed behind a tractor (Avalle, 2004d). Thereby, a combination of potential and kinetic energy is derived from the impact mechanism, which provides a series of impact blows as the roller traverses the ground. Consequently, the soil beneath the surface is densified into a state of lower void ratio by expelling the pore air and fluid. However, the major benefit of RDC is its capability of influencing the ground to a greater depth, when compared to conventional static and vibratory compaction, which is more than 1 m beneath the ground surface and sometimes as deep as 3 m in some soils (Clegg and Berrangé, 1971; Clifford, 1976, 1978b; Avalle and Carter, 2005; Jaksa et al., 2012). In addition to the greater depth of compaction, RDC is capable of achieving the required density in thicker lifts, which are generally in excess of 500 mm, as compared with traditional layer thicknesses of 200 to 500 mm (Avalle, 2004d, 2006). Moreover, RDC can operate with larger particle sizes and the surface corrugations produced as a result of its operation provide a measure of interlocking between the adjacent soil layers, which helps to overcome lateral shearing effects. The economics of the use of RDC has also been found to be favourable because of its speed of operation, i.e. 9 to 12 km/h, which is substantially greater than the traditional vibratory roller which travels at approximately 4 km/h (Pinard, 1999). These inherent characteristics of RDC make it very effective for many civil, mining and agricultural applications, including pavement rehabilitation; in situ densification of existing fills, such as on brownfield sites and landfills; subgrade proof-rolling (Avalle, 2004a); construction of tailing dams at mine sites (Avalle, 2006); rock demolition in open cut mine waste tips (Scott and Jaksa, 2012); and improvements of existing water storages, channels and embankments (Avalle, 2004d).

To date, RDC has been studied experimentally through a number of field-based and case studies. However, until recently, as a result of work undertaken by the authors, which is discussed below, no rational means for obtaining a *priori* estimation of the

degree of densification or the extent of the influence depth by RDC in different ground conditions. Indeed, development of such a reliable theoretical model for prior estimation of the effectiveness of RDC is complex due to the heterogeneous nature of soil and of the various site-specific factors that can potentially affect the improvement process. Subsequently, the performance design and application of RDC currently relies heavily on the geotechnical engineer's experience and judgment. Field trials are often carried out on site to ascertain the operational parameters, especially the optimal number of impact roller passes required to achieve the required percentage of maximum dry density.

However, to address this problem, recent studies conducted by the authors in relation to RDC have proposed models by means of the artificial intelligence (AI) technique known as artificial neural networks (ANNs) (Ranasinghe et al., 2016c, 2016a). Two distinct ANN models have been developed based on cone penetration test (CPT) (Ranasinghe et al., 2016c), and dynamic cone penetration (DCP) test (Ranasinghe et al., 2016a), data and results obtained from previous ground improvement projects associated with the Broons 4-sided 'impact roller'. Except for a few restrictions imposed on the model utilisation, they have been shown to be successful in providing reliable predictions of the effectiveness of RDC in various ground conditions. Despite the fact that ANNs provide acceptable performance in many geotechnical engineering applications, they suffer from a few shortcomings. Essentially, ANNs require the network structure and the parameters to be recognised in advance, which usually entail the implementation of somewhat ad-hoc, trial-and-error methods. Moreover, a common criticism levelled at ANNs is their lack of transparency, in that they often fail to explain the underlying physical processes associated with the phenomenon under investigation; in this case, compaction.

This paper investigates the applicability of a relatively new, machine learning technique called genetic programming (GP), that is reported to overcome many of the shortcomings associated with ANNs and other conventional modelling approaches for the prediction of the effectiveness of RDC. The GP models are developed using a reliable dataset of CPT results that have also been utilised previously for ANN model

development by the authors (Ranasinghe et al., 2016c). Explicit formulae based on the optimal GP model are presented. In addition, a comparative study is conducted where the GP- and ANN-based models are compared in terms of a range of performance measures.

6.2 GENETIC PROGRAMMING (GP)

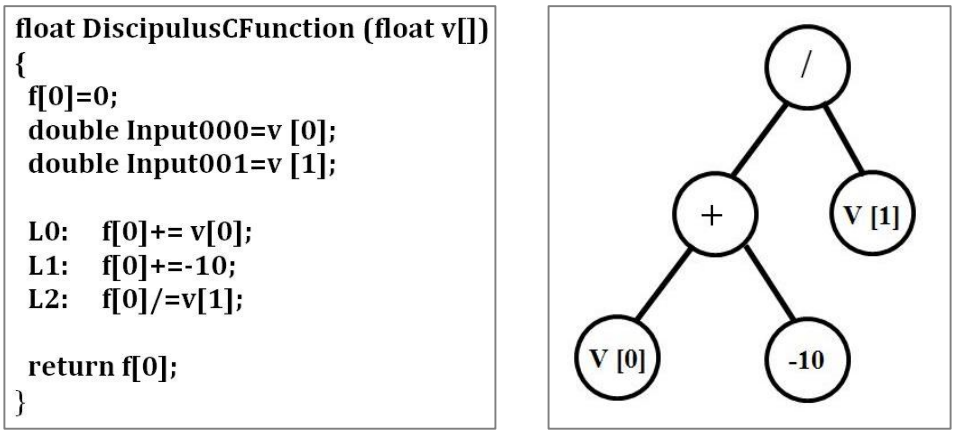
Genetic Programming (GP) (Koza, 1992) is one of a number of approaches based upon evolutionary algorithms (EAs) that mimic the concept of Darwin's evolution theory in relation to optimising a solution to a predefined problem. Similar to ANNs, GP is part of the AI class of modelling techniques, which can be considered as an alternative to conventional methods, such as for example, statistical and finite element modelling, because of its ability to approximate any linear/non-linear relationship among a set of observed input and output data in the absence of former knowledge on the underlying mechanisms of the system. However, GP is still considered a relatively new evolution-based optimisation technique in the field of geotechnical engineering. Nonetheless, several applications of GP can be found in the literature (Javadi et al., 2006; Johari et al., 2006; Heshmati et al., 2008; Taskiran, 2010; Alavi and Gandomi, 2012; Alavi et al., 2013).

GP was first introduced by Koza (1992) in the early 1990s and it is often considered as a generalisation/extension of genetic algorithms (GAs). However, GAs are essentially recognised as individuals represented by fixed-length binary strings (Holland, 1975) whilst, GP represents the individuals as computer programs of variable size and shape (Koza, 1992). The individuals in a GP population are hierarchically composed of a set of functions and terminals fitted to a particular problem domain. The function set may consist of arithmetic functions (+, –, ×, /), mathematical functions (*sin*, *cos*, *ln*), Boolean logic operators (AND, OR, NOT), logical expressions (IF or THEN), iterative functions (DO, CONTINUE, UNTIL) and/or other user-defined functions (Sette and Boullart, 2001). The terminal set typically comprises of input variables attached to the problem domain and pre-specified or randomly generated numeric constants.

The earliest Koza style of GP is now widely known as a tree-based GP (TGP), where the programs are represented as tree structures. In addition to the traditional TGP, there are several distinguished variants of GP, where programs are represented in different forms: i.e. linear GP and graph-based GP (Banzhaf et al., 1998; Poli et al., 2007). The emphasis of the present study is placed on the linear GP technique.

6.2.1 Linear Genetic Programming

Linear genetic programming (LGP) (Brameier and Banzhaf, 2007) is a subset of GP, where the individuals are represented in a linear format. The main distinguishing feature of LGP over TGP is that LGP evolves programs of an imperative language (e.g. C, C++ or Java) or machine language instead of the standard TGP expressions in a functional programming language (i.e. Lisp) (Brameier and Banzhaf, 2001; Brameier and Banzhaf, 2007). Moreover, the data flow of evolved programs in LGP has a more general register-based graphical representation at the functional level, compared to the rigidly determined tree representation of traditional TGP. A comparison of the typical program structures giving the same end result, produced by LGP and TGP, is presented in Figure 6.1.



(a) (b)

Figure 6.1 Comparison of the GP structures: (a) LGP (b) TGP [modified from Alavi et al. (2013)].

As described earlier, the LGP individuals are evolved either from an imperative programming language (e.g. C, C++ or Java) (Brameier and Banzhaf, 2001; Brameier and Banzhaf, 2007) or a direct machine language (Nordin, 1994). The latter variant is also known as automatic induction of machine code by genetic programming (AIMGP), where the evolved programs are stored as linear strings of native binary machine code (Nordin et al., 1999). In contrast, the individuals in AIMGP are directly executable by the processor (Francone and Deschaine, 2004). AIMGP is found to be more memory efficient and significantly faster than other GP variants because there is no need for an interpreter for the evaluation of individuals (Nordin et al., 1999; Alavi and Gandomi, 2012). As a consequence of these advantages, this study makes use of AIMGP.

6.2.2 LGP Evolutionary Algorithm

LGP performs a multi-directional simultaneous search for an optimal solution from a pool of many potential solutions, collectively known as a 'population'. The individuals in the population compete with each other, such that the fittest individuals survive and eventually evolve to do well in the given environment. In general, the basic steps of the LGP evolutionary algorithm (Brameier and Banzhaf, 2007) can be summarised as follows:

1. Initialising a population of randomly generated programs and evaluating their fitness;
2. Performing two fitness tournaments with randomly selected programs from the population and selecting the winning programs;
3. Making temporary copies of the two winning programs;
4. Transforming the two winning programs into offsprings subjected to genetic operations, i.e. crossover and mutation with certain probabilities;
5. Replacing the two tournaments losing programs with the temporary copies of the winning programs; and
6. Repeating Steps 2 to 5 until the termination or convergence criteria are satisfied.

6.3 LGP-BASED MODELLING APPROACH

The following section provides an overview of the data used in the modelling process, followed by a detailed assessment of the methodology adopted in developing the LGP-based models.

6.3.1 Database and Data Pre-processing

A comprehensive database containing the results of several field trials undertaken by Broons, as presented in the authors' recent work (Ranasinghe et al., 2016c), is again utilised in the present study. The database is comprised of in situ soil strength data in the form of cone penetration test (CPT) results with respect to a varying number of roller passes. CPT results are presented in terms of cone tip resistance (q_c) and sleeve friction (f_s) measurements. It is considered that the differences between the individual measurements obtained at essentially the same location prior to compaction (0 roller passes) and after compaction (10, 20 roller passes), effectively quantify the variation in soil strength and density resulting from RDC. In total, 1,977 data records are available after averaging the CPT values over 0.2 m depth intervals from 103 CPT soundings. Further details of the CPT data are given by Ranasinghe et al. (2016c).

This study takes into account the fundamental factors that influence ground density improvement by means of soil compaction when deciding the appropriate model inputs and outputs. The input variables used to develop the prediction correlations are: the depth of measurement (D), cone tip resistance (q_{ci}) and sleeve friction (f_{si}) prior to compaction and the number of roller passes (P), whilst the single output variable is cone tip resistance after compaction (q_{cf}). It is considered that these input variables, either individually or collectively, address the key factors that are most influential to the degree of soil compaction, including the inherent physical properties of the soil, such as initial density, moisture content and soil type, which are accounted for in the CPT data, and the amount of energy imparted to the ground, which is accounted for with the parameter P. The ranges of the input and output variables involved in model development are presented in Table 6.1.

Prior to LGP modelling, the available dataset is divided into a series of subsets. In order to conduct a fair comparison between the results obtained herein and those from the previous ANN model (Ranasinghe et al., 2016c), this study utilises the same data subsets employed in the previous ANN model development by the authors. In summary, the entire dataset has been divided into two sets: a modelling dataset (consisting of 1,755 records from 91 CPT soundings); and a verification dataset (consisting of 222 records from 12 CPT soundings). The modelling dataset is used to train and validate the LGP models and is adopted in the modelling phase. However, the verification dataset is not a part of the modelling phase in any capacity but is introduced to the selected optimal LGP model in order to further verify its capabilities.

Table 6.1 Input/output variable ranges used in model development.

Variables	Range
Input	
Depth of measurement, $\boldsymbol{D}$ (m)	0.2 – 4.0
Cone tip resistance prior to compaction, $\boldsymbol{q_{ci}}$ (MPa)	0.19 – 50.65
Sleeve friction prior to compaction, $\boldsymbol{f_{si}}$ (kPa)	1.67 – 473.86
No. of Roller Passes, $\boldsymbol{P}$	5 – 40
Output	
Cone tip resistance after compaction, $\boldsymbol{q_{cf}}$ (MPa)	0.17 – 50.36

For the LGP analysis, it is necessary to divide the modelling dataset into 3 subsets: training, testing and validation. The learning subset incorporates 80% of the entire dataset and consists of the training and testing subsets. The programs are genetically evolved and optimised to learn the input/output relationships with respect to the training subset, whilst the model's generalisation capability is evaluated periodically using the testing subset. Upon the completion of the LGP model calibration, the remaining 20% of the data, included in the validation subset, is presented to the optimal program as a set of unseen data to assess its performance. However, it is important to ensure that these 3 subsets are statistically consistent so that they effectively represent the same population, which is the ideal form of data division. It can be seen in Table 6.2, that the subsets used in this study effectively represent the same population as

evidenced by the similar values of mean, standard deviation, minimum, maximum and range.

Table 6.2 Statistical properties of the data used in the LGP model development.

Variable	Data subset	Statistical parameters				
		Mean	*SD	Minimum	Maximum	Range
Inputs						
Depth, $\boldsymbol{D}$ (m)	Training	1.95	1.11	0.20	4.00	3.80
	Testing	2.03	1.14	0.20	4.00	3.80
	Validation	2.03	1.14	0.20	4.00	3.80
Cone tip resistance prior to compaction, $\boldsymbol{q_{ci}}$ (MPa)	Training	9.33	8.23	0.19	50.65	50.46
	Testing	9.32	8.37	0.30	47.39	47.09
	Validation	9.39	7.83	0.32	47.94	47.63
Sleeve friction prior to compaction, $\boldsymbol{f_{si}}$ (kPa)	Training	103.36	71.76	1.67	473.86	472.19
	Testing	99.23	71.29	8.70	441.04	432.34
	Validation	103.54	70.37	7.08	470.29	463.21
No. of Roller Passes, $\boldsymbol{P}$	Training	26.59	9.94	5.00	40.00	35.00
	Testing	27.21	9.64	5.00	40.00	35.00
	Validation	26.62	10.30	5.00	40.00	35.00
Output						
Cone tip resistance after compaction, $\boldsymbol{q_{cf}}$ (MPa)	Training	10.42	8.30	0.17	50.36	50.20
	Testing	10.50	8.66	0.29	45.12	44.83
	Validation	10.43	8.03	0.39	46.20	45.81

The entire dataset is rescaled using the min-max normalisation method. Although such data transformation is not strictly necessary, it is usually recommended as it often improves the effectiveness and the performance of the algorithm (Alavi and Gandomi, 2012). Thus, prior to model development, both the input and output variables are rescaled into the range of 0 to 1 using the following Equation:

$$x_{scaled} = \frac{(x_{unscaled} - x_{min})}{(x_{max} - x_{min})} \quad (6.1)$$

Where x_{min} and x_{max} are, respectively, the minimum and maximum values of the *x* variable with respect to the training dataset, as given in Table 6.2.

6.3.2 Model Development Using LGP

In this study, the commercially available software *Discipulus* version 5.2 (Francone, 2010) is used for the LGP-based model development. It is a supervised learning system, which operates on the basis of the AIMGP platform. It can be considered to be an efficient modelling tool for complex problems, but requires careful consideration in terms of parameter selection.

It has been identified from the literature, that the selection of control parameters affects the model's generalisation ability. Therefore, different parameter settings, in terms of population size, crossover rate and mutation rate, are investigated in this study. Most of the other minor parameters are maintained at the values recommended from similar applications (Baykasoğlu et al., 2008; Heshmati et al., 2008; Gandomi et al., 2010b; Alavi and Gandomi, 2012). Furthermore, the preliminary modelling observations are used when selecting the parameters, as listed in Table 6.3.

In this study, a relatively large number of LGP projects are carried out with the different parameter combinations discussed above. Furthermore, each parameter combination is tested for 5 replications to permit different random initial conditions. It is worth mentioning that a LGP project consists of a successively generated series of runs which may begin with short runs and be permitted to increase the length of the runs as the project continues. This study makes use of MSE as the fitness function, as discussed above, and thus the evolved programs are monitored for minimum error. Each of the LGP projects is given a reasonable time (ranging from few minutes to 20+ hours) to evolve and the project is terminated when no further improvement in model performance is likely to occur.

Table 6.3 Parameter settings for the LGP algorithm.

Parameter	Settings
Function set	+, –, ×, /, Absolute, Square Root, Trigonometric (*sin*, *cos*), Exponential
Population size	500, 1,000, 2,000, 5,000, 7,500, 10,000
Number of demes	10, 20
Initial program size	80 bytes
Maximum program size	512 bytes
Mutation frequency	50%, 90%
Block mutation frequency	40%
Instruction mutation frequency	30%
Instruction data mutation frequency	30%
Crossover frequency	50%, 95%
Homologous crossover frequency	95%

The resulting LGP models are evaluated using several performance measures with respect to each of the 3 data subsets and compared. The criteria used to evaluate the performance of the evolved program models include the coefficient of correlation (*R*), root mean square error (*RMSE*) and mean absolute error (*MAE*), which are determined as follows:

$$R = \frac{\sum_{i=1}^{n}(d_i-\bar{d})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(d_i-\bar{d})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{6.2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(d_i-y_i)^2}{n}} \tag{6.3}$$

$$MAE = \frac{\sum_{i=1}^{n}|d_i - y_i|}{n} \quad (6.4)$$

Where d_i and y_i are respectively the actual and model predicted output values for the i^{th} sample; $\bar{d}$ and $\bar{y}$ are the averages of the actual and model predicted output values, respectively, over n number of samples.

6.4 OPTIMAL MODEL RESULTS

In this section, the details of the optimal LGP-based model are presented along with the performance analysis results. The robustness of the optimal model is investigated using a parametric study, followed by the explicit formulation of an LGP-based numerical equation. In addition, the selected optimal LGP model for predicting the cone tip resistance after compaction (MPa), is presented in computing code in the C language in Appendix A.

6.4.1 Performance Analysis

In selecting the optimal model, the program models generated from the LGP projects are compared based on the performance measures in terms of *R*, *RMSE* and *MAE*, as discussed above. The model yielding the lowest error and highest *R* with respect to the validation data subset is considered to be optimal and Table 6.4 presents the statistical performance of the selected optimal LGP model.

Table 6.4 Performance statistics of the optimal LGP model.

Data subset	**Performance criteria**		
	R	***RMSE* (MPa)**	***MAE* (MPa)**
Training	0.87	4.05	2.72
Testing	0.88	4.08	2.73
Validation	0.87	4.03	2.71

It is evident, that the selected optimal model is able to predict accurately the target values as evidenced by the high values of R and low prediction errors indicated by *RMSE* and *MAE*. According to Smith (1986), when $R > 0.8$ and the errors are relatively small, there exists a strong correlation between the measured and predicted values. Thus, it can be considered that the optimal LGP model yields reliable estimates of the ground's cone tip resistance due to RDC.

Figure 6.2 compares the measured and predicted q_{cf} values with respect to the testing and validation set data. It is apparent that the model has learnt the input/output mapping very effectively, demonstrated by the very good estimates when presented with a new set of unseen data. The optimal model predictions are scattered within an envelope of 0.5 to 2 times the measured values, which can be considered as a reasonable band of accuracy for the ground improvement predictions given the uncertainties involved.

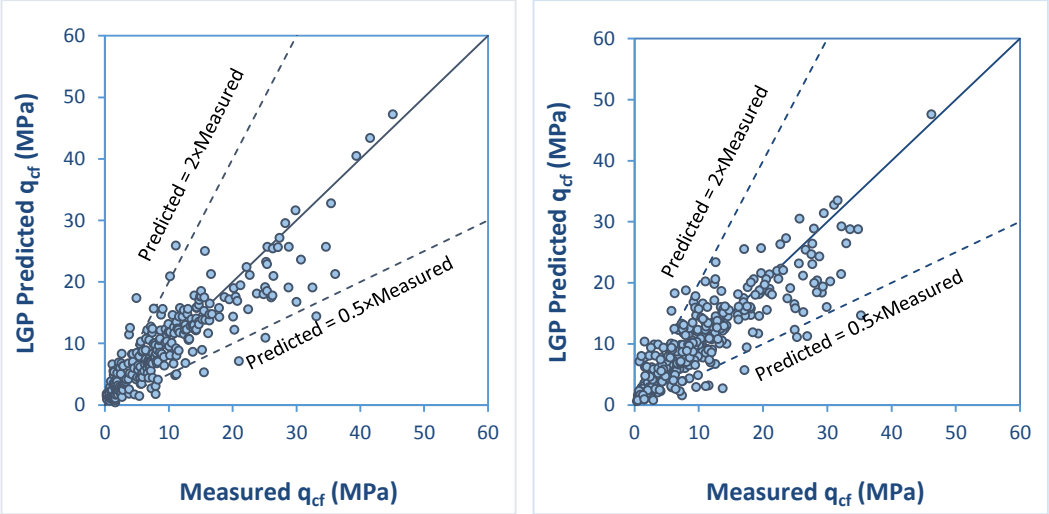


Figure 6.2 Measured versus predicted q_{cf} for the optimum LGP model with respect to: (a) testing; and (b) validation data sets.

6.4.2 Parametric Study

In order to assess the generalisation ability and the robustness of the selected optimal LGP-based model, a parametric study is conducted, which evaluates the sensitivity of the model output; i.e. q_{cf} to the variations in the input parameters D, q_{ci}, f_{si} and P. This involves investigating the model's response to a hypothetical input dataset, where the input variables are varied one at a time, whilst all other input variables remain constant at a pre-defined value. It is important to ensure that the variables fluctuate only within the range fixed by the training dataset since the model performs best as an interpolation predictor rather than by extrapolation beyond the calibrated range. In this study, the output, q_{cf} is examined while the input variables adopt the following values: q_{ci} = 2, 5, 8, 10, 15, 20 MPa; f_{si} = 50, 100, 150, 200 kPa and the number of roller passes, P = 10, 20, 30, 40. Figures 6.3 to 6.8 present the optimal model predictions of q_{cf} with respect to the variations in the input variables.

The results of the parametric study indicate that the soil strength continuously improves with increasing numbers of roller passes at a given location. For instance, in Figure 6.3(a), it can be observed that q_{cf} consistently rises when the number of roller passes increases systematically from 10 to 40 passes while q_{ci} and f_{si} remain constant at the pre-defined values of 2 MPa and 50 kPa, respectively. A similar trend is also observed when f_{si} is varied between 50 to 200 kPa, while q_{ci} remains constant at 2 MPa [Figures 6.3(b)–(d)]. Furthermore, the effect of varying q_{ci} is also investigated as illustrated in Figures 6.3 to 6.6. It is evident, when q_{ci} increases from 2 to 10 MPa, at a given depth, q_{cf} always improves to a value higher than q_{ci}, consistently indicating some level of ground improvement. However, from Figures 6.3 to 6.6, it is also evident that q_{cf} is less sensitive to variations in f_{si}, as indicated by the modest changes in q_{cf} when f_{si} increases from 50 to 200 kPa, whilst the other variables remain constant. Nevertheless, it can be concluded that the model predictions are reliable in the sense that the model replicates the expected underlying physical behaviour of RDC compaction and can be considered to be robust.

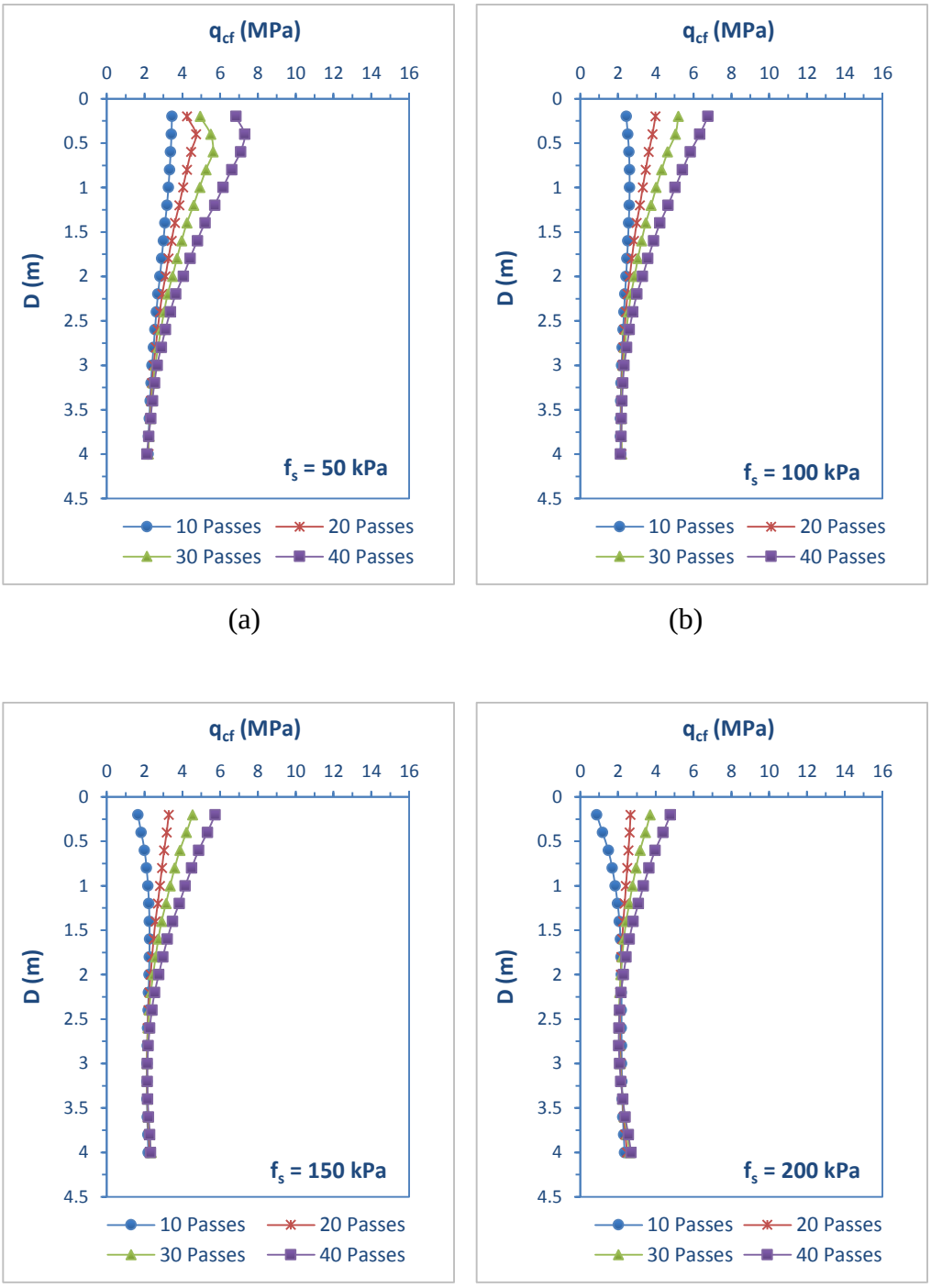
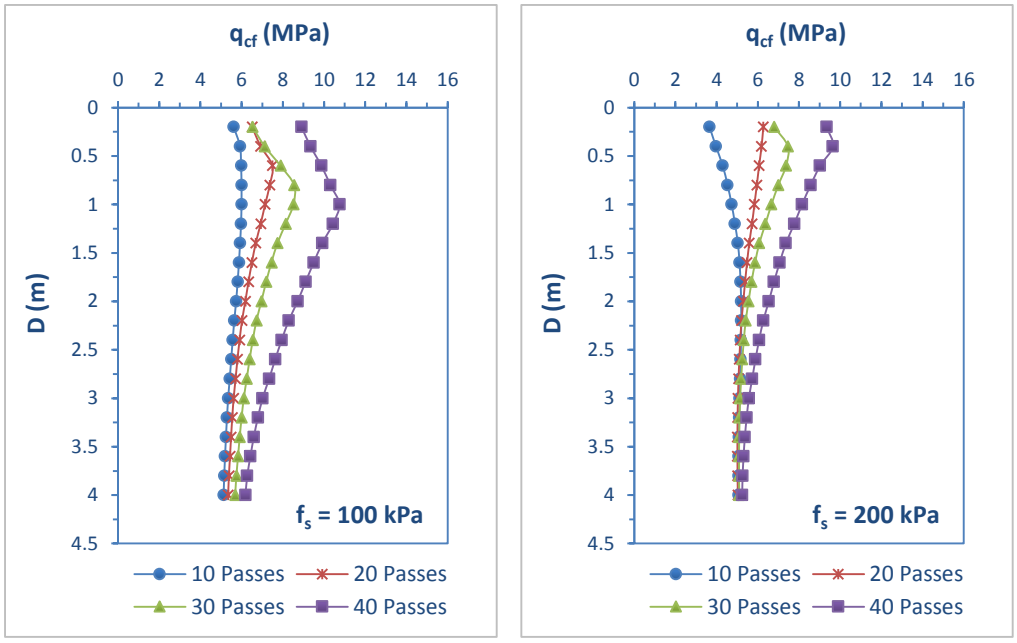
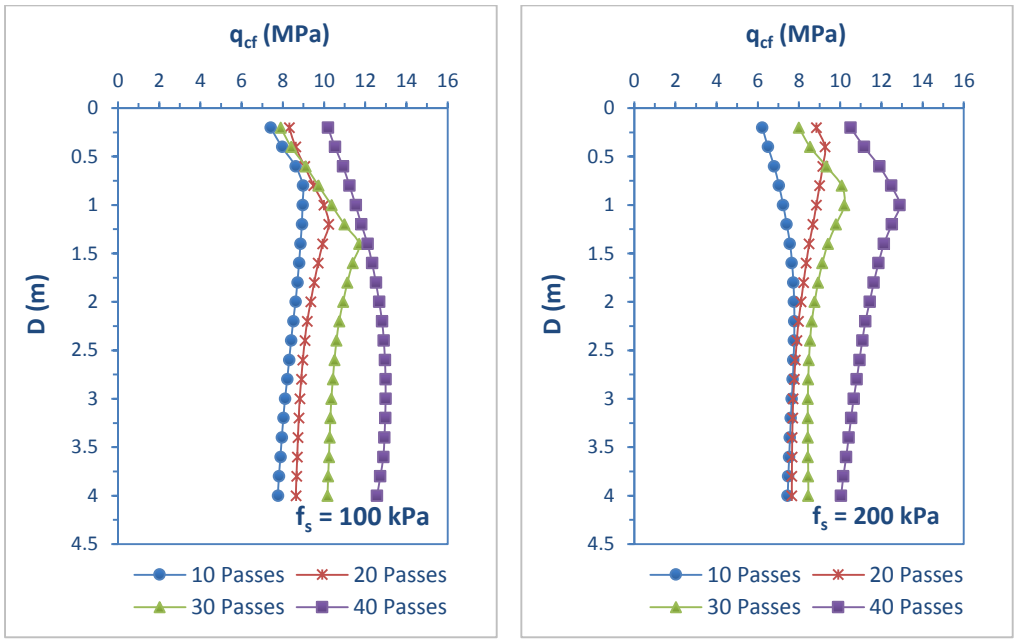


Figure 6.3 Variation of q_{cf} with different number of roller passes at q_{ci} = 2 MPa and: (a) f_{si} = 50 kPa; (b) f_{si} = 100 kPa; (c) f_{si} = 150 kPa; and (d) f_{si} = 200 kPa.



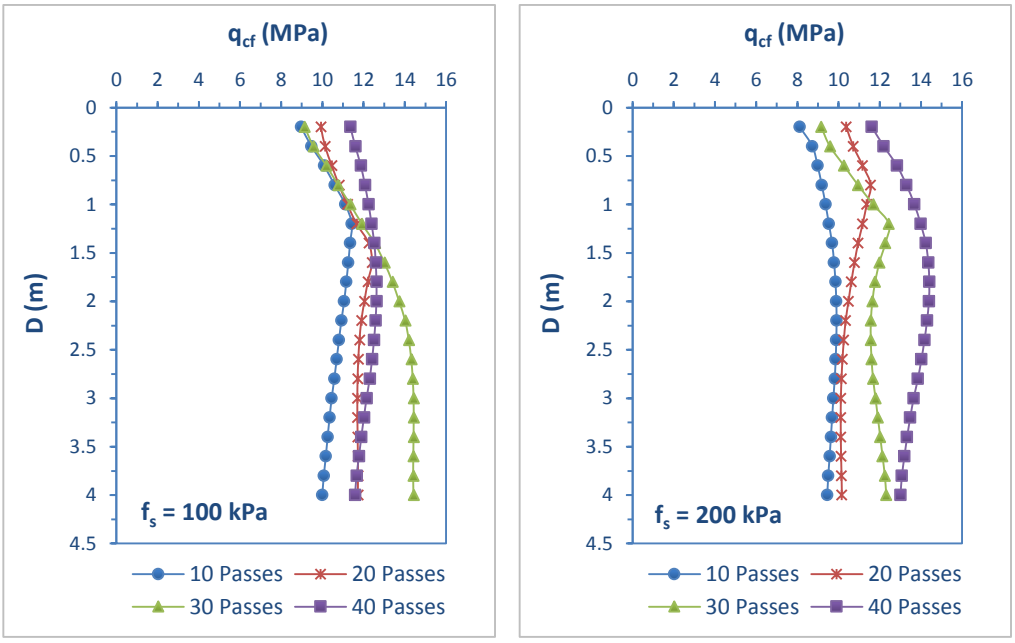
(a) (b)

Figure 6.4 Variation of q_{cf} with different number of roller passes at q_{ci} = 5 MPa and: (a) f_{si} = 100 kPa; and (b) f_{si} = 200 kPa.



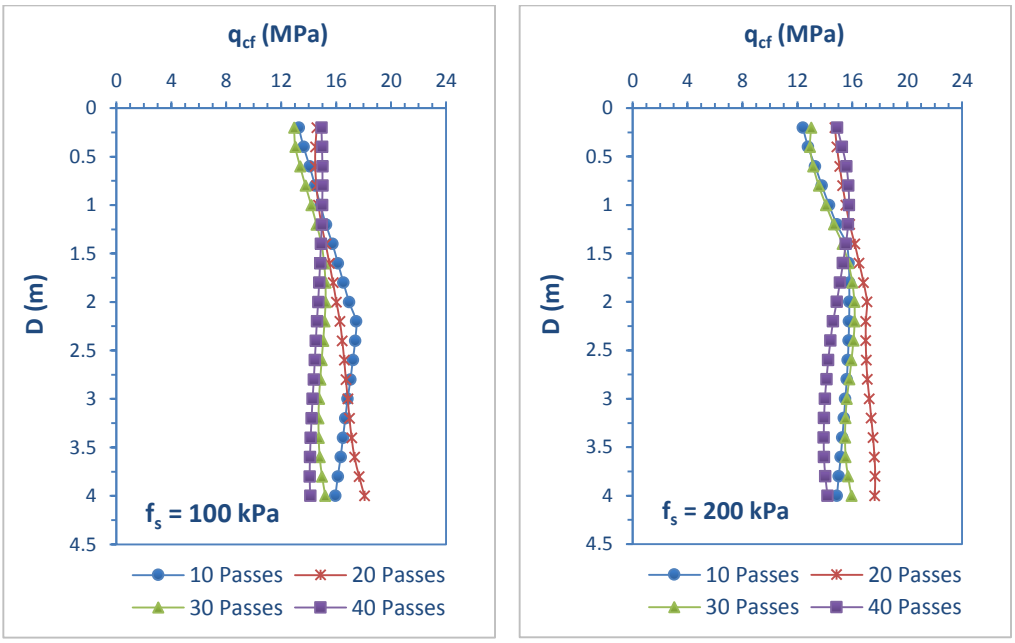
(a) (b)

Figure 6.5 Variation of q_{cf} with different number of roller passes at q_{ci} = 8 MPa and (a) f_{si} = 100 kPa; and (b) f_{si} = 200 kPa.



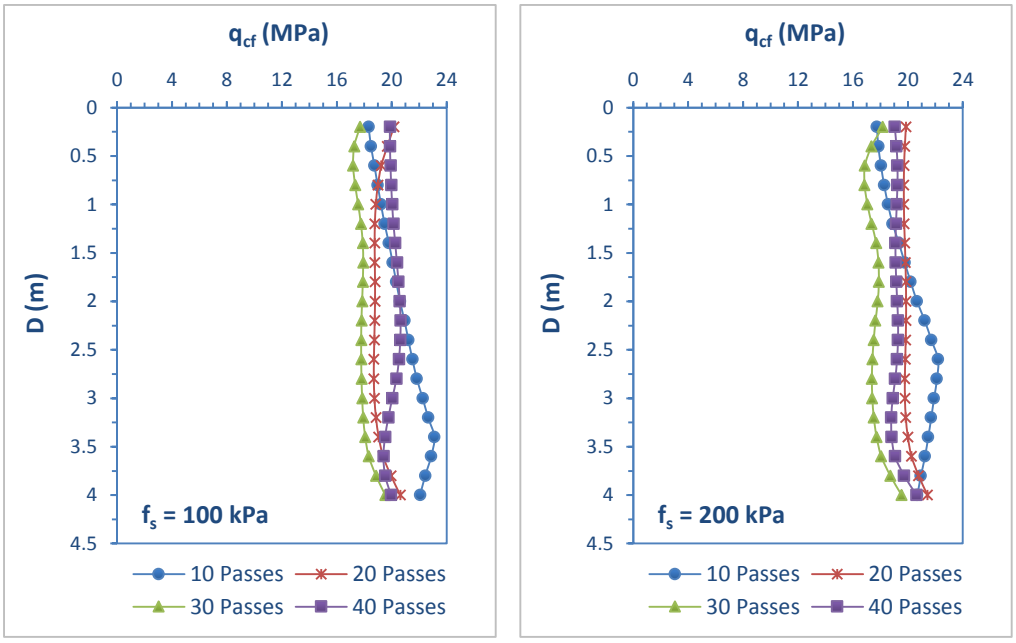
(a) (b)

Figure 6.6 Variation of q_{cf} with different number of roller passes at q_{ci} = 10 MPa and (a) f_{si} = 100 kPa; and (b) f_{si} = 200 kPa.



(a) (b)

Figure 6.7 Variation of q_{cf} with different number of roller passes at q_{ci} = 15 MPa and: (a) f_{si} = 100 kPa; and (b) f_{si} = 200 kPa.



(a) (b)

Figure 6.8 Variation of q_{cf} with different number of roller passes at q_{ci} = 20 MPa and: (a) f_{si} = 100 kPa; and (b) f_{si} = 200 kPa.

As can be seen from Figures 6.3 to 6.6, the trends and relationships between the q_{cf} predictions and the variations of the other input parameters are appropriate and conform to the expected behaviour. However, in contrast, the q_{cf} predictions are less satisfactory when the model is exposed to relatively high q_{ci} values. As shown in Figures 6.7 and 6.8, the predicted q_{cf} values are negatively correlated with the number of roller passes when q_{ci} is either 15 or 20 MPa. This is likely due to the poor representation of high q_{ci} values in the database, since RDC in most projects is usually applied to initially loose ground. Thus, the majority of the samples in the dataset used for model calibration represent the lower range of q_{ci}. Therefore, the selected LGP model is suggested to be well suited for the cases with low q_{ci} values, i.e. below 10 MPa, where the initial ground condition is loose to medium-dense.

6.4.3 Sensitivity Analysis

In this study, a sensitivity analysis is also carried out to investigate the contribution of each input variable on the final model predictions. *Discipulus* is capable of examining

the frequency of each input variable appearing in the 30 best selected programs (Francone, 2010). In this LGP-based modelling process, the frequency obtained for all the input variables; i.e. D, q_{ci}, f_{si} and P, is equal to 1, which indicates that these variables have been appeared in all of the 30 best programs evolved using LGP. Nonetheless, the average and maximum effect of removing the corresponding variable from the 30 best programs is calculated relative to each input variable and the results are presented in Figure 6.9. As can be observed, all the input variables have an almost identical effect on the output if they are removed and therefore, it is considered that all the selected input variables are highly significant with respect to the q_{cf} predictions. However, as indicated by the average impact measure, q_{ci} has the greatest effect when compared to the other input parameters.

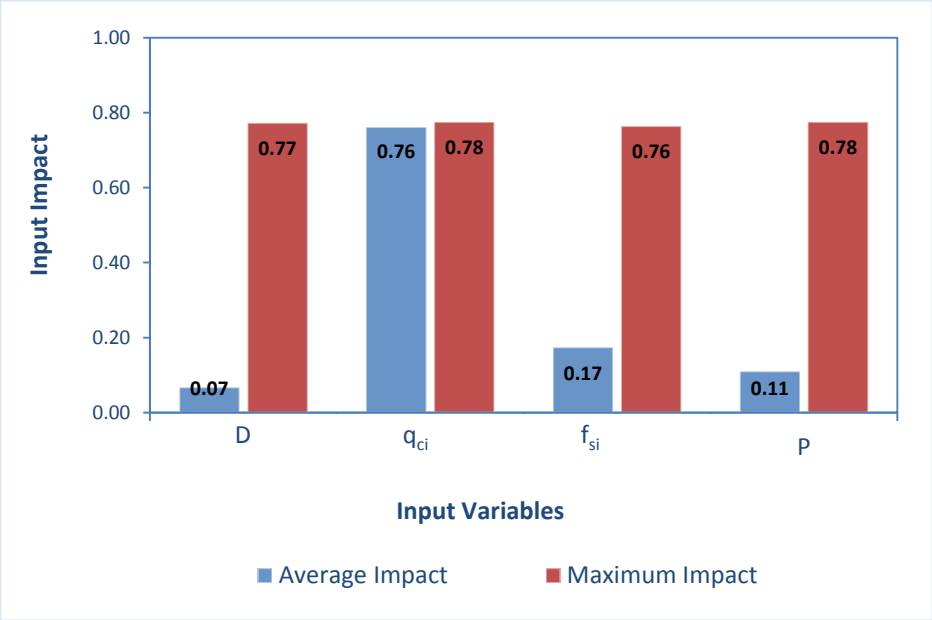


Figure 6.9 Impact of the input variables on optimal model predictions.

6.4.4 LGP-based Numerical Equation

Once the model has been developed, *Discipulus* (Francone, 2010) provides the facility to translate the evolved optimal LGP program into a C code computer program as presented in Appendix A. In order to facilitate its use in practice, the obtained code is converted into a series of tractable equations as follows:

$$q_{cf} = 50.19X(1 - Y) + 1.004q_{ci} - 0.371 \quad (6.5)$$

Where:

$$X = -0.085 \times (0.02q_{ci} + 0.0006f_{si} + C - 0.0924D + 2\sin(0.0032(q_{ci} - 0.19)) - 1.296sin^2(-A) - 0.716)$$

$$Y = C - 0.04q_{ci} + 0.008 + [1 - 0.592A]sin^2(-A) + \sin(A) + 0.0012A(f_{si} - 2) + [Z + 0.478\cos(Z)][0.251 - 0.092D][2^n - 1]$$

and

$$A = [0.02q_{ci} - 0.092D + 0.247] \times [0.004f_{si} - 0.135P - 0.263D - 1.182\sin(0.0032q_{ci} - 0.0006) + 3.121]$$

$$B = [A(0.311D - 0.063) + 1.296]sin^2(-A) - 2sin(0.0032q_{ci} - 0.0006) + (0.02q_{ci} - 0.004)[1.578D - 1.318] + 0.092D - 0.251 - [A(0.311D - 0.063) + 0.296](0.002f_{si} - 0.004)$$

$$C = \left|(0.029P - 0.143).\, sin\left(\frac{B^2 + (0.02q_{ci} - 0.004)}{0.87}\right) + (0.029P + 1.857)(0.02q_{ci} - 0.004) - 0.263D - 0.002f_{si} + 0.057\right|$$

$$Z = |0.002f_{si} - 0.068P - 1.182\sin[0.0032q_{ci} - 0.0006] + 0.738|$$

$$n = [0.02q_{ci} - 0.092D + 0.247]^2 \times [3.092 - 0.129P + 0.004f_{si} - 0.263D - 1.182\sin(0.0032q_{ci} - 0.0006)] + 2\sin(0.0032q_{ci} - 0.0006) - 0.092D + 0.02q_{ci} + 0.247$$

Where D is the depth of measurement (m); q_{ci} and f_{si} are the initial cone tip resistance (MPa) and sleeve friction (kPa) prior to compaction, respectively; P is the number of roller passes and q_{cf} is the cone tip resistance after compaction (MPa), as detailed previously.

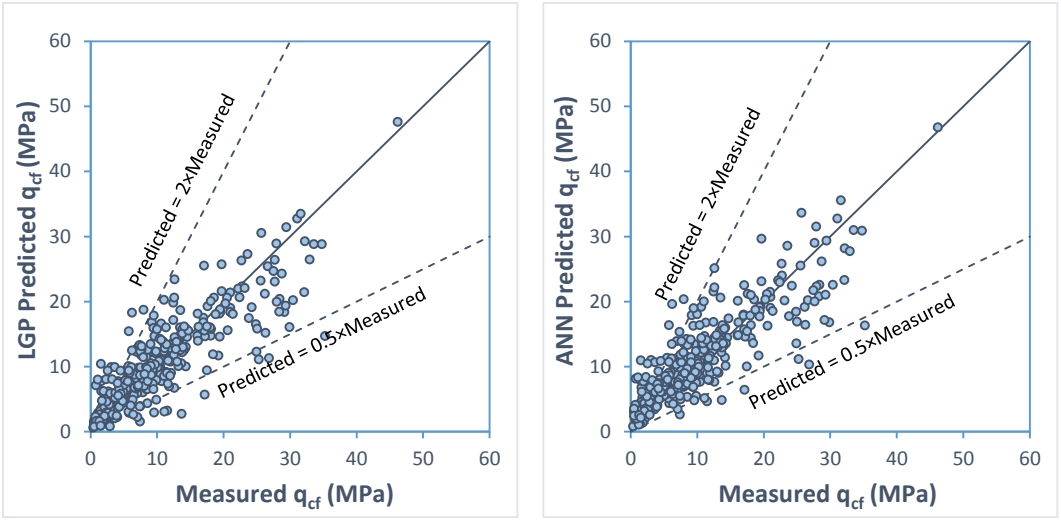
6.5 COMPARATIVE STUDY

In this section, the results obtained from the LGP simulations are compared with those obtained from the ANN-based model recently developed by the authors (Ranasinghe et al., 2016c) using several measures. Firstly, the performances of both models are again evaluated using *R*, *RMSE* and *MAE* and the results are presented in Table 6.5. It is observed that, overall, both models exhibit similar performance. Therefore, it can be considered that both models are capable of predicting the target values to a high degree of accuracy, as indicated by the strong correlation coefficients; i.e. $R > 0.8$ (Smith, 1986), together with the relatively low error values with respect to each of the datasets. However, it is evident that the LGP model yields slightly better *R* values and lower error values compared to the ANN model and thus, the LGP-based model marginally outperforms the ANN model.

Table 6.5 Comparison of the performance statistics of optimal LGP- and ANN-based models.

Data subset	**Performance criteria**					
	R		***RMSE* (MPa)**		***MAE* (MPa)**	
	LGP	**ANN**	**LGP**	**ANN**	**LGP**	**ANN**
Training	0.87	0.87	4.05	4.19	2.72	2.89
Testing	0.88	0.87	4.08	4.33	2.73	3.03
Validation	0.87	0.86	4.03	4.16	2.71	2.93

Figure 6.10 compares the measured and predicted q_{cf} values of optimal LGP and ANN model with respect to the validation set data. It can be clearly seen that there is a minimal scatter given the measured and predicted values are in relatively close agreement. Thus, it is evident that both the models perform very favourably.



(a) (b)

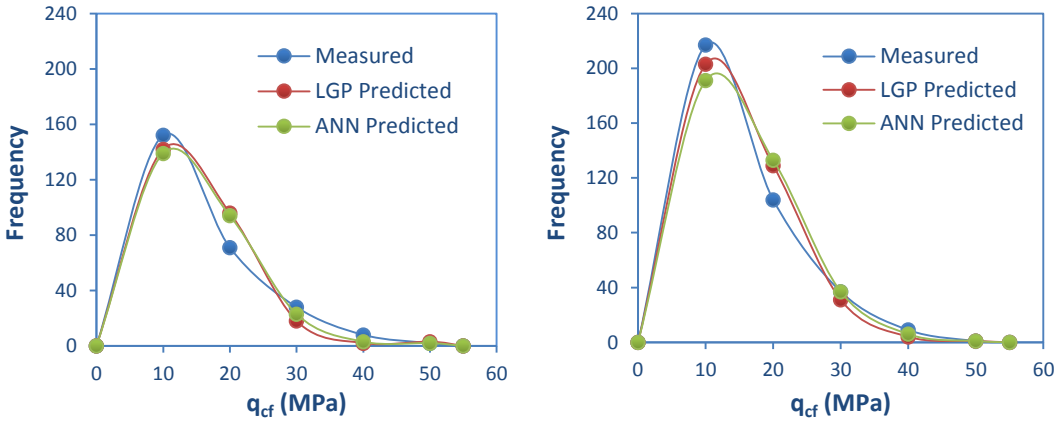
Figure 6.10 Measured versus predicted q_{cf} with respect to validation data set for: (a) LGP model predictions; and (b) ANN model predictions.

In order to assess more extensively the relative performance of both models additional measures, as suggested in literature, are examined. The validation criteria and the results obtained from both of the model predictions are presented in Table 6.6. It is evident from these results that all performance measures confirm the above conclusion that the LGP and ANN models provide accurate predictive capability and that the LGP model slightly outperforms the ANN model.

As a further measure, the distribution properties of the optimal LGP and ANN model predictions are compared with measured data with respect to the testing and validation data sets. The distribution properties considered here include the mean, standard deviation (SD), coefficient of variation (CV), skewness, minimum, maximum and range. The statistics are summarised in the Table 6.7 and represented graphically in Figure 6.11. It is evident that the distributions of the LGP and ANN model predictions are very similar to the measured data distribution and, in essence, the distribution properties are in very close agreement. However, the LGP model again yields slightly superior distribution properties and more closely matches the measured data distribution.

Table 6.6 Additional performance measures of the LGP and ANN models for the validation data set.

Item	Formula	Reference	Condition	LGP	ANN
1	$k = \frac{\sum_{i=1}^{n}(d_i \times y_i)}{\sum_{i=1}^{n} {d_i}^2}$	Golbraikh and Tropsha (2002)	$0.85 < k < 1.15$	0.91	0.95
2	$k' = \frac{\sum_{i=1}^{n}(d_i \times y_i)}{\sum_{i=1}^{n} {y_i}^2}$	Golbraikh and Tropsha (2002)	$0.85 < k' < 1.15$	0.99	0.95
3	${R_0}^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - {d_i}^0\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y_i})^2}$ Where; ${d_i}^0 = k \times y_i$	Roy and Roy (2008)	Should be close to 1	0.98	0.99
4	${R_0}'^2 = 1 - \frac{\sum_{i=1}^{n}(d_i - {y_i}^0)^2}{\sum_{i=1}^{n}\left(d_i - \bar{d_i}\right)^2}$ Where; ${y_i}^0 = k' \times d_i$	Roy and Roy (2008)	Should be close to 1	1.00	0.99
5	$m = \frac{R^2 - {R_0}^2}{R^2}$	Golbraikh and Tropsha (2002)	$m < 0.1$	–0.30	–0.34
6	$n = \frac{R^2 - {R_0}'^2}{R^2}$	Golbraikh and Tropsha (2002)	$n < 0.1$	–0.34	–0.34



(a) (b)

Figure 6.11 Data distribution of predicted q_{cf} for the LGP and ANN models with respect to the: (a) testing; and (b) validation data sets.

Table 6.7 Distribution properties of LGP and ANN model predictions with respect to the validation set.

Parameter	Testing data			Validation data		
	Measured	LGP	ANN	Measured	LGP	ANN
Mean	10.50	10.34	11.02	10.43	10.49	11.20
Standard Deviation	8.66	7.54	7.34	8.03	7.03	6.96
Coefficient of Variation	0.82	0.73	0.67	0.77	0.67	0.62
Skewness	1.34	1.47	1.50	1.23	1.11	1.22
Minimum	0.29	0.40	1.04	0.39	0.59	0.82
Maximum	45.12	47.24	45.88	46.20	47.61	46.78
Range	44.83	46.84	44.84	45.81	47.02	45.96

Finally, further verification of the LGP-based model's predictive capability is carried out using a completely new, additional dataset, unseen by the model that lies within the data limits of the LGP model, as explained earlier. The verification dataset was discussed previously in §6.1 and further details are given there. The results are summarised in Table 6.8, together with the corresponding statistics obtained from the ANN model for comparison purposes. It is clear that both models perform very favourably with respect to this series of unseen CPT data, although the results in relation to *R*, *RMSE* and *MAE* present a somewhat inconsistent picture, as compared with those obtained thus far. Nevertheless, it can be concluded that the LGP-based model yields marginally superior performance to that of the ANN-based model.

Table 6.8 Performance statistics of LGP and ANN models with respect to verification data.

CPT location	*R*		*RMSE* (MPa)		*MAE* (MPa)	
	LGP	ANN	LGP	ANN	LGP	ANN
Port Botany – 30	0.96	0.96	4.21	3.63	3.74	3.39
Port Botany – 11	0.86	0.84	3.37	3.65	2.65	2.86
Port Botany – 45	0.96	0.97	7.11	6.06	5.97	5.30
Port Botany – 35	0.71	0.72	3.96	3.75	3.15	2.51
Potts Hill – 37/44	0.27	0.42	2.86	3.15	1.79	2.12
Potts Hill – 27/54	0.51	0.44	1.60	2.23	1.27	1.93
Potts Hill – 24/57	0.58	0.53	2.89	2.88	2.29	2.17
Outer Harbor – EFC 5	0.85	0.84	2.31	3.15	1.89	2.84
Banksmeadow – C 3	0.20	0.14	2.47	2.49	1.97	1.90
Cairns – CPT 2	0.54	0.61	4.68	4.26	3.32	2.95
Cairns – CPT 5	0.78	0.79	2.61	2.83	1.53	2.31
Cairns – CPT 8	0.96	0.94	2.02	2.55	1.59	2.27

6.6 SUMMARY AND CONCLUSIONS

This paper presents a unique approach for the prediction of the effectiveness of rolling dynamic compaction (RDC) based on genetic programming (GP). A reliable database consisting of cone penetration test (CPT) results obtained from several ground improvement projects, associated with the Broons 4-sided, 8 tonne ‘impact roller’, is utilised for the model development. The emphasis of the present study is placed on a particular variant of GP, namely linear-based genetic programming (LGP), which has significant benefits over most other modelling approaches. The models incorporate 4 input variables; depth (m), cone tip resistance (MPa) and sleeve friction (kPa) prior to compaction, and the number of roller passes, which together, are considered to be the

most effective in predicting the cone tip resistance after compaction (MPa) as the single model output.

The selected optimal LGP model yields high accuracy in model predictions with a coefficient of correlation (*R*) of 0.87, a root mean square error (*RMSE*) of 4.03 MPa, and a mean absolute error (*MAE*) of 2.71 MPa, when assessed using a set of unseen data. The optimal model is evaluated by means of a parametric study and it is apparent that the model is robust and has appropriately captured the input/output non-linear relationships. However, the investigation has revealed that the model performs best with initial cone tip resistance (q_{ci}) values less than or equal to 10 MPa. Moreover, the contributions of the each of the input variables with respect to model predictions are investigated in a sensitivity analysis and it is observed that each of the input variables are highly relevant to the prediction of cone tip resistance after compaction (q_{cf}). Furthermore, a series of numerical equations is formulated based on the optimal model parameters, which can readily be adopted in practice. Finally, the LGP simulations are compared with the existing ANN model subjected to several criteria suggested in the literature and both models are compared using a series of unseen CPT data. The results indicate that the LGP-based model marginally outperforms the ANN model and overall produces slightly more accurate predictions.

The LGP approach presented in this paper is considered to be valuable during the pre-planning and pre-design phases. However, it is not expected to replace or undervalue the importance of field trials but, nevertheless, it is a worthwhile additional tool for ground improvement projects involving RDC.

Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP120101761). The authors wish to acknowledge Mr. Stuart Bowes at Broons Hire (SA) Pty Ltd for his kind assistance and continuing support, especially in providing access to the in situ test results upon which the numerical models are based. The authors are also grateful to Mr. Brendan Scott for his contribution to this work.

Chapter 7.

Use of Genetic Programming for the Predictions of Effectiveness of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results

INTENTIONALLY BLANK

Statement of Authorship

Title of Paper	Use of Genetic Programming for the Predictions of Effectiveness of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results.
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016). Use of Genetic Programming for the Predictions of Effectiveness of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results. Geotechnical and Geological Engineering, Under Review.

Principal Author

Name of Principal Author (Candidate)	R. A. T. M. Ranasinghe		
Contribution to the Paper	Data collection and pre-processing, Model development based on Genetic Programming (GP), Analysis of results and interpretation, and Preparation of the manuscript.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	20/02/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

x. the candidate's stated contribution to the publication is accurate (as detailed above);

xi. permission is granted for the candidate in include the publication in the thesis; and

xii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mark B. Jaksa		
Contribution to the Paper	Overall percentage – 10% Initial concept, research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Yien Lik Kuo		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

Name of Co-Author	Fereydoon Pooya Nejad		
Contribution to the Paper	Overall percentage – 5% Assisting with research direction and editing manuscript.		
Signature		Date	21/02/2017

ABSTRACT

Rolling dynamic compaction (RDC), which employs non-circular module towed behind a tractor, is an innovative soil compaction method that has proven to be successful in many ground improvement applications. RDC involves repeatedly delivering high-energy impact blows onto the ground surface, which improves soil density and thus soil strength and stiffness. However, there exists a lack of methods to predict the effectiveness of RDC in different ground conditions, which has become a major obstacle to its adoption. This study develops a prediction model based on linear genetic programming (LGP), one of the common approaches in the application of artificial intelligence for non-linear forecasting. The models are based on in situ density-related data in terms of dynamic cone penetrometer (DCP) results obtained from several projects that have employed 4-sided RDC modules. It has been shown that the model is accurate and reliable over a range of soil types. Furthermore, a series of parametric studies confirms its robustness in generalising data. In addition, the results of the comparative study have indicated that the optimal LGP model has a better predictive performance than the existing artificial neural network (ANN) model developed earlier by the authors.

Keywords: *Ground improvement, rolling dynamic compaction, linear genetic programming, dynamic cone penetrometer test*

7.1 INTRODUCTION

Rapid urban and industrial growth has created a demand for construction on ground which has previously been considered unsuitable, such as collapsible and loose natural soils, former landfills, fill from mine workings and sites with prior uncontrolled filling. Rolling dynamic compaction (RDC) has found to be useful to improve such problematic soils and is now in widespread use globally in the construction industry. This technique involves towing a heavy (6 to 12 tonnes), non-circular (3-, 4- and 5-sided) module behind a tractor, where the module rotates about its corners as it is drawn forward (Avalle, 2004d). As a result, a series of high-energy impacts is imposed repeatedly onto the ground surface (Pinard, 1999) by which the soil is densified into a state of lower void ratio due to pore air expulsion. The high energy waves generated by the impact blows penetrate deeply into the ground resulting a greater influence depth, which is more than 1 m beneath the ground surface and sometimes in excess of 3 m in some soils (Avalle and Carter, 2005). This deep compaction effect is beneficial when compared to conventional static and vibratory compaction (Clegg and Berrangé, 1971; Clifford, 1976; Avalle and Carter, 2005; Jaksa et al., 2012), where the influence depth is limited to depths less than 0.5 m below the ground. In addition, it is efficient to employ RDC in large and open sites as the modules are drawn at the comparatively higher optimal speed of 9 to 12 km/h, whereas the traditional compaction rollers travel at 4 km/h speed (Pinard, 1999). Furthermore, RDC can also compact thicker lifts, i.e. in excess of 0.5 m whilst, with the conventional compaction, the lift thickness is usually limited between 0.2 and 0.5 m (Avalle, 2006). Thus, the improved ground compaction capability of RDC, especially with respect to a greater influence depth and a higher speed of compaction, results in increased productivity. In addition, the prudent use of RDC can also provide significant cost savings, reduced infrastructure costs and environmental benefits. Given these significant advantages over the traditional approaches of soil compaction, RDC applications are found to be successful in a variety of fields, including earthworks and pavement construction (Avalle, 2006); the agricultural sector (Avalle, 2004d); and in mining applications (Scott and Jaksa, 2012).

The estimation of the influence depth of RDC is of prime importance, in particular, if multi-layered soil profiles are encountered. Although RDC has been studied

experimentally through a number of field-based case studies, until recently as a result of work undertaken by the authors, there has been no rational means available for the prior estimation of the effectiveness of RDC in different ground conditions. Subsequently, current practice associated with estimating site-specific operational parameters relies heavily on the judgment of geotechnical engineering practitioners. In addition, field trials are often carried out, where a testing pad is arranged, which is representative of a large-scale operation of the compaction procedure. The efficacy of RDC is verified using a combination of surface settlement surveys, soil sampling and in situ tests, such as penetrometer, field density and geophysical testing, that are undertaken after different numbers of module passes. As such, field trials are valuable for ascertaining the relevant operational parameters, especially the optimal number of impact roller passes needed to achieve the required percentage of maximum dry density, but incurs a non-trivial cost and time commitment.

Until recently, as described below, no method was available to predict, a *priori*, the density improvement at a specified depth below ground due to RDC, based on subsurface conditions and the number of roller passes. With this in mind, this research investigates the feasibility of using linear genetic programming (LGP), which is one of the well-known machine learning techniques available to develop predictive models. Recently, the authors have also suggested an approach based on artificial neural networks (ANN) (Ranasinghe et al., 2016c, 2016a), in which two distinct models have been developed based on in situ soil test data in terms of cone penetration test (CPT) and dynamic cone penetrometer (DCP) test results obtained from projects that have employed the Broons 4-sided ‘impact roller’. Whilst the developed ANN models were demonstrated to be accurate and reliable, it has been suggested in the literature (Rezania and Javadi, 2007; Alavi and Gandomi, 2012; Alavi and Sadrossadat, 2016) that evolutionary computation, such as LGP, offers a number of advantages over ANNs and thus, may yield improved predictive capability with respect to RDC. Furthermore, many of the applications related to geotechnical engineering, including the recent study by the authors in relation to RDC (Ranasinghe et al., 2016b), where LGP models were developed using CPT data, have suggested that LGP outperforms ANNs, in addition to other benefits.

The developed LGP-based model in this study utilises a reliable database consisting of DCP results obtained from several ground improvement projects, associated with the Broons 4-sided, 8 tonne 'impact roller'. Since this dataset has also been employed previously in the ANN model developed recently by the authors (Ranasinghe et al., 2016a), a comparative study is conducted between the results obtained herein and with those obtained from the existing ANN model. In addition, a parametric study is conducted, by which the reliability of the developed model is further verified.

7.2 LINEAR GENETIC PROGRAMMING (LGP)

Genetic programming (GP) is an evolutionary computational approach of non-linear modelling, where the computer programs evolve automatically to optimise a solution towards a pre-defined goal. This supervised machine learning technique aligns with the theory of Darwinian natural selection and was first introduced by Koza (1992). Generally, GP is considered as an extension to genetic algorithms (GAs), in which most of the genetic operators used in GAs are also applicable, albeit with slight modifications (Alavi et al., 2013). However, the main differences between GP and GAs lie in the representation of the solution. GAs are often recognised by individuals represented by fixed-length binary strings (Holland, 1975) and the solutions require post-processing prior to execution. In contrast, GP represents the individuals as computer programs whose size, shape and complexity is dynamically varied during evolution and are usually executable without post-processing (Koza, 1992). Moreover, GAs are generally used for parameter optimisation, where the best values are evolved for a pre-defined set of model parameters, whilst GP, on the other hand, evolves the program structure of the approximation model together with the values of its parameter setting (Torres et al., 2009; Mousavi et al., 2011; Alavi et al., 2013). However, as with GAs, GP performs a multi-directional simultaneous search for an optimal solution from a pool of many potential solutions, collectively known as a 'population'. The fact that these methods operate from a population enables them to escape local minima in the error surface and are thus able to find optimal or near optimal solutions (Selle and Muttil, 2011).

In the traditional GP approach, which is also referred to as tree-based genetic programming (TGP), the computer programs (individuals) have a symbolic representation of a rooted tree-like structure with ordered branches in which the root node and internal nodes are comprised of functions whereas, external nodes (leaves) contain the input values or constants (Koza, 1992). Thus, they are often expressed in a functional programming language like LISP (Koza, 1992). However, in addition to the traditional TGP, there are several other distinguished subsets of GP that have a different form of program structure representation, i.e. linear GP and graph-based GP (Banzhaf et al., 1998; Poli et al., 2007; Alavi et al., 2013). In the present study, emphasis is placed on linear genetic programming (LGP). In this particular variant, the evolved programs are represented by a sequence of instructions, either from an imperative language (e.g. C, C++ or Java) (Brameier and Banzhaf, 2001; Brameier and Banzhaf, 2007) or from a machine language (Nordin, 1994). In contrast to the rigidly determined, tree-structured data flow in TGP, LGP has a more general, specifically-directed graphical structure at the functional level resulting from multiple usage of register contents (Brameier and Banzhaf, 2007; Alavi et al., 2013; Gandomi et al., 2014). Moreover, the existence of noneffective code segments, which are also referred as introns in LGP, makes them different from their traditional tree-based counterparts. As such, these structurally noneffective codes denote the instructions, which manipulate the registers that have no influence on the output calculation (Gandomi et al., 2010a). Although these noneffective code segments coexist with the effective code, they are not connected to the data flow unlike in TGP, where the structural introns do not exist because all the program components have a connection with the root node (Brameier and Banzhaf, 2007). However, because of the imperative program structure in LGP, the structural introns can be detected efficiently and completely (Francone and Deschaine, 2004; Alavi et al., 2013).

There is a special variant of LGP, named automatic induction of machine code by genetic programming (AIMGP), where the individuals are represented and manipulated as native binary machine code (Nordin, 1994; Banzhaf et al., 1998). During fitness evaluation in GP, the programs are executed multiple times or at least once, which is considered to be the most time-critical step in evolutionary algorithms

(Brameier and Banzhaf, 2007). Program execution refers to the interpretation of internal program representation. However, in AIMGP the individuals are directly executable by the processor and that avoids the use of an expensive interpreter (Francone and Deschaine, 2004; Brameier and Banzhaf, 2007). As a result, AIMGP is found to be significantly faster and more memory efficient when compared with other interpreting GP variants (Nordin, 1994; Brameier and Banzhaf, 2001). Given these advantages, AIMGP is also utilised in this study.

In general, the basic steps of the LGP evolutionary algorithm (Brameier and Banzhaf, 2007) can be summarised as follows:

1. Initialising a population of randomly generated programs and evaluating their fitness;
2. Performing two fitness tournaments with randomly selected programs from the population and selecting the winning programs;
3. Making temporary copies of the two winning programs;
4. Transforming the two winning programs into offsprings subjected to genetic operations, i.e. crossover and mutation with certain probabilities;
5. Replacing the two tournaments losing programs with the temporary copies of the winning programs; and
6. Repeating Steps 2 to 5 until the termination or convergence criteria are satisfied.

7.3 METHODOLOGY

The details of the database that is used to develop the LGP-based model, as well as the methodology adopted for model development, are discussed separately below.

7.3.1 Database, Data Analysis and Data Pre-processing

This study utilises a comprehensive database that comprises in situ strength data in the form of dynamic cone penetrometer (DCP) test results. The DCP (Standards

Association of Australia, 1997) is one of the most commonly used in situ test methods, which provides an indication of soil strength in terms of rate of penetration (blows/mm). The database contains the results of DCP tests with respect to the number of roller passes obtained from various sites and soil types. The relevant data have been extracted from the results of several field trials undertaken using the 4-sided, 8 tonne 'impact roller' (BH-1300), which is operated by the Australian company Broons (SA) Hire. In total, the database contains 2,048 DCP records from 12 separate projects.

Given the problem at hand, the model is established to predict the degree of soil improvement of the ground with respect to the number of roller passes. Thus, the single model output should necessarily define the ground density at a particular location resulting after several passes of the impact roller. However, in selecting the model input variables, it is essential to incorporate the factors that are most influential on the model output variable. There are several fundamental parameters that significantly affect soil compaction: the geotechnical properties at the time of compaction, such as ground density, moisture content, soil type; and the amount of energy imparted to the ground during compaction. Consequently, the model input variables are defined so that they effectively address each of the aforementioned factors that influence soil behaviour upon the application of RDC.

Whilst the standard DCP procedure involves recording the number of blows for each 50 mm of penetration, a compromise must be achieved between model parsimony and predictive accuracy. In this study, the average DCP blow count per 300 mm is used to indicate the average density with depth. Therefore, the initial density at the point of interest is selected to be described by the input variable of *initial DCP count (blows/300 mm)*, whilst the single output variable is described by the *final DCP count (blows/300 mm)*. In addition, the amount of energy imparted to the ground during RDC is described in terms of the number of roller passes so that two input variables are specified: the *initial*; and the *final number of roller passes* corresponding to the initial and final density at a particular location, respectively. The *average depth (m)* is established as another input variable, whilst the *soil type* is also adopted. The soil type is defined in a generalised form at each DCP location by implementing a primary

(dominant) and a secondary soil type. With the availability of project data in the DCP database, 4 distinct soil types are characterised as: (i) Sand–Clay; (ii) Clay–Silt; (iii) Sand–None; and (iv) Sand–Gravel. However, it is worth noting that soil moisture content is not included as a model input due to the paucity of data, as it is usually not measured routinely in practice in ground improvement projects. However, the moisture content is considered to be described implicitly by the DCP data, given that penetrometer test results, including those from the DCP, are affected by soil moisture. The input and output variables involved in LGP-based model development and their statistics are presented in Table 7.1, with Figure 7.1 summarising the histograms of the model input and output variables.

Table 7.1 Descriptive statistics of the dataset used in LGP model development.

Non-numerical input						
Soil type	Sand–Clay, Clay–Silt, Sand–None, Sand–Gravel					
Numerical input[+] and output[#]						
Variables	**Mean**	***SD**	**Minimum**	**Maximum**	**Skewness**	**Kurtosis**
Average depth, D (m) [+]	0.82	0.52	0.15	1.95	0.33	–0.74
Initial no. of roller passes[+]	7.89	10.65	0.00	50.00	1.63	2.06
Initial DCP (blows/300 mm) [+]	16.41	10.69	3.00	65.00	1.51	2.64
Final no. of roller passes[+]	21.13	16.25	2.00	60.00	1.01	–0.03
Final DCP (blows/300 mm) [#]	18.14	11.25	2.00	84.00	1.42	2.94

* SD – Standard Deviation

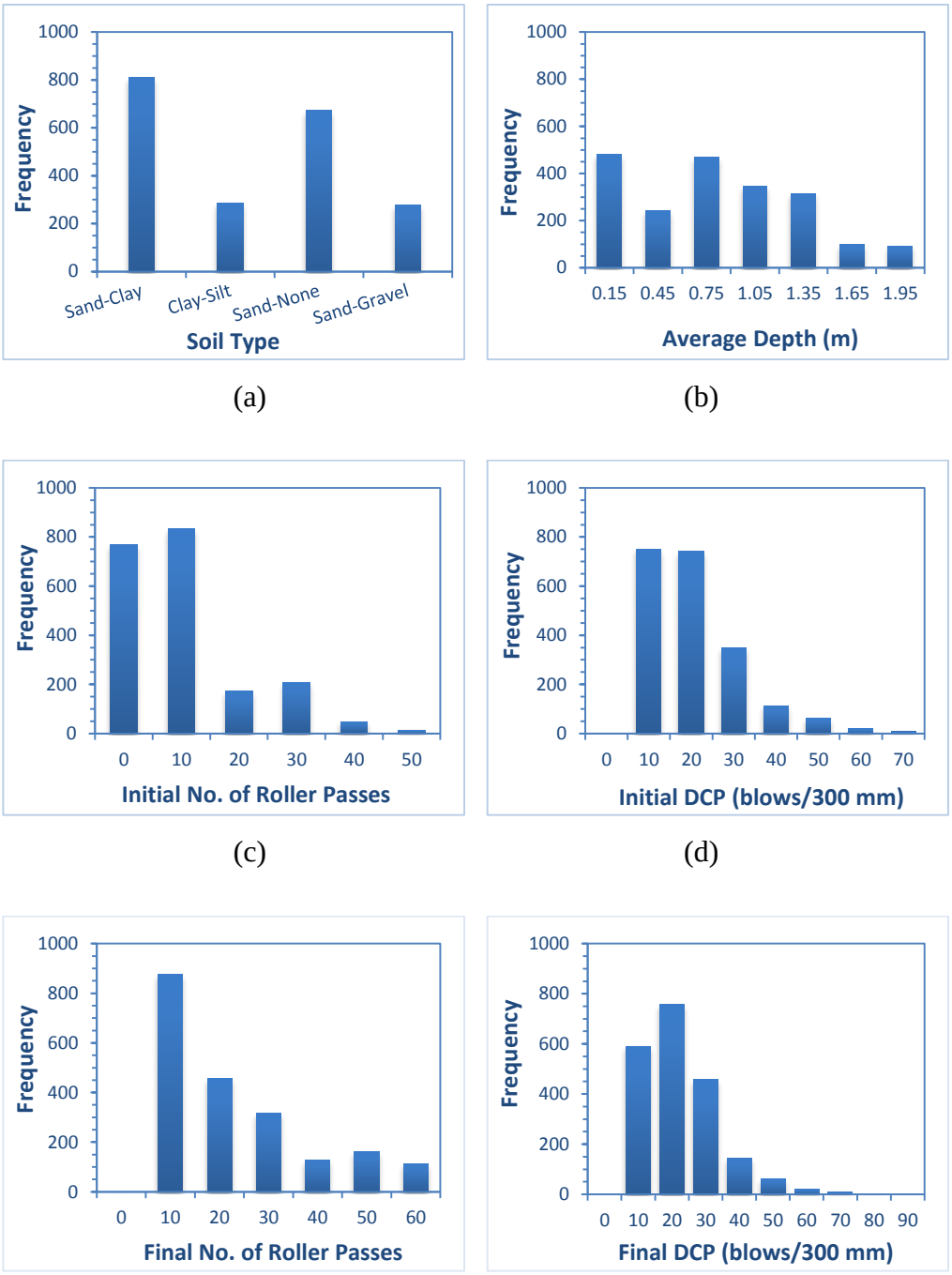


Figure 7.1 Histograms of the model variables used in LGP model development: (a) soil type; (b) average depth; (c) initial number of roller passes; (d) initial DCP; (e) final number of roller passes; and (f) final DCP.

Prior to model development, the entire dataset is subdivided into 2 subsets, calibration and validation data. The calibration dataset is further subdivided into training and testing sets, by which the models are respectively trained and the best program is selected by testing. The testing set provides an estimate of the prediction error for a set of unseen data during the model calibration phase and this information is useful when selecting the optimal program model. The validation dataset is not used during model development, and thus, it is optional to provide this separate and additional dataset. However, it is with the validation dataset that the selected optimal model is assessed for its generalisation capabilities. Since the optimal model is evaluated with respect to an unseen data set, the results are significant for the evaluation of model performance.

In order to allow a fair comparison between the results obtained from the proposed LGP-based model and those from the existing ANN model (Ranasinghe et al., 2016a), the same 3 data subsets are employed in the present study as was undertaken by the authors in the earlier ANN model study. In summary, 80% of the data are used for training and testing (1,310 and 319 records, respectively), whilst the remaining 20% of the data (419 records) are used for model validation. However, it is important to maintain similar statistics between these 3 subsets, which ensures that they belong to the same population, which is the ideal form of data division. As can be observed from the summary statistics in Table 7.2, the subsets effectively represent the same population by the similar values of mean, standard deviation, minimum, maximum and range.

7.3.2 LGP-based Modelling Approach

In this study, the commercially available software *Discipulus* version 5.2 (Francone, 2010) is used for the LGP-based model development. It is a supervised learning system, which operates on the basis of the AIMGP platform. The selection of control parameters is considered to be vital in LGP modelling, since it has a direct impact on the model's generalisation capacity. Therefore, in this study, the control parameters are defined in accordance with the suggested values from similar LGP applications (Gandomi et al., 2010a; Rashed et al., 2012; Alavi et al., 2013; Babanajad et al., 2013;

Gandomi et al., 2014; Alavi and Sadrossadat, 2016) and also from observations obtained from preliminary runs. As presented in Table 7.3, several different parameter settings, including population size, probabilities of genetic operations and program size, are investigated, whilst most of the other minor parameters are defined by the software default values.

Table 7.2 Statistical properties of the data subsets.

Model variables	Subset	Statistical parameters				
		Mean	*SD	Minimum	Maximum	Range
Input						
Average depth (m)	Training	0.81	0.51	0.15	1.95	1.80
	Testing	0.82	0.51	0.15	1.95	1.80
	Validation	0.83	0.52	0.15	1.95	1.80
Initial no. of roller passes	Training	7.69	10.61	0.00	50.00	50.00
	Testing	7.65	10.44	0.00	50.00	50.00
	Validation	8.71	10.93	0.00	50.00	50.00
Initial DCP (blows/300 mm)	Training	16.57	10.86	3.00	65.00	62.00
	Testing	15.88	10.64	3.00	59.00	56.00
	Validation	16.31	10.20	3.00	61.00	58.00
Final no. of roller passes	Training	21.14	16.25	2.00	60.00	58.00
	Testing	21.16	16.49	2.00	60.00	58.00
	Validation	21.08	16.11	2.00	60.00	58.00
Output						
Final DCP (blows/300 mm)	Training	18.30	11.29	2.00	84.00	82.00
	Testing	17.80	10.81	2.00	73.00	71.00
	Validation	17.93	11.47	3.00	75.00	72.00

* SD – Standard Deviation

Table 7.3 Parameter setting used in LGP modeling.

Parameter	Settings
Function set	+, –, ×, /, Absolute, Square Root, Trigonometric (*sin*, *cos*), Exponential
Population size	500, 1,000, 2,000, 5,000
Initial program size	80 bytes
Maximum program size	128, 256 bytes
Mutation frequency	50%, 95%
Block mutation frequency	40%
Instruction mutation frequency	30%
Instruction data mutation frequency	30%
Crossover frequency	50%, 95%
Homologous crossover frequency	95%

In this study, several LGP projects are carried out including only the arithmetic functions (+, –, ×, /). Furthermore, in order to permit the evolution of highly non-linear models, inclusion of mathematical functions (*sin*, *cos*, *exponential*, *absolute* and *square root*), in addition to basic arithmetic operators, is also considered. This study applies the mean square error (*MSE*) as the fitness measure. Equation 7.1 defines the *MSE*, where d_i and y_i are respectively the actual and the model predicted output values for the i^{th} sample and n denotes the number of samples.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(d_i - y_i)^2 \qquad (7.1)$$

The population size parameter is regulated at several different levels; i.e. 500; 1,000; 2,000 and 5,000. However, it has been found that the evolutionary process converges faster in semi-isolated sub-populations, named ‘demes’ than in a single population of equal size (Brameier and Banzhaf, 2007; Alavi and Sadrossadat, 2016). Thus, the parameter that determines the number of demes into which the population is subdivided is set at 20 (Alavi and Gandomi, 2012; Alavi and Sadrossadat, 2016). As

discussed earlier, in LGP essentially two genetic operations, crossover and mutation, are used. In this study, the frequencies of these genetic operations are considered at two levels 50% and 95%. These frequencies are the overall probabilities of genetic operations applied to the tournament winning programs (Koza, 1992). It has been suggested in the literature, that the success of the LGP algorithm usually rises with the increasing program size parameter (Rashed et al., 2012; Alavi et al., 2013). However, at the same time, as the complexity of the evolved programs increases the convergence speed decreases. Considering these trade-offs, the initial program size is set to 80 bytes, whilst the maximum program size is tested at two optimal levels, 128 and 256 bytes.

Likewise, in this study, many numbers of LGP projects are carried out, where all of the above described combinations of parameters are tested. Each LGP project is made up of a series of runs, which progressively increases in length during the course of a project. Each run is allowed to evolve in generations, while *MSE* is being monitored continuously. However, the projects are terminated manually, given a reasonable time (ranging from a few minutes to more than 20 hours on a standard PC) to evolve into an accurate model and when no further improvement in model performance is likely to occur. Finally, the resulting LGP models are evaluated using several performance measures with respect to each of the 3 data subsets and compared to select the optimal program. The criteria used to evaluate the performance of the evolved program models include the coefficient of correlation (*R*), root mean square error (*RMSE*) and mean absolute error (*MAE*).

7.4 RESULTS AND DISCUSSION

The following sections summarise the results of the optimal LGP model along with a comparison with those obtained from the existing ANN model. Moreover, the details of parametric study and sensitivity analysis are also discussed. In addition, the selected optimal LGP model for predicting the final DCP (blows/300 mm) is presented in computing code in the C language in Appendix B.

7.4.1 Performance Analysis

The performance statistics in terms of *R*, *RMSE* and *MAE* associated with the selected optimal LGP-based model, with respect to the 3 data subsets (training, testing and validation), are presented in Table 7.4. The selected model's performance and reliability is assessed based on the criteria suggested by Smith (1986), as in the following:

Given that the error values (e.g. *RMSE* and *MAE*) are minimum, when:

- $|R| \geq 0.8$, there exists a strong correlation;
- $0.2 < |R| < 0.8$, there exists a correlation; and
- $|R| \leq 0.2$, there exists a weak correlation between the two variables.

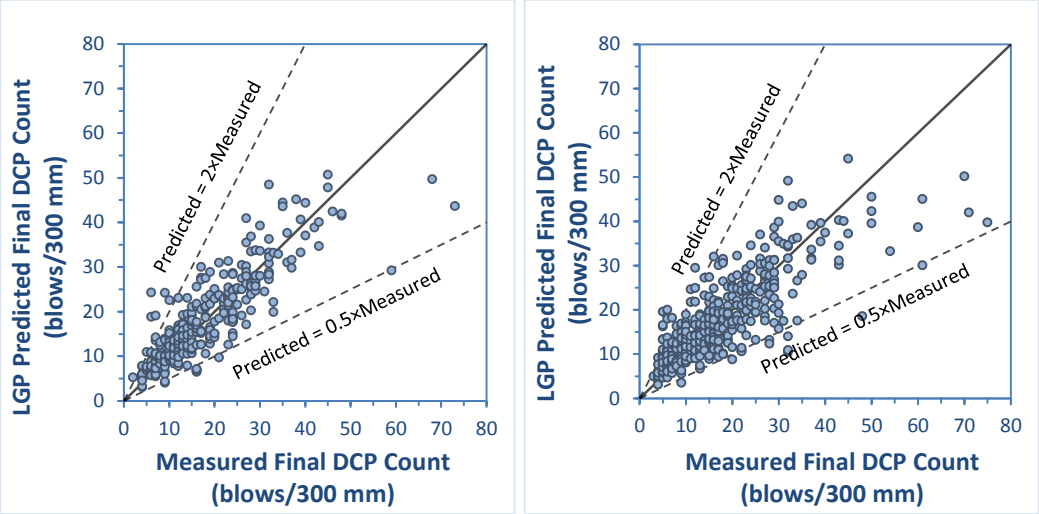
Table 7.4 Performance statistics of the selected optimal LGP model.

Data subset	**Performance criteria**		
	R	***RMSE* (blows/300 mm)**	***MAE* (blows/300 mm)**
Training	0.84	6.22	4.18
Testing	0.87	5.35	3.70
Validation	0.81	6.80	4.74

Accordingly, it can be concluded that there exists a strong correlation between the model's predicted results and the measured data since $R \geq 0.8$ and measures of error (i.e. *RMSE* and *MAE*) are relatively small. It is also evident that the above criterion is valid and not limited to the data subsets that were used during the model calibration phase (i.e. training and testing sets), but also the new unseen data in the validation set as well. This implies that the model predicts the target values accurately and also incorporates a generalisation capability.

Figure 7.2 presents scatter plots of the final DCP count predicted from the LGP model and compared against the measured values in the testing and validation sets. It can be observed that the results are scattered around the solid line that indicates the line of

equality and the spread exhibits classic heteroscedasticity. The spread is confined to 2 standard error (SE) envelops of 0.5 to 2 times the measured values. These SE bands can be considered reasonable for such a model that yields predictions based on DCP results given the uncertainties associated with the dataset and the method itself.



(a) (b)

Figure 7.2 Measured versus LGP predicted final DCP count with respect to: (a) testing data set; and (b) validation data set.

In order to investigate further the model's performance, the LGP predictions and the measured data, with respect to the validation data set, are assessed subject to several additional measures, as suggested in literature. Table 7.5 presents the validation criteria along with the results obtained from the LGP model. It is evident that satisfactory results are obtained from each of the criteria. This provides further evidence that the optimal LGP model yields accurate predictions.

A comparative study is conducted between the results obtained from the LGP-based model and with those obtained from the ANN-based model recently developed by the authors (Ranasinghe et al., 2016a). The performance indices in terms of *R*, *RMSE* and *MAE* are again adopted and the results are presented in Table 7.6. As can be seen, the results for both models are very promising, as indicated by the strong correlation coefficient ($R \geq 0.8$), along with the relatively low error values with respect to all 3

data subsets. However, it is also evident that the LGP predictions are slightly superior to those from the ANN model.

Table 7.5 Additional performance measures of the LGP model for validation data.

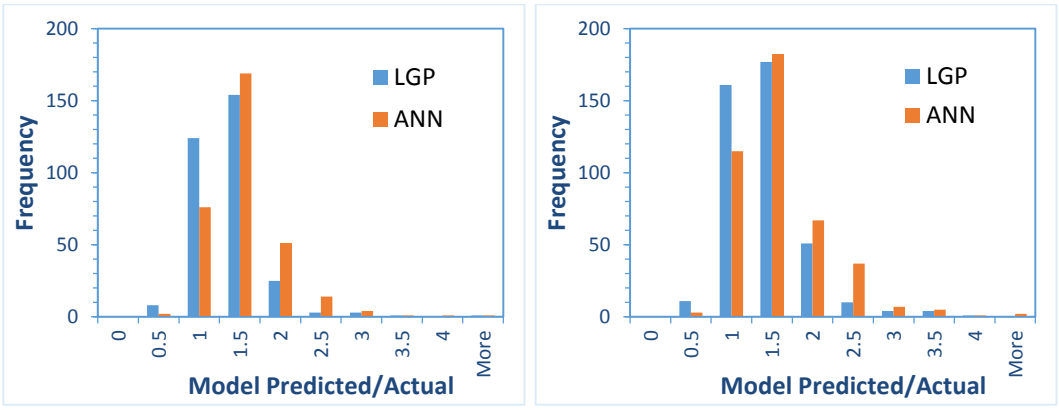
Item	Formula	Reference	Condition	Result
1	$k = \frac{\sum_{i=1}^{n}(d_i \times y_i)}{\sum_{i=1}^{n} {d_i}^2}$	Golbraikh and Tropsha (2002)	$0.85 < k < 1.15$	0.92
2	$k' = \frac{\sum_{i=1}^{n}(d_i \times y_i)}{\sum_{i=1}^{n} {y_i}^2}$	Golbraikh and Tropsha (2002)	$0.85 < k' < 1.15$	0.98
3	${R_0}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - {d_i}^0)^2}{\sum_{i=1}^{n}(y_i - \bar{y_i})^2}$ Where; ${d_i}^0 = k \times y_i$	Roy and Roy (2008)	Should be close to 1	0.97
4	${R_0}'^2 = 1 - \frac{\sum_{i=1}^{n}(d_i - {y_i}^0)^2}{\sum_{i=1}^{n}(d_i - \bar{d_i})^2}$ Where; ${y_i}^0 = k' \times d_i$	Roy and Roy (2008)	Should be close to 1	1.00
5	$m = \frac{R^2 - {R_0}^2}{R^2}$	Golbraikh and Tropsha (2002)	$m < 0.1$	–0.49
6	$n = \frac{R^2 - {R_0}'^2}{R^2}$	Golbraikh and Tropsha (2002)	$n < 0.1$	–0.54

Table 7.6 Comparison of the performance statistics of LGP and ANN model.

Data subset	Performance criteria					
	R		***RMSE*** (blows/300 mm)		***MAE*** (blows/300 mm)	
	LGP	**ANN**	**LGP**	**ANN**	**LGP**	**ANN**
Training	0.84	0.85	6.22	6.45	4.18	4.88
Testing	0.87	0.83	5.35	6.52	3.70	4.74
Validation	0.81	0.79	6.80	7.54	4.74	5.59

Further to the above assessment, the predictions of final DCP from the LGP and ANN models are compared graphically with the measured final DCP values, with respect to the testing and validation datasets, and the resulting histograms are shown in Figure 7.3. The *x*-axis indicates the ratio of the predicted data to measured values, with ideal

performance being indicted by a ratio of unity. As can be seen, both the LGP and ANN model predictions from the testing and validation data suggest that they have strong predictive abilities and generalisation performance, as given by the high frequencies around the ratio of 1. In addition, it can be clearly seen that the LGP model slightly outperforms the ANN model.



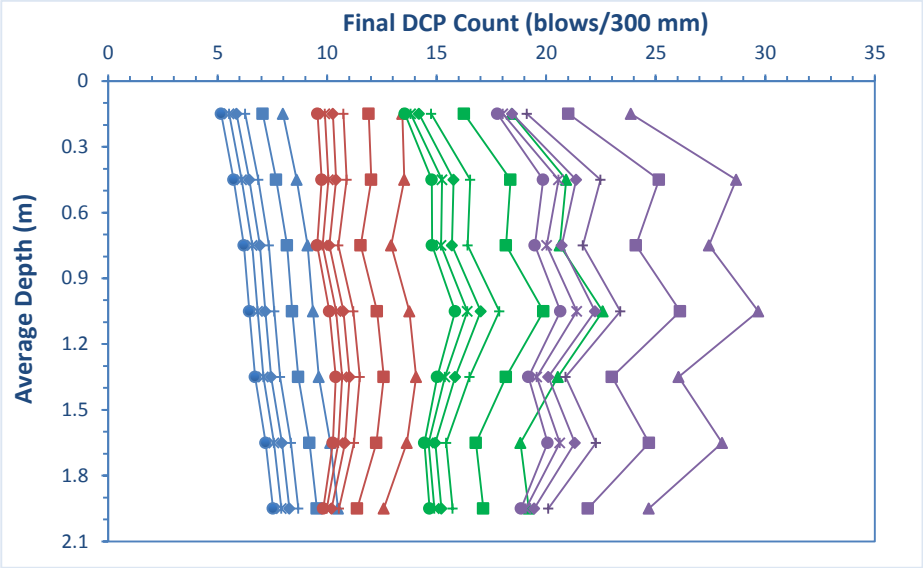
(a) (b)

Figure 7.3 Frequency histograms for model predicted to measured final DCP with respect to: (a) testing data; and (b) validation data.

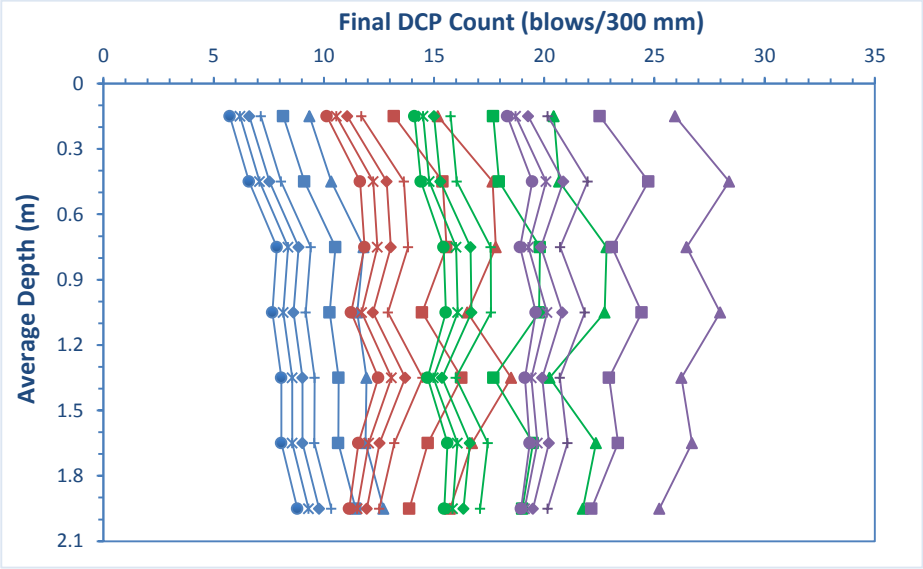
7.4.2 Parametric Study

Although the LGP model yields satisfactory performance in terms of the measures discussed above, it is essential to investigate the model's behaviour in a parametric study to further test its robustness. To this end, the LGP model is implemented to predict the output for a synthetic input data set that lies within the range that model is trained against, in order to examine whether the results conform to the known physical behaviour of the system. Each of the model input variables is varied successively, while maintaining all other input variables at a pre-defined value. In this case, the model output of final DCP count (blows/300 mm), which expresses the post compaction condition, is examined given the different initial conditions as described by initial DCP count (i.e. 5, 10, 15, 20 blows/300 mm), soil types (i.e. Sand–Clay, Clay–Silt, Sand–None and Sand–Gravel) and final numbers of roller passes (i.e. 5, 10, 15, 20, 30, 40 passes), while the initial number of roller passes is set at zero.

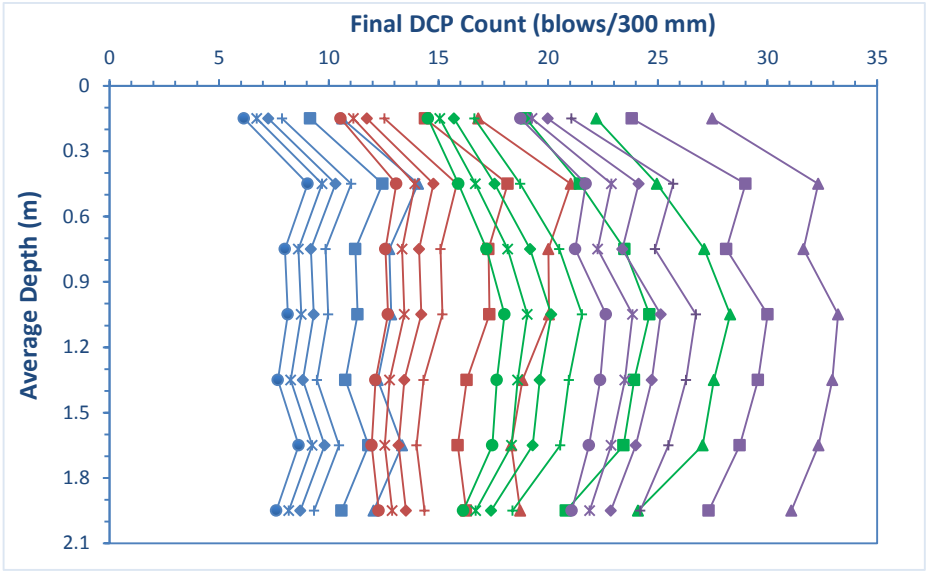
The results of the parametric study are shown in Figure 7.4 and indicate that the final DCP count continuously increases when the final number of roller passes is increased for a given initial DCP value in each soil type. This parametric behaviour demonstrates that the ground is significantly improved with the application of RDC, as evidenced by the increasing DCP count, which is consistent with the behaviour observed from field trials. Therefore, it can be concluded that the optimal LGP model developed in this study is robust, accurate and reliable within the specified range of the input variables (i.e. data ranges in training set).



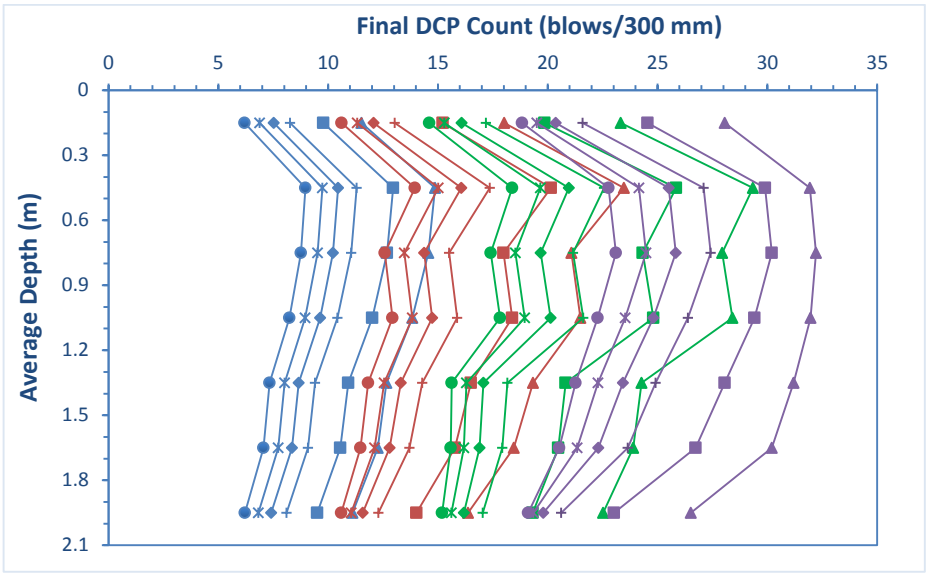
(a)



(b)



(c)



(d)

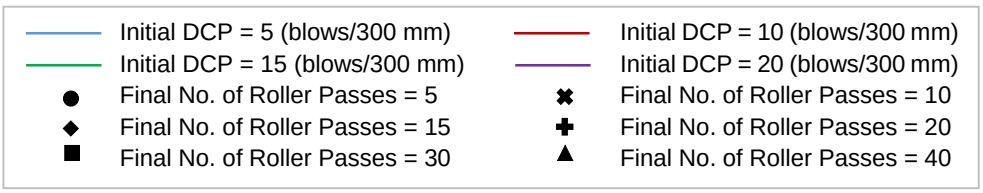


Figure 7.4 Variation of final DCP with respect to initial DCP and final number of roller passes in: (a) Sand–Clay; (b) Clay–Silt; (c) Sand–None; and (d) Sand–Gravel.

7.4.3 Sensitivity Analysis

It is informative to conduct a sensitivity analysis to evaluate the contribution of each input variable in predicting the target output. This is a common approach and has been utilised in number of applications (Gandomi et al., 2010a; Rashed et al., 2012; Alavi et al., 2013; Alavi and Sadrossadat, 2016). At the end of each LGP project, the software *Discipulus* calculates the frequencies, average and maximum impacts of each of the input variables, with reference to the 30 best selected programs (Francone, 2010). The frequency indicates the proportion of times (expressed as a percentage) that each input variable appears, in the best 30 evolved programs, in a way that contributes to the fitness of the programs that contain them. In this particular project, a frequency of 100% was obtained for each of the specified input variables (i.e. soil type, average depth, initial number of roller passes, initial DCP count and final number of roller passes), indicating that they are all important. In addition, the average and maximum impact values with respect to each of the input variables, which describes the average and maximum effect that results from the removal of each corresponding input variable from the 30 best programs, are also calculated and the resulting histogram is presented in Figure 7.5. It is again evident that all the input variables are significant, with respect to the predictions of the final DCP count, and the initial DCP count and the final number of roller passes are the most significant.

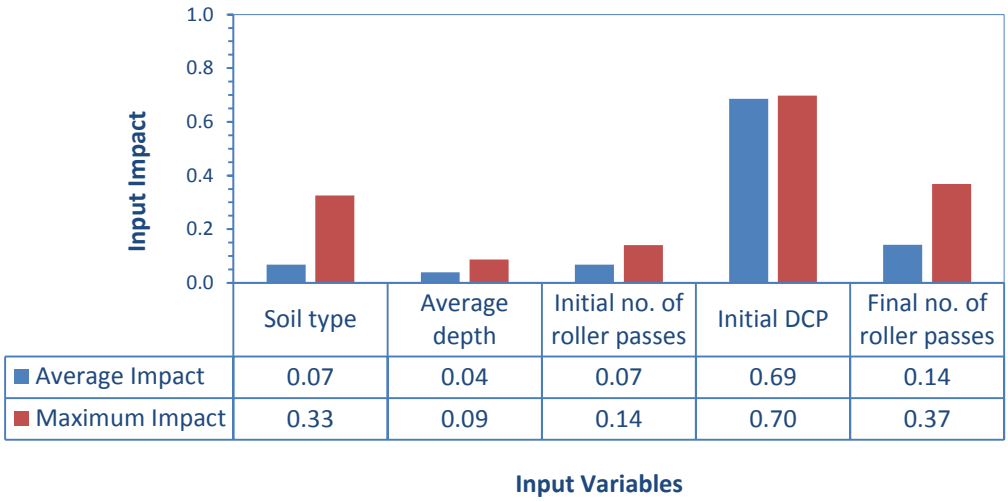


Figure 7.5 Contribution of the input variables on optimal LGP model predictions.

7.5 SUMMARY AND CONCLUSIONS

This paper presents a new approach based on genetic programming (GP) for the predictions of the efficacy of rolling dynamic compaction (RDC), which is considered to be an alternative to conventional soil compaction technology. A particular variant of GP, namely linear genetic programming (LGP), is used to develop the model. A comprehensive database consisting of in situ density data in terms of dynamic cone penetrometer (DCP) test results, associated with various ground improvement project records involving the Broons' 4-sided, 8 tonne 'impact roller' is utilised. In model development, 5 input variables [soil type, average depth (m), initial number of roller passes, initial DCP count (blows/300 mm) and the final number of roller passes] are considered to be the most influential with respect to predictions of the final DCP count (blows/300 mm), which is the sole output of the predictive models.

The selected optimal LGP model is found to yield accurate estimates of the final DCP count, with a coefficient of correlation (*R*) of 0.81, a root mean square error (*RMSE*) of 6.80 (blows/300 mm), and a mean absolute error (*MAE*) of 4.74 (blows/300 mm), when assessed against unseen data in the validation set. These outcomes confirm that the LGP model yields accurate predictions and demonstrates very good generalisation capability. Moreover, when the selected optimal LGP model results are compared with those obtained from the ANN model developed by the authors in a previous study, the LGP model demonstrates superior performance. In addition, a parametric study has been carried out for further verification of the LGP model and it is evident that the model predictions are accurate and robust. In addition, a sensitivity analysis has been conducted that examines the contribution of each input variable on the final model predictions. The results indicate that the input variables utilised in this study are highly significant with respect to the predictions of the final DCP count resulting from the application of RDC. The optimal LGP program is provided in a C code in order to disseminate the model and facilitate its use in practice.

The LGP model presented in this study it is expected to provide initial estimates of the effectiveness of RDC in different ground conditions, which are likely to be of particular value in the pre-design phase. It is, nevertheless, recommended that the

model predictions be validated on site using a traditional field trial, as the data upon which the model is based incorporates a limited number of soil types.

Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP120101761). The authors wish to acknowledge Mr. Stuart Bowes from Broons Hire (SA) Pty Ltd for his kind assistance and continuing support, especially in providing access to the in situ test results upon which the numerical models are based. The authors are also grateful to Mr. Brendan Scott for his contribution to this work.

Chapter 8.

Optimal Artificial Intelligence Models and Model Development

8.1 INTRODUCTION

As described in Chapters 4 – 7, two sets of AI models have been developed by means of ANNs and LGP incorporating two distinct databases that consisted of CPT and DCP test results obtained from RDC using the 4-sided, 8-tonne impact roller. In this chapter, these optimal models are synthesised and compared with each other to select the most feasible approach for predicting the effectiveness of RDC in different ground conditions with respect to CPT and DCP test data. This comparison is followed by a section presenting a comprehensive set of guidelines for each of the AI techniques employed in this research, i.e. ANN and LGP. With that, it is intended to provide guidance for potential and current users of these techniques, specifically the feed-forward type MLPs trained with error backpropagation algorithm and the LGP method, to facilitate the development of such AI models. In addition, treatment is given to the currently available approaches, areas that require consideration during model development, challenges encountered during the research, as well as means by which to address these challenges.

8.2 SYNTHESIS OF THE OPTIMAL AI MODELS

As discussed previously, two comprehensive databases, consisting of CPT and DCP test data, divided into 3 subsets (training, testing and validation), were employed for the development of ANN and LGP models for predicting the effectiveness of RDC in different ground conditions. Each of the selected optimal models of these two techniques was tested in a parametric study to assess generalisation ability and model robustness. Moreover, optimal model predictions were further examined by using several other means such as analysing performance measures and distribution properties. In addition, the performance of the optimal ANN and LGP-based models, as well as other aspects, were compared with one another. The results of these analyses and evaluations are summarised below. A more detailed review was included in an international, peer-reviewed conference paper given in Appendix C.

8.2.1 Optimal AI Models based on CPT Data

The CPT data analysis, model development and optimal model results were detailed previously in Chapters 4 and 6 in regards to the ANN and LGP techniques, respectively. However, overall, the following results and conclusions are obtained:

- These AI models utilised 4 input variables: depth (m); cone tip resistance prior to compaction (MPa); sleeve friction prior to compaction (kPa); and the number of roller passes, along with a single output of cone tip resistance after compaction (MPa).

- The resulting optimal ANN- and LGP-based models were found to be capable of predicting the target values to a high degree of accuracy, with a high coefficient of correlation ($R > 0.8$) and with low prediction error values, i.e. *RMSE* and *MAE*, when validated against a set of unseen data. Specifically, the optimal ANN and LGP models yielded values of *R* of 0.86, 0.87; *RMSE* of 4.16, 4.03 MPa; and *MAE* of 2.93, 2.71 MPa, respectively, when assessed against the validation set data. Overall, both models were found to have similar performance but it was also evident that the LGP-based model marginally outperforms the ANN model.

- By means of a set of additional performance measures, the relative performance of both models was further examined, by which it was again confirmed that the ANN and LGP models provide accurate predictive capabilities.

- It was observed that the distributions of the ANN and LGP model predictions with respect to the testing and validation set data are very similar to the measured data distribution. Nonetheless, it was also observed that the LGP model predictions more closely match the

measured data distribution and yields slightly superior distribution properties than those of the ANN model predictions.

- It was observed that both the optimal AI models provide very good estimates of q_{cf}, when presented with a completely new, additional series of unseen CPT datasets. The model predictions were found to have a strong correlation with the measured data. However, it was observed that the results in relation to *R*, *RMSE* and *MAE* are somewhat inconsistent with those of the previous observations, given that in some cases, the ANN-based model outperforms the optimal LGP model.

- By all these means, it was confirmed that both the optimal ANN- and LGP-based models yield reliable estimates of cone tip resistance relative to different numbers of impact roller passes.

- When the optimal models were evaluated in a parametric study, it was observed that the model predictions essentially replicate the expected underlying physical behaviour of RDC compaction. It was observed that the distinct non-linear relationship between parameters was appropriately captured by the optimal models and thus, the optimal model predictions are considered to be robust and reliable.

- In contrast, it was found that the model predictions are less likely to be satisfactory when forecast with respect to high values of initial cone resistance, specifically, when the q_{ci} is greater than 10 MPa. It was assumed that both the ANN and LGP models have been calibrated more appropriately for lower values of q_{ci} due to the lack of data in the CPT database in relation to higher values of q_{ci}. Therefore, both these optimal AI models are recommended to be used with confidence when the q_{ci} values are less than or equal to 10 MPa.

- By means of a sensitivity analysis on the optimal LGP-based model, the contributions of each of the input variables with respect to model predictions were investigated. It was confirmed that all the selected input variables have a substantial effect on the model predictions of q_{cf}.

- Finally, the optimal AI models were translated into a series of numerical equations to facilitate their adoption in practice. Formulation of explicit formulae employed the optimal weight combination of the ANN model, whilst the LGP model was translated into C code by the *Discipulus* software.

8.2.2 Optimal AI Models based on DCP Data

The details of model development by utilising DCP data and the optimal model results were described previously in Chapters 5 and 7 in relation to the ANN and LGP techniques, respectively. Chapter 7 also provided a comparison between the results obtained from both of these AI techniques. In summary, the following results and conclusions are obtained:

- The AI models using DCP data consisted of 5 input variables: soil type; average depth, *D* (m); initial number of roller passes; initial DCP count (blows/300 mm); and the final number of roller passes, whilst the sole output was the final DCP count (blows/300 mm) at depth *D* after compaction.

- As with the AI models based on CPT data summarised above, AI models developed by utilising the DCP dataset were found to be effective when predicting the density development of the ground with respect to the number of impact roller passes and the associated subsurface conditions.

- The model predictions were found to be accurate as evidenced by the high correlation coefficients and relatively lower error values, when validated against a set of unseen data. More specifically, the ANN- and LGP-based models yielded: *R* of 0.79, 0.81; *RMSE* of 7.54, 6.80 (blows/300 mm); and *MAE* of 5.59, 4.74 (blows/300 mm), respectively. However, it was also observed that the LGP predictions were slightly superior to those from the ANN model.

- It was further verified that the optimal LGP model yielded accurate predictions when subjected to several additional measures recommended in literature:

- By means of the scatter plots of model predictions versus the measured DCP values with respect to the data in the testing and validation sets, it was found that the results are modestly scattered about the line of equality. The dispersion was considered to be classic heteroscedastic and reasonable given the uncertainties associated with the input DCP dataset and the method itself.

- The optimal model predictions were also graphically compared by means of frequency histograms with the measured final DCP values with respect to the testing and validation datasets. It was observed that both these models have a strong predictive capability and generalisation performance. However, it was also observed that the LGP model was slightly superior to the ANN model.

- In addition to the above assessments of the accuracy of the optimal models, their validity and generalisation capability were further confirmed by means of a parametric study. The model predictions were found to be consistent with the known underlying physical behaviour, as observed from field trials and thus, it was confirmed that the

predictions of these optimal AI models are robust and reliable within the specified ranges of the input variables.

- By analysing the parametric study results given by the optimal ANN model, the effect of different soil types on the model predicted output, i.e. final DCP count (blows/300 mm), was investigated. It was observed that the conclusions drawn from the optimal model predictions are consistent with conventional wisdom.

- By means of a sensitivity analysis, the contribution of each input variable on the final model predictions was examined and it was observed that the input variables utilised in this study are highly significant with respect to the predictions of the final DCP count resulting from the application of RDC. In addition, both the ANN and LGP models confirmed that the contribution made by the initial DCP count is more significant than the other input variables on the model predictions.

- Finally, the optimal models were again presented in a format that facilitates their application in practice. In particular, the optimal ANN model was translated into a series of numerical equations and the optimal LGP program was provided in C code format.

8.2.3 Recommended Optimal Models

From this study, the applicability of ANN and LGP techniques for predicting the effectiveness of RDC in different ground conditions with respect to CPT and DCP data has been investigated. Both these AI techniques possess several merits over other conventional approaches. As discussed in Chapter 3, several research studies have already demonstrated that they outperform the most commonly used traditional methods. The work presented in this thesis has also demonstrated that the application

of ANNs and LGP in relation to RDC is entirely appropriate and useful, given that no other predictive tools, conventional or AI-based, are available. The developed models are considered to be valuable additional tools for the necessary prior estimations during the pre-planning and pre-design phases in ground improvement projects involving RDC. However, as with all AI-based models, their accuracy can be improved by incorporating more data from additional RDC-related projects. This can be readily undertaken as additional data become available in the future.

Whilst the developed ANN models were demonstrated to be accurate and reliable, comparatively, the LGP-based models yielded marginally superior predictive capability. In terms of transparency, the ANN model provides comparatively a more transparent numerical formulation than that of the LGP-based model, which is translated into C code, which is more difficult to interpret. This could be attributed to the fact that LGP penalises complex models to a greater extent. However, in this study, it has been demonstrated that ANNs, particularly backpropagation MLPs, provide information on the functional relationship between input and outputs in the form of a numerical equation. Hence, this emphasises that ANNs need not be treated as *black box* models, which is a criticism that they often receive. In addition to these conclusions, it has been reported in literature that LGP approach overcomes many of the shortcomings associated with ANNs and other conventional modelling approaches. Given the slightly superior performance of the LGP-based models, it is recommended that they be adopted in preference to their ANN-based counterparts.

8.3 GUIDELINES FOR ANN MODEL DEVELOPMENT

The ANN model development considered in this research employed the feed forward type MLPs trained with the back-propagation algorithm, which is also the most common form used in prediction and forecasting applications. According to the results discussed above, the optimal models were found to be both accurate and robust. The feasibility of the ANN approach was further confirmed, particularly the MLPs trained with the back-propagation algorithm for predictions of geotechnical engineering variables. Nonetheless, the potential of ANNs for modelling different aspects of

geotechnical engineering has already been demonstrated by several researchers, as discussed earlier in Chapter 3. However, based on the success and superior performance of ANNs in comparison to most of traditional methods, ANNs are highly suitable for modelling complex systems and likely to see greater application in the future.

This section provides a comprehensive set of guidelines to assist in developing ANN models. In particular, the methods presented below are applicable for modelling the feed forward type of MLPs using the error back-propagation algorithm. The major steps in ANN model development outlined by Maier et al. (2010), as illustrated previously in Figure 3.3, are presented and re-examined with respect to the currently available approaches. However, it is not the intention to review all the available options in detail in this regard, but to highlight the most common approaches that are currently available. Finally, a discussion of the appropriateness and limitations of these methods is presented.

8.3.1 Selection of Model Inputs

The selection of an appropriate set of model input variables is the initial stage of the ANN model development process and is considered to be one of the most important steps. Essentially, appropriate data representation has a direct impact in developing successful ANN models (Basheer and Hajmeer, 2000). Unlike statistical modelling, ANNs, which are data driven, does not require a *priori* rationalisation about the input/output mapping (Lachtermacher and Fuller, 1994) and instead, it has the capability to determine the critical inputs (Maier and Dandy, 2000). However, both the inclusion of too many variables and the exclusion of significant model variables from the input set have undesirable effects on model performance. As such, the presence of too many variables in the input set sometimes provides redundant information, as these variables are generally correlated with one another. Moreover, the network becomes complex with an increase in the network size resulting in a decrease in processing speed. At the same time, the number of connection weights in the network increases and consequently, the networks with redundant inputs become susceptible to

overfitting of the training data. Moreover, inclusion of too many variables in the input set introduces additional local minima in the error surface in weight space, which in turn makes it difficult to reach the global minimum (Maier et al., 2010). On the other hand, use of fewer input parameters may be preferable since it effectively reduces the network size. However, omission of significant variables from the input set causes the inadequacy of knowledge presented to ANNs, which makes it difficult to capture the exact input-output relationship. Subsequently, the selection of an appropriate and adequate set of input variables is of prime importance.

The currently available techniques used for selecting an appropriate set of input variables can be categorised as model-based and model free approaches. One of the model based approaches is the selection of model inputs in an ad-hoc manner, where the modeller chooses several combinations of input parameters to be tested. A number of models are developed with these different input combinations and the results are compared when selecting the network with the best performance. This approach has already been employed in ANN modelling of several geotechnical engineering aspects by researchers [e.g. Goh (1994), Najjar et al. (1996)]. Another model-based approach is to identify the optimal combination of input variables in a stepwise manner, which could be either a constructive or pruning process (Maier and Dandy, 2000). If the constructive stepwise approach is used, several networks are developed, each using only a single potential input variable. The best performing network is selected and then retrained by combining another additional input variable. This process is repeated until no further significant improvement in model performance results from the addition of any extra variables. In addition to this stepwise method, there is another model-based approach for input selection, where a model is developed with a relatively large number of input variables. Later, the relative importance of the variables is investigated by means of a sensitivity analysis to determine the insignificant variables. Apart from the methods discussed above, global optimisation algorithms, such as genetic algorithms, can also be used for the selection of the input set that optimises the ANN model performance.

The use of the model-based approaches discussed above has several drawbacks. They have been found to be time consuming as they require the development of a number of ANNs with different combinations of potential input variables in order to determine the optimal combination. Moreover, as the name suggests, model-based approaches are influenced by the ANN model itself. As such, they are affected by the ANN model structure (i.e. number of hidden nodes) and also by the optimisation method, which is a function of several internal parameters, such as learning rate and the momentum term. Consequently, identification and isolation of the effect of the different input combinations is difficult (Maier et al., 2010).

Conversely, several model free methods are also available for the selection of the appropriate input set. In model free approaches, the input selection does not depend upon the performance of the developed ANN model. These methods are based either on analytical approaches or ad-hoc methods. When analytical methods are employed, a statistical measure of the significance of the input variables, such as correlation or mutual information, is used to assess the model input-output relationship, whereby the most appropriate set of input variables is then determined (Maier et al., 2010). However, the most common approach of selection of data inputs in the geotechnical engineering discipline is simply the ad-hoc method based on the prior knowledge on system. In this case a set of parameters is selected in advance and assumed to be the most influential on the desired model output (Maier and Dandy, 2000).

The present study has also employed the ad-hoc, model free approach. Here, the factors affecting ground improvement resulting from RDC was considered to be well understood and thus, the model input variables were defined in such a way that they include these parameters. It has been considered that during impact rolling, the degree of soil compaction depends upon a number of factors, including the inherent physical properties of soil, the number of times the module passes over the site, the weight and characteristics of the machinery, and the soil moisture at the time of compaction. Therefore, this knowledge has been incorporated in defining the model input variables. In general, it can be considered that the effective parameters are relatively few in number in the case of this geotechnical engineering application and therefore, the

model input variables can be defined in advance. Given that, it is recommended to explore the literature, which has contributed significantly to the understanding of the important factors that control the problem at hand and thereby, to identify the most significant parameters in advance.

8.3.2 Data Division

Having determined the appropriate model inputs and output variables, the ANN model development process requires partitioning the dataset into subsets. However, the method used for data division is usually affected by the amount of records available in the data set. It is a common practice to divide the dataset into 2 subsets, i.e. training set and the validation set. The training set is used to train and calibrate the model and it is with this data subset that the model's connection weights are optimised. Subsequent to the model calibration phase, the validation dataset is used to validate the performance of the model, whereby its performance is assessed with respect to an unseen data set (Twomey and Smith, 1997). Usually, two thirds of the overall data set is allocated to training and the remaining data are used for validation (Hammerstrom, 1993). However, in the presence of a limited dataset, the implementation of data in the form of subsets might be difficult. As such, it would not be possible to assemble a comprehensive validation set aside from the training set. In such instances, several approaches can be used to maximise the utilisation of the available data as follows:

- Use the holdout method (Masters, 1993), where the model should be trained withholding a small subset for validation. Once the model is validated, another different subset of data should be held back and the model trained again. This process should be repeated until the generalisation ability has been tested for the whole set of available data; or

- implementation of another synthetic dataset for model validation, which possesses similar statistical properties to the training dataset (Lachtermacher and Fuller, 1994); or

- as suggested by Maier (1995), a trial should be conducted by using a small data subset as a testing set, whereby the time duration required by the model to achieve an acceptable generalisation ability is determined. Thereafter, the network should be retrained with the whole dataset for the predefined period.

However, the modeller should consider using the cross-validation technique, whenever possible in ANN modelling. Cross-validation (Stone, 1974) is a frequently used technique, which greatly affects the way the available data are divided. However, this approach is only a modification of the above discussed data division method that partitions the dataset into only 2 subsets. As such, in cross-validation, the training subset should be subjected to a further partitioning, so as to construct 3 data subsets; i.e. training, testing and validation. The model performance should be assessed periodically with validation. In such a way, the testing set is used to determine the termination of model calibration and to ensure that the model does not become overtrained. Although, cross-validation is a more efficient data division method, it should be acknowledged that this method is very data intensive since the dataset needs to be divided into 3 comprehensive subsets. In the case studies presented in this thesis, data division has been carried out in such a way that the training set contains 80% of the data and the remaining 20% are used for validation purposes. The training set has been further subdivided into two subsets; 80% for training and 20% for testing.

As with empirical models, ANNs are considered to be an interpolation technique and do not perform well when they are used to extrapolate beyond the boundaries of training data set (Flood and Kartam, 1994; Minns and Hall, 1996). Consequently, the training dataset, that is used for model calibration, should essentially include all the available patterns contained in the parent database. On the other hand, as the testing subset is used for determining the termination of model training, and/or for selecting the optimal network geometry, it needs to be distinct but should also contain all the available patterns of the parent dataset. Therefore, the testing set should be representative of the training subset. In addition, the most challenging assessment of the generalisation ability of the ANN model could be possible only if the validation

subset contains all the patterns that constitute the training subset. Consequently, it is recommended that the 3 subsets should represent the same statistical population, which is evidenced by similar statistical parameters, i.e. mean, standard deviation, minimum, maximum and range (Masters, 1993). Several studies [e.g. Shahin et al. (2002)], including the present study, have confirmed that there is a direct relationship between the statistical consistency of the subsets and the model performance with respect to each of those subsets, which in turn ensures that the best possible network has been developed given the available data. Based on these considerations, it is recommended to employ a systematic approach for data division that ensures statistical consistency between the subsets. Some of the approaches that can be used are:

- the SOM method, as proposed by Bowden et al. (2002), which should be used for data clustering to identify the similarities associated with the available raw dataset. The data records from each cluster should be allocated to each of the training, testing and validation subsets so that these 3 subsets are similar to each other, whilst statistical consistency is ensured. However, the determination of the optimal SOM size is a challenge since there is no precise rule to select the most favourable map size. Therefore, as conducted in this research, it is suggested to explore several map sizes, such as 8×8 and 10×10, in order to ensure that the optimal map size is neither too large nor too small; or

- the trial-and-error approach, where statistical consistency between the subsets is achieved by manual adjustments to the composition of the subsets. This approach has been commonly employed in geotechnical engineering aspects [e.g. Pooya Nejad et al. (2009)]; or

- by using a formal optimisation approach, such as genetic algorithms, to optimise the similarity of the statistical parameters between subsets; or

- by using the fuzzy clustering method. A detailed procedure of this approach is given by Shahin (2003).

8.3.3 Data Normalisation

It is essential to pre-process the data prior to presenting them to the ANN model. Data normalisation, or scaling, is the most common form of pre-processing, whereas other forms, such as data transformation, are not necessarily required. Thus, scaling the data into a range that is appropriate to the transfer function used at the output layer is recommended. For an example, if a sigmoidal type transfer function, such as the logistic or hyperbolic tangent, is adopted at the output layer, the scaled data should be commensurate within the limits of 0 to 1 and –1 to 1, respectively. However, it is further recommended to normalise the data slightly offset from these actual limits. For instance, the range of 0.1 to 0.9 or 0.2 to 0.8, is often used when the logistic sigmoidal transfer function is adopted because otherwise the weight updates may be too small during training and may slow down the training speed (Masters, 1993; Hassoun, 1995; Maier and Dandy, 2001). Scaling the data to a uniform range emphasises that each variable receives similar attention during the training process and prevents larger numbers from overriding smaller values (Basheer and Hajmeer, 2000; Maier and Dandy, 2000). The following equation has been used to scale the data in the case studies presented in this thesis:

$$x_{scaled} = a + (x_{unscaled} - A)\frac{(b-a)}{(B-A)} \qquad (8.1)$$

Where, A and B are the minimum and maximum values of the unscaled dataset, respectively. Similarly, a and b are the minimum and maximum values of the scaled dataset, respectively.

In addition to the linear normalisation approach presented above, several other studies [e.g. Masters (1993), Dowla and Rogers (1995) and Swingler (1996)] have proposed other techniques for data normalisation. However, these techniques are

computationally intensive and on the other hand, Masters (1993) indicated that they do not produce substantially better results than the linear normalisation approach.

8.3.4 Selection of Model Structure

The following treatment relates to the feed forward type MLP, which is the ANN architecture employed in the present study, which is the most popular architecture adopted by researchers, and has also been found to be successful in a large number of geotechnical engineering applications. A more detailed description on MLPs was provided earlier in Chapter 3.

Having selected the ANN architecture, i.e. MLP, the next step is to define the network structure. The combined form of network architecture and its structure defines the functional relationship between the input and output variables (Maier et al., 2010). However, the optimisation of the network structure is a crucial aspect of ANN modelling, which involves the selection of the number of layers and the appropriate number of nodes contained in each layer. The number of nodes in the input and output layers of the network are restricted by the number of model input and output variables, respectively, as specified by the modeller. Thus, the model structure selection reduces to the selection of the number of hidden layers and the number of hidden nodes in each layer, which also defines the number of connection weights incorporated in the network.

It is essential that the intermediate hidden layer(s) contain a sufficient number of hidden nodes to capture the complexity and non-linearity of the system. However, it is considered that the optimal ANN structure generally strikes a balance between network complexity and its generalisation capabilities (Maier et al., 2010). It is important that the structure is neither too complex nor too simple but adequately captures the nuances included in the training data. When the network structure is too simple, the network might not be able to represent the relationship between variables adequately and thus, predictive performance is compromised (Maier et al., 2010). On the other hand, an excessively complex network with too many hidden nodes may

overfit the data and thus lack the ability to generalise (Maren et al., 1990; Masters, 1993; Rojas, 1996) and also may have a decreased processing speed and diminished transparency (Shahin et al., 2008).

Determination of number of layers

A single hidden layer network with sufficient connection weights will be capable of approximating any continuous function (Cybenko, 1989; Hornik et al., 1989). As discussed in Chapter 3, several geotechnical engineering applications, including the present study, have yielded reliable predictions using the simplest form of MLP that consists of 3 layers, including one hidden layer comprised of an adequate number of hidden nodes. Conversely, there have been suggestions to use more than a single hidden layer to provide more flexibility to the network to capture complex functions (Flood and Kartam, 1994; Ripley, 1994; Sarle, 1994). Masters (1993) stated, however, that multi-hidden layer networks often possess a slow training speed and are also more susceptible to getting trapped in some local minima. Nevertheless, it is emphasised that the selection of an optimal network structure is highly problem-dependent. Therefore, it is recommended to begin the modelling process by using a single hidden layer, which incorporates an adequate number of hidden nodes. However, it will also be useful to examine the networks with multi-hidden layers with the intention of improving prediction accuracy. On the other hand, it is also important to note that several previous MLP applications have experienced difficulties in approximating the exact functional relationship between the variables, as well as a lack of robustness, in the presence of a single hidden layer, and in these situations multi-hidden layer networks need to be adopted [e.g. Pooya Nejad et al. (2009)]. Thus, the selected optimal model should always be tested using a parametric study to ensure robustness. If model performance is unsatisfactory, further modelling should be conducted to determine whether multi-hidden layers provides a solution.

Determination of number of nodes

The number of nodes in the input and output layers is again defined by the number of input and output variables used in the model. Hence, the most critical aspect is the selection of optimum size of the hidden layer that relates to the number of connection

weights in the network. Several guidelines are provided in the literature to assist with determining the optimum number of nodes in the hidden layers. For networks consisting of a single hidden layer, Salchenberger et al. (1992) suggested that the number of hidden nodes should be 75% of the number of input nodes, whilst Berke and Hajela (1993) suggested that it should be between the average and sum of the input and output nodes. In addition to these suggestions, Caudill (1988) proposed an upper bound to the number of hidden nodes of ($2I$+1), for a network with I number of input nodes. Apart from that, several empirical relationships between the number of connection weights and the number of training samples have also been suggested by several researchers [e.g. Rogers and Dowla (1994)] to ensure the development of a model with good generalisation capability. As such, it has been proposed to maintain the number of connection weights below the number of training records. Additionally, the ratio of the number of training records to the number of connection weights in the network has been suggested to be 2:1 (Masters, 1993) or 10:1 [e.g. Weigend et al. (1990)]. Furthermore, Amari et al. (1997) stated that the networks trained with a number of samples of at least 30 times of the number of connection weights do not tend to have an issue of overfitting. However, apart from the above guidelines, it should be noted that the network structure is generally highly problem-dependent. Therefore, it is recommended not to restrict modelling in accordance with the above guidelines but to consider exploring the optimal network structure, which is defined by the smallest possible network that adequately captures the relationship in the training data (Maier and Dandy, 2000).

For identifying the optimal network structure, stepwise approaches, which could be either a constructive or pruning method, can be used. If the constructive method is used, modelling should be started with a small network, to which nodes and connections should be added successively. The resulting models should be calibrated and the process should be continued until no further improvement in model performance is likely to occur. With the pruning algorithm, the same procedure should be applied in the converse direction, where one begins with a sufficiently large network and unnecessary nodes and connection weights are removed successively. Other approaches of determination of number of nodes in ANN models include global

methods, such as genetic algorithms, and simulated annealing, which have been rarely adopted in geotechnical engineering applications. However, the usual practice of identifying the optimal network structure is to employ an ad-hoc trial-and-error approach. Therefore, several models should be tested with different numbers of hidden nodes to develop a network yielding satisfactory performance. However, it is important whenever possible to execute the guidelines discussed above, especially with regards to the upper limit of the number of hidden nodes and also the relationships between the number of training samples and the size of the hidden layer. For instance, in the case study of ANN modelling with CPT data, several ANN models were trained, starting with the smallest possible network involving a single hidden node and successively increasing the number of hidden nodes to a maximum of 9, where 9 nodes were considered to be the maximum number required for the models with 4 input nodes, as suggested by Caudill (1988).

8.3.5 Model Calibration

The objective of model calibration is to investigate the optimal set of connection weights that permits the ANN model in the given functional form to best represent the desired input/output mapping (Maier et al., 2010). During model calibration, a suitable error measure between the actual and predicted output is minimised and accordingly, the set of connection weights are adjusted so that the network can efficiently produce the best fit to the training set data. However, the determination of the optimal weight combination that minimises the error function is difficult. The different connection weight combinations produce different error surfaces. Unlike a smooth error surface, more rugged error surfaces are complex with many local minima and thus, achieving the global optimum might be difficult. However, the degree of ruggedness of the error surface is again problem-dependent and may affected by many factors including the number of connection weights. Consequently, for ANN model calibration, it is important to follow a systematic approach that strikes a balance between computation cost and model performance (Parisi et al., 1996; Maier and Dandy, 2000).

There are many approaches available for model calibration, as discussed by Maier et al. (2010). However, in the present study, the MLPs were trained using the back-propagation algorithm (Rumelhart et al., 1986). It is by far the most frequently used approach for optimising feed-forward ANNs and a detailed description of its operation was provided earlier in Chapter 3. The back-propagation algorithm is a deterministic, local optimisation approach that operates based on first-order gradient information (Minns and Hall, 1996). However, model optimisation with the back-propagation algorithm is affected by the size of the steps taken in the weight space, which is generally fixed throughout training (Maier and Dandy, 2000). If the step size is too small, networks tend to reach the local minimum slowly and steadily but also become vulnerable to being trapped in a local minimum. On the other hand, if the step size is too large, the networks reach local minima quickly but at the same time they are likely to fall into oscillatory traps (Rojas, 1996; Maier and Dandy, 1998; Maier and Dandy, 2000) or to diverge completely. Therefore, it is essential to optimise the size of the steps and the common approach is to use a trial-and-error method.

The step size taken during model calibration is a function of the internal parameters, including the transfer function, epoch size, error function, learning rate and momentum term (Maier and Dandy, 1998; Maier and Dandy, 2001). Thus, the choice of internal parameters has an adverse effect on the performance of the back-propagation algorithm (Dai and MacBeth, 1997). Therefore, the modeller needs to have a good understanding of the impact of different internal parameters. In order to assist with this, a brief overview of each parameter is given below.

Transfer function

Any transfer function can be employed in ANNs trained using the back-propagation algorithm as far as they possess the distinctive properties of continuity and differentiability on $(-\infty, \infty)$ (Basheer and Hajmeer, 2000). However, the majority of the applications use sigmoid type transfer functions, such as the logistic sigmoid and hyperbolic tangent transfer functions (Maier and Dandy, 2000). The logistic sigmoid function transforms the values into the range of 0 to 1, or –1 to 1 by the tangent

hyperbolic function. Usually, the same transfer function is used for all nodes in a particular layer.

Epoch size

The number of training records presented to the network prior to each weight update is called the epoch. Depending on the epoch size, the network can be trained either in an online mode or in a batch mode. If the network is trained with the online mode, weights are updated immediately after the presentation of each training pattern, which in the case of epoch size, is equal to one. Conversely, in the batch mode, the weight updates take place after a certain number of training samples, or the entire training set, is presented to the network. In comparison, the online mode requires less storage for the weights and adopts a stochastic search path that prevents entrapment in local minima in the error surface (Hassoun, 1995). On the other hand, the batch mode is advantageous in that it can estimate the error gradient vector better (Basheer and Hajmeer, 2000). Moreover, when the epoch size is equal to the size of the training set, it forces the steps taken in the weight space to move in the direction of the true gradient at each weight update (Maier and Dandy, 2001). However, the effectiveness of both these training methods is again problem-dependent (Basheer and Hajmeer, 2000).

Error function

The error function, which accounts for the deviation of the model predictions from the corresponding target values is progressively reduced to a desired minimum during the training phase. The common approach is to employ the *MSE,* which is the mean of the squares of deviations. In addition, other error functions such as sum of squared errors (*SSE*), coefficient of determination (R^2) can also be used. Employing the *MSE* as the error function has several advantages, such as simplicity of calculations, that it penalises large errors, easy calculation of subsequent derivatives with respect to the weights and that it lies close to the heart of the normal distribution (Masters, 1993).

Learning rate

The learning rate is one of the parameters that affects the size of the steps taken in the weight space during model calibration. As a result, the learning rate affects the training

speed. If the learning rate is set too high, the resulting larger steps accelerate the training speed but also can result in large oscillations or even a complete divergence from the desired minimum (Basheer and Hajmeer, 2000). On the other hand, if the learning rate is set too small, the training speed becomes slow, however, this may result in a steady search in the direction of the global minimum. In general, the learning rate is allowed to remain constant during model calibration (Warner and Misra, 1996), whereas the optimal value of the learning rate is usually obtained from a trial-and-error procedure. Several heuristics are also available that suggest the optimal range of the learning rate; i.e. according to Wythoff (1993) it should be between 0.1 to 10, Zupan and Gasteiger (1993) suggest between 0.3 to 0.6, whereas Fu (2003) recommends the range of 0 to 1.

Momentum term

The momentum term determines the fraction of the weight update from the previous stage to the current stage of weight update. It enables the search to escape the local minima and also reduces the likelihood of search instability (Zupan and Gasteiger, 1993; Basheer and Hajmeer, 2000). As with the high values of learning rate, a high momentum can accelerate the training process by several orders of magnitude (Masters, 1993) but at the same time, it increases the risk of overshooting a near optimal weight vector (Basheer and Hajmeer, 2000). Conversely, an extremely low momentum reduces the convergence speed. Thus, an optimal level of momentum should be utilised in model calibration and a trial-and-error approach is normally preferable. However, it must be always less than 1 for the convergence (Dai and MacBeth, 1997). Wythoff (1993) recommends the momentum term to be between 0.4 to 0.9, whereas Hassoun (1995), as well as Fu (2003), suggest to adopt a value between 0 to 1.

Initial connection weights

The initial weight distribution is also a major concern in model calibration. Since the back-propagation algorithm employs a gradient descend rule to adjust the connection weights, the results are sensitive to the initial conditions. As such, it has the possibility of getting trapped in a local minimum if training is commenced from an unfavourable

position in the weight space. In general, the weights are uniformly initialised to zero-mean random numbers (Rumelhart et al., 1986; Maier and Dandy, 2000). However, the weights initialised in an extremely smaller range may cause the training to be paralysed, whereas a too large region causes premature saturation of the nodes that results in cessation of training at sub-optimal levels (Maier and Dandy, 2000). According to empirical studies, it is considered that there is no exact optimal range for which similar performance can be obtained. Therefore, it is recommended to calibrate the network several times having started at different random initial conditions and to examine whether the network performance is consistent.

8.3.6 Stopping Criteria

ANN model calibration should be terminated by using an appropriate stopping criterion, by which it is decided whether the network has been trained optimally or sub-optimally. Stopping criteria can be defined based on an error measure between the actual and predicted output, either with respect to the training set or testing set. However, when selecting the stopping criterion, it is important to consider whether overfitting is a possibility. Depending upon the available number of training records and the number of connection weights in the network, sometimes overfitting might be an issue and requires the modeller's attention for necessary precautions. Therefore, it is important to reconsider the guidelines that examines the possibility of overfitting in advance. Especially, the empirical relationship provided by Amari et al. (1997), who recommended that networks do not tend to overfit the data if the ratio of the number of training records to the number of connection weights is more than 30.

If overfitting is unlikely, model calibration can be stopped when the error function with respect to the training set data has been minimised. As such, model training should be terminated either after the training error has reduced to a sufficiently small value or when the changes in training error become negligible. On the other hand, if overfitting might be a problem, the best option is to employ the cross-validation approach (Stone, 1974), whereby model calibration can be stopped using the error with respect to testing set data. As discussed in §8.3.2, in cross-validation, the testing subset

evaluates the generalisation ability of the network from time-to-time, whilst the network is being calibrated against the training dataset. At the initial stage, the testing error will decrease similar to the training error, but when the model overfits the training data there will be an increase in the error for testing set data, as the network begins to memorise rather generalise. Therefore, model calibration should be terminated at the onset of an increase in the test data error. However, as discussed in some studies [e.g. Masters (1993), Ripley (1994)], there is no guarantee that the error will reach its lowest value just before the testing error starts to increase for the first time and it might achieve some further lower value with continued training. Therefore, it is always recommended to continue training for some time, even after the testing error starts increase for the first time (Maier and Dandy, 2000).

8.3.7 Performance Evaluation

The evaluation of network performance is required at different stages during the ANN model development process, such as for the selection of the optimal set of input variables by using the model-based approaches (as discussed in §8.3.1) and more importantly, for the selection of the optimal network structure (as discussed in §8.3.4). The performance of ANN models can be evaluated based on several factors, including prediction accuracy, training speed and the processing speed (Maier and Dandy, 2000). However, most commonly, ANNs are evaluated by using one or more measures that assess the prediction accuracy of the network.

A number of measures of prediction accuracy are available to use, as suggested in the literature. For instance, a list of commonly employed metrics has been given by Maier et al. (2010) and that includes the functions of squared errors, absolute errors, relative errors and product differences. In the case studies presented in this thesis, the networks' prediction accuracy is evaluated by using *RMSE*, *MAE* and *R*. When using the squared error functions, such as *RMSE*, the errors with high magnitude receive much greater attention than the smaller errors (Maier et al., 2010). On the other hand, the absolute errors, such as *MAE*, provide information on the magnitude of the error, but they are incapable of providing information regarding the performance of the

model in the form of overall under- or over-prediction (Maier et al., 2010). In addition, by using the product difference statistics, such as the *R* value, the relative correlation between the predicted and actual results can be effectively measured.

8.3.8 Model Validation

It is the good practice to validate the performance of the trained network once model calibration has been completed. At this stage, the generalisation capability of the network is assessed with an independent validation set, which should not have been used in any capacity during model calibration. The performance of the network with respect to the validation set data can be evaluated again using the above selected criteria (as discussed in §8.3.7). The network is expected to generate non-linear relationships between the inputs and output(s) rather than simply memorising the patterns. Since the model is assessed by an unseen data set, the results are significant for the evaluation of network performance. However, it is expected that the results obtained with respect to the validation dataset to have general consistency with respect to those obtained from the training and testing subsets. If it is not, it can be either because of the model is overfitting to the training data or sometimes the data subsets may not represent the same population (Masters, 1993). In addition, poor validation can resulted from inadequate or a lack of data pre-processing and data normalisation, or else it could be due to the selected network architecture (Maier and Dandy, 2000).

8.3.9 Parametric Study

In several previous applications of ANNs, it has been demonstrated that good ANN network performance with respect to the validation set does not guarantee robust model predictions [e.g. Pooya Nejad et al. (2009)]. Therefore, it is important to ensure the generalisation ability of the selected optimal ANN model having tested it using a parametric study that examines the conformity of the model predictions with the known physical behaviour of the modelled system. For this purpose, the approach outlined by Shahin et al. (2000) can be employed. This method investigates the response of the optimal model to variations in the input variables. If this method is

used for parametric study, the model output should be examined over a given set of synthetic input data by varying the input variables one at a time, whilst all other input variables are kept constant at a pre-defined value, i.e. the mean value of training set data. However, it is essential to vary the input variables only within the range over which the model is calibrated, as ANNs should be utilised only as an interpolation, rather an extrapolation, technique.

The results of the parametric study are critical in assessing model performance. If the selected optimal network is robust and possesses adequate generalisation capability, the model can be implemented for future predictions with confidence. On the other hand, if the model response does not conform to the underlying physical behaviour of the problem, the model predictions should not have considered to be reliable. In that case, the possible scenarios, such as model overfitting, should be investigated further.

8.4 GUIDELINES FOR LGP MODEL DEVELOPMENT

The development of a predictive model to forecast the effectiveness of RDC in different ground conditions presented in this thesis employed another particular variant of AI, namely LGP. It is evident from the results analysed in Chapters 6 and 7 that the LGP models provide accurate predictive capability and also slightly outperform the ANN models. Thereby, the feasibility of LGP approach for modelling the complex non-linear relationships between many parameters has been confirmed. In addition, it has been highlighted in the discussion that GP techniques overcome most of the difficulties associated with conventional approaches and with the ANN modelling. For instance, it is difficult to produce straightforward and practical prediction equations in ANNs, especially when the network is less parsimonious. As such, despite the fact that ANNs perform adequately in the vast majority of applications, the lack of transparency of ANN models that fails to explain the underlying physical processes of the system in question, in many cases limits the applicability of ANNs, whereas the GP-based techniques often overcome such limitations. Nonetheless, GP has already emerged as a promising approach for non-linear modelling, which further encourages future applications in the broader geotechnical engineering context.

Therefore, with the intention of providing a basic understanding and guidance for readers interested in LGP, the following discusses briefly the fundamental theory of the evolutionary algorithm of LGP. It is considered that the knowledge of the underlying algorithm facilitates improved modelling capability. Later in this section, a description is provided of the primary steps involved in developing LGP models. Details are provided regarding the measures and the control parameters that the modeller should specify prior to a GP run. However, it is beyond the scope of this treatment to address all aspects of LGP.

8.4.1 Executional Steps of LGP Algorithm

The essential executional steps of the LGP evolutionary algorithm have been summarised previously in Chapter 3, whereas a comprehensive description is readily available in the literature [e.g. Brameier and Banzhaf (2007), Koza and Poli (2005), Alavi et al. (2013)]. However, in the following treatment, a series of well-defined, problem-independent executional steps of the LGP algorithm is presented.

In GP, a potential solution (computer program) to the problem in question is referred to as an individual. GP performs a multi-directional simultaneous search for the optimal solution from a pool of many individuals, collectively known as a 'population'. However, during the LGP evolutionary process, having been subjected to a fitness evaluation, each individual in a population is assessed from time-to-time in terms of its capability of performing in the particular problem environment. The fitness measure specifies the desired goal of the search process (Poli et al., 2007) and typically depends on the nature of the problem. More details about the fitness measure is discussed later in this chapter.

LGP begins with an initial population of individuals (computer programs) that are composed of randomly chosen functions and terminals appropriate to the problem domain. Typically, these individual programs in the initial population emerge with a diversity in fitness, and as such some of the individuals inherit an exceedingly poor fitness, whilst other individuals appear to be much fitter. Therefore, the creation of the

initial population is considered to be a blind random search for a possible solution in the search space of the problem in question (Koza, 1994). Nonetheless, the initial population yields the baseline for judging future search efforts (Koza, 1994). However, the differences in the fitness of the individuals are exploited in the course of program evolution.

After the fitness of each individual in the population is ascertained, several individuals are probabilistically selected in order to be subjected to the genetic operations. The selection methods are fitness based, such that the individuals that are relatively higher in fitness are more likely to be selected (Selle and Muttil, 2011). Among the several different methods of selecting the individuals, tournament selection and fitness-proportionate selection methods are the most common. In both these methods, it is considered that the selection is not greedy, where the individuals that are considered to be inferior are also selected to a certain degree (Koza and Poli, 2005). Moreover, it cannot be guaranteed that the individual with the highest fitness is also necessarily selected and on the other hand, the individual having the worst fitness is excluded. Throughout the case studies presented in this thesis, the tournament selection method has been employed.

In the standard LGP evolutionary algorithm discussed here, two tournaments take place in parallel, where the most fit individual is selected as the winner of each tournament. The resulting two winning programs turn out to be the parental programs that undergo the genetic operations. Nevertheless, prior to the transformation of the selected winning programs, they reproduce within the population at a certain probability and thereby, the copies of tournament winning programs replace the corresponding tournament losing programs. Therefore, with the full reproduction of the winning programs, it can be ensured that the fittest individuals, i.e. better solutions, always survive within the population (Brameier and Banzhaf, 2007).

In the next executional step in the LGP algorithm, the selected individual programs are modified subject to the variation operators, again with certain probabilities. The variation operators change the contents and the size of the individuals (Aminian et al.,

2013; Gandomi et al., 2014) and thereby, the chosen parental programs are transformed into new offspring programs that are more effective in solving the problem. The variation operators used for the LGP evolution include crossover and mutation.

Crossover operates by exchanging sequences of instructions between two selected parent programs and creates two new offspring programs. Crossover can be either homologous or non-homologous according to the selection of the position and the length of the instruction blocks to be exchanged (Gandomi et al., 2010a). In homologous crossover, a segment of a parent program with an arbitrary length is selected at a random position and swapped with an equally sized segment starting from the same position in the other parent program (Francone, 2010). In contrast, non-homologous crossover exchanges program segments in the two parent programs with no reference to the length or position (Rashed et al., 2012). As such, segments of random position and arbitrary length in each of the parent programs are exchanged and thereby, non-homologous crossover can change the length of either or both of the parent programs (Francone, 2010). However, when either of the resulting offspring programs exceeds the maximum allowable length, the crossover operation is terminated abruptly but resumed with swapping equally sized segments (Brameier and Banzhaf, 2001; Alavi and Sadrossadat, 2016).

Crossover is the standard macro-operator since it can alter the length of programs only at the level of instructions but not inside the instructions (Brameier and Banzhaf, 2007). Conversely, the variation operation of mutation causes random changes in the selected parent programs in 3 different ways, i.e. block mutation, instruction mutation or data mutation (Gandomi et al., 2010a). In block mutation, an entire instruction block in a program is replaced by a new randomly generated instruction block. In instruction mutation, an existing instruction in a program is replaced with a new, randomly chosen instruction of the same length whereas, data mutation modifies an existing instruction by replacing one of the terminals with a new randomly selected terminal (Francone, 2010).

Once, after the above described operations of selection, reproduction and variation are applied to the individuals in the current population, a new population is created with the offspring programs. In a steady-state LGP algorithm that constitutes the executional steps described so far, the newly created offspring programs replace the existing individuals in the same population, unlike a generational evolutionary algorithm, where the offspring programs migrate to a separate pool of population (Brameier and Banzhaf, 2007). However, each new individual is measured for fitness, and these steps are repeated many times until a run termination criterion is satisfied. In a multiple-run system, a number of runs are executed iteratively and the best individual ever encountered within the LGP project is typically designated as the optimal or near optimal solution for the problem in question.

8.4.2 Preparatory Steps

As discussed in earlier chapters, an evolutionary algorithm like LGP, requires minimal user input – only to specify the problem requirements – whereas, subsequently, the computer programs automatically evolve to an optimal/near optimal solution for the problem in question in accordance with the executional steps discussed above. Therefore, as described by Koza and Poli (2005), it is important that the modeller implements certain, well defined steps in preparing to use GP, through which the high-level statement of the problem is communicated to the algorithm. These major preparatory steps have been summarised previously in Chapter 3. Briefly, prior to the evolutionary process, the modeller should specify: the function and terminal set; the fitness measure; certain parameters that control a GP run; and lastly the termination criterion. Below, these preparatory steps are discussed in detail with the aim of assisting the modeller develop robust LGP models.

The configuration of the GP system fundamentally requires the modeller to specify the function and terminal set appropriate to the particular problem domain (Koza, 1994). An individual program in LGP, which is generally interpreted as a sequence of instructions, is formed by the combination of function and terminal set. The function set (instruction set) of the system may consist of arithmetic functions (+, –, ×, /), any

mathematical function (e.g. *sin*, *cos*, *exponential*, *square root*), Boolean logic operators (AND, OR, NOT), conditionals (IF, THEN, ELSE), iterative functions (DO, CONTINUE, UNTIL) and/or any other user-defined functions. Typically, the composition of the function set is driven by the nature of the problem domain (Koza and Poli, 2005) and the choice of the function set has a significant effect on the complexity of the evolved program (Gandomi et al., 2010a). For instance, the function set that consists merely of arithmetic functions may solve simple numerical problems, whereas inclusion of mathematical functions permits the evolution of highly non-linear models (Mehr et al., 2014). In fact, the assignment of the function set requires careful attention since the capability of GP to find an optimal solution depends heavily on the composition of the function set that determines the expressiveness of the GP (Brameier and Banzhaf, 2007). On the other hand, the choice of terminal set is straightforward to the problem. Terminals contain the arguments for the functions, such that the model independent variables (external inputs) and numerical constants together form the terminal set (Brameier and Banzhaf, 2001; Gandomi et al., 2011).

As with the above two preparatory steps, the modeller specifies the primitive set for LGP, by which the search space of the problem is indirectly defined for the GP system to explore (Mehr et al., 2013). The search space may consist of possible solutions composed of random combinations of the primitive set. However, at this stage, there remains a lack of information regarding which individuals or regions in the search space are more appropriate in solving or approximately solving the problem. Thus, it is necessary to assign the fitness measure that specifies the desired goal of the search.

The fitness measure is often considered as the sole mechanism of communicating the problem requirements to the GP system (Poli et al., 2007). The measurement of the fitness of programs depends on the nature of the problem (Koza, 1994). Usual practice is to derive the fitness measure based on a mapping error between the predicted and the desired output. In that case, the closer the fitness value is to zero, the better the program. On the other hand, in a problem of optimal control, the fitness can be derived by the amount of time, or any other suitable quantity required to bring the system to a desired state. As such, the lower the amount or quantity, the higher the fitness.

However, when it is required to recognise patterns or to classify the samples, the fitness can be measured in terms of the number of correct instances. In such cases, the higher the number of instances that correctly recognise or classify the samples, the better the program. In addition to the above, the fitness measure can be multi-objective so that it accounts for a combination of factors, such as model correctness, parsimony or efficiency (Koza, 1994). Likewise, depending on the problem, the modeller should specify a suitable fitness measure that evaluates the individuals in a population in terms of the capability of performing the task at hand.

Typically, throughout the LGP run, the individual programs in a population are evaluated over a number of different fitness cases, with the fitness being measured as the sum or an average over a set of different input-output pairs. For example, this study has utilised the *MSE* as the fitness function, which calculates the squared difference between the predicted output and the actual output averaged over a number of training samples.

The fourth major step in developing a GP model entails the specification of control parameters. In the LGP algorithm, several different parameters, such as population size, number of demes, program size and frequencies of genetic operations, are involved. Details regarding these parameters and their influence on the subsequent model, are discussed briefly below. However, it is considered that the parameter selection affects the generalisation capability of the model (Alavi and Gandomi, 2012; Gandomi et al., 2014) and thus, careful attention should be given to the selection. It is recommended that the modeller conduct several runs with different parameter combinations to obtain the parameterisation of the LGP that provides satisfactory robustness and generalisation capability.

Population size

The LGP individuals evolve through a series of generations, where the number of individual programs in a generation is defined by the population size. As the initial population consists of randomly generated individuals, some of the programs may have exceedingly poor fitness, whilst others will be fitter. However, as the evolution

progresses, new offspring populations emerge with an increased chance of evolving an acceptable solution. Therefore, it is important to permit an appropriate size for the population and that allows a sufficient search in the problem space. However, a proper population size is dependent upon the number of possible solutions and the complexity of the problem (Gandomi et al., 2014). Although the population size parameter has no upper limit, the run time will be longer for a large population. However, on the other hand, a larger population expands the exploration of the search space and increases the likelihood of evolving to the solution (Walker, 2001). Nonetheless, it is necessary to assign an adequate size for the population in order to capture the complexity of the problem. However, the practical limitations of the available computer capacity should also be considered when specifying the population size. In the case studies conducted in this thesis, several population sizes were attempted ranging from 500 to 10,000 individuals.

Demes

Brameier and Banzhaf (2007) stated that tournament selection together with the steady-state LGP evolutionary algorithm is well suited for parallelisation. Thus, it is usually recommended to subdivide the population of individuals into multiple subpopulations, named demes. Therefore, the individuals could migrate between subpopulations causing evolution to occur in the population as a whole (Gandomi et al., 2010a). The allocation of demes is always beneficial, since it has been found that evolution progresses faster in these semi-isolated subpopulations in comparison to an equally sized single population (Alavi and Gandomi, 2012; Rashed et al., 2012). This may be attributed to the fact that the multiple subpopulations with restricted migration may enhance genetic diversity (Brameier and Banzhaf, 2001). However, when specifying the number of demes, it is important that each deme is allocated enough programs to engage in useful evolution.

Program size

As discussed earlier, in LGP, the initial population is composed of randomly generated individuals. However, it is important that the sizes of these individuals are controlled. The lower bound of the program length may be assigned to be equal to the absolute

minimum length of a program, i.e. one instruction. In addition, it is necessary to specify the maximum size of the initial program length. As such, the upper limit of the length of a program evolved in the initial population is defined by the initial program size parameter. Accordingly, the individuals in the initial population are generated with the sizes chosen within this predefined range with a uniform probability. However, when specifying the sizes for the initial program length, there is a trade-off. Initialising an exceedingly lengthy program is not recommended since this may reduce the variability of the programs during evolution. Conversely, initialising smaller size programs allows a better exploration of the search space at the initial stage (Brameier and Banzhaf, 2007).

As the evolutionary process continues, the programs are likely to evolve in different sizes, as non-homologous crossover is allowed in LGP, which permits the programs to change their length. However, it is necessary to specify an upper bound for the program length to avoid unnecessary growths (Azamathulla et al., 2011). Therefore, the upper limit of the program length in any other population, except the initial population, is specified by the maximum program size parameter (Aminian et al., 2013).

It is apparent from other previous LGP applications, that the LGP algorithm is more likely to find the optimal solution with increased program sizes (Alavi and Gandomi, 2012; Rashed et al., 2012; Aminian et al., 2013). However, at the same time, exceedingly lengthy programs are more susceptible to becoming complex and thus, decreases the convergence speed. Thus, the trade-off between running time and the complexity of the evolved programs should be taken into consideration when specifying the program size parameters (Gandomi et al., 2014). For example, in this study, the initial program size was tested at 80 bytes, whilst the maximum program size was tested at two optimal levels of 256 and 512 bytes.

Crossover and mutation rate

As discussed, in the LGP algorithm, new offspring populations are created having the selected programs from the current population and are subjected to the genetic operations, i.e. crossover and mutation. The frequency parameters of crossover and

mutation define the overall probabilities of the corresponding genetic operations that take place with the tournament winning programs (Koza, 1992; Aminian et al., 2013). It can be commonly seen in many LGP applications that both these frequencies are regulated and tested at two different levels, such as 50% and 95%. Although the mutation frequency of 95% may seem high, an improved generalisation capability has been observed in several applications (Banzhaf et al., 1996; Brameier and Banzhaf, 2001; Gandomi et al., 2010a). This may be attributed to the fact that variable exchanges resulting from the mutation operator could have a significant effect on data flow and its fitness (Brameier and Banzhaf, 2001).

The final preparatory step requires the modeller to specify the criterion for terminating a LGP run, along with the method for designating the result of a run. The run termination criterion can be defined based on the number of generations. For instance, a run can be terminated either after a predefined sufficient number of generations have evolved without any improvement in fitness or otherwise after a pre-specified maximum number of generations have evolved since the beginning. On the other hand, run termination can be defined based on a problem specific success measure (Koza and Poli, 2005). As such, when the fitness measure achieves a predefined satisfactory level, a run can be terminated, given that the individuals have evolved sufficiently. Finally, the best individual ever encountered during a run is designated as the 'winner' of that run.

However, during a LGP run, it is important that the generalisation ability of the individual solutions is continuously monitored with the intention of avoiding overfitting, which is a common problem for machine learning methods. As mentioned previously, model overfitting can be observed when the individuals perform well on the training data but have a comparatively poor performance in the presence of a new dataset. Thus, to prevent overfitting, it is necessary to employ a suitable data division approach, similar to that adopted in ANN modelling, which involves the division of the available dataset into 3 subsets: training, testing and validation. The training process, or genetic evolution, is conducted with respect to the training dataset, whilst the generalisation capability of the evolved programs is examined during the model

training phase with respect to the testing subset. Thereby, the program with the best performance, with respect to both the training and testing data subsets, should be selected as the best evolved program. After evolution, the selected program is validated against the independent validation set, which plays no role in developing the LGP model. However, it is recommended that the modeller employs a statistically consistent data division approach for the similar reasons discussed previously in ANN modelling.

8.5 SUMMARY

This chapter has summarised the details of the optimal AI models developed in this thesis. It is evident that the two distinct sets of optimal models: two involving ANNs – one for the CPT and the other for the DCP; and two LGP models – again, one for the CPT and the other for the DCP – are successful in providing reliable predictions of the effectiveness of RDC in various ground conditions. In addition, it has been demonstrated that the LGP-based models slightly outperform their ANN counterparts and overall produce slightly more accurate predictions. Thus, considering many other factors when compared to ANNs, the value of the LGP-based models has been emphasised over ANNs in relation to RDC.

Furthermore, based on the success and usefulness of these two AI applications in relation to RDC, similar future applications are recommended. Therefore, in this chapter, a set of guidelines were provided in regards to each the AI techniques, i.e. ANN and LGP, in order to assist modellers. Step-by-step guidance of the modelling methodology, along with the currently available approaches and their feasibility, has also been presented.

8.6 RECOMMENDATIONS FOR FUTURE WORK

The research presented in this thesis has focused on predicting the effectiveness of RDC by means of two AI techniques in the form of ANNs and LGP. It is important to note that no such predictive models exist for RDC, neither empirical, theoretical nor numerical. This study is therefore the first to investigate the applicability of AI methods for the prediction of the effectiveness of RDC. As discussed above, the results of the selected optimal models are pleasing. Nevertheless, there remains room for improvement and modifications, which could potentially provide fruitful research opportunities in the future. Several recommendations for future work are described below.

As with all other empirical methods, the applicability of the ANN- and LGP-based models are constrained to the range of the input variables used during the model calibration phase. For instance, the DCP database consists of measurements where the input variable of average depth is limited to 1.95 m, so that the optimal DCP models are restricted to the predictions above depths of 1.95 m. On the other hand, when the CPT models are considered, the model predictions are satisfactory only with regards to soils with values of $q_{ci} < 10$ MPa. This is because the models are not well calibrated for the higher values of q_{ci} due to the paucity of such data in the existing CPT database. However, the applicability and the accuracy of the developed models can be further enhanced by incorporating more data from additional RDC-related projects that may become available in the future. Consequently, it is desirable to include a dataset, where the input parameters span a wider range than the existing dataset.

Furthermore, in relation to the development of DCP models, the soil type has been defined in a more generalised manner using a primary (dominant) and secondary soil type for each DCP record. With the availability of project data in the DCP database, 4 distinct soil types are characterised as: (i) Sand–Clay; (ii) Clay–Silt; (iii) Sand–None and (iv) Sand–Gravel. Therefore, other variations of soil types have yet to be applied to the DCP models developed in this thesis. In addition, these models can be further enhanced by defining a more specific form of soil type rather including a generalised form, perhaps by incorporating particle size distribution data and Atterberg limits

measurements, for example, which were not available in any great quantity to allow them to be included in the present database.

Moreover, it is important to note that soil moisture content is not included as an input in the developed models, again due to the paucity of data. As such, the moisture content was considered to be described implicitly by the CPT and DCP measurements, given that penetrometer test results are affected by soil moisture. Nevertheless, efforts could be made in the future to incorporate the moisture content at the time of compaction directly as an input variable in the models to enhance the models' accuracy.

Therefore, it is of great value to produce a database with an extensive set of CPT and DCP data in relation to RDC. It is important that the testing uncertainties in the form of operator, procedural and equipment, that could have associated with random errors in the dataset, are minimised. Consequently, more flexible and reliable predictive models, having a wide range of applicability, can be obtained. However, these would only be possible when the appropriate set of data are readily available to form a comprehensive database, which is very difficult at present with the currently available resources.

Overall, the application of AI techniques, particularly ANNs and LGP, for prediction purposes, in relation to RDC was shown to be promising. It was evident that these techniques have the ability to model complex interactions between many parameters and are also capable of producing tractable equations and/or computer codes facilitating their use in practice. Thus, it is recommended that consideration be given to these AI techniques as being useful means of forecasting other aspects of RDC, such as ground vibrations associated with impact rolling. Moreover, the research presented in this thesis has solely focused on the 4-sided, 8 tonne impact roller (BH-1300), whereas other module variants having different shapes and weights are also required to be assessed in terms of their performance and the effectiveness in different ground conditions.

In addition, research effort is needed in regards the theoretical enhancements of these AI techniques, such as ANN model transparency, knowledge extraction and model uncertainty (Shahin, 2016), that may improve their usefulness and applicability. As it was evident in this research, as well as in other previous applications in the literature [e.g. Pooya Nejad et al. (2009)], ANNs can possibly produce more accurate predictions with multi-hidden layers, but at the same time, they become less transparent resulting in difficulties in knowledge extraction. In this regard, alternative methods, such as neurofuzzy networks (Shahin et al., 2009) could also be explored in terms of model transparency. Although, LGP is also considered to be an alternative approach, in the RDC case studies presented in this thesis, some difficulties have been experienced in terms of producing relatively simple formulae. However, this could be attributed to the fact that LGP penalises complex models to a greater extent and also due to the spatial variability associated with the parameters, which may include sudden fluctuations, discontinuities and uncertainties. In this regard, further modelling with LGP may be beneficial. It is suggested to incorporate a dataset of carefully controlled measurements with the intention of producing a relatively simple formula based on LGP.

INTENTIONALLY BLANK

References

Abdel-Rahman, A. H. (2008). Predicting Compaction of Cohesionless Soils Using ANN. *Proceedings of the ICE-Ground Improvement,* Vol. *161*(1), pp. 3-8.

Al-saffar, R. Z., Khattab, S. I., and Yousif, S. T. (2013). Prediction of Soil's Compaction Parameter Using Artificial Neural Network. *Al-Rafidain Engineering,* Vol. *27*(3), pp. 15-27.

Alavi, A. H., and Gandomi, A. H. (2012). Energy-Based Numerical Models for Assessment of Soil Liquefaction. *Geoscience Frontiers,* Vol. *3*(4), pp. 541-555.

Alavi, A. H., Gandomi, A. H., Mollahasani, A., and Bazaz, J. B. (2013). Linear and Tree-Based Genetic Programming for Solving Geotechnical Engineering Problems. In *Metaheuristics in Water, Geotechnical and Transport Engineering*, Yang, X.-S., Gandomi, A. H., Talatahari, S. and Alavi, A. H. (Eds.), Elsevier, Waltham, MA, USA, pp. 289-310.

Alavi, A. H., Gandomi, A. H., Sahab, M. G., and Gandomi, M. (2010). Multi Expression Programming: A New Approach to Formulation of Soil Classification. *Engineering with Computers,* Vol. *26*(2), pp. 111-118.

Alavi, A. H., and Sadrossadat, E. (2016). New Design Equations for Estimation of Ultimate Bearing Capacity of Shallow Foundations Resting on Rock Masses. *Geoscience Frontiers,* Vol. *7*(1), pp. 91-99.

Amari, S., Murata, N., Muller, K. R., Finke, M., and Yang, H. H. (1997). Asymptotic Statistical Theory of Overtraining and Cross-Validation. *IEEE Transactions on Neural networks,* Vol. *8*(5), pp. 985-996.

Aminian, P., Niroomand, H., Gandomi, A. H., Alavi, A. H., and Esmaeili, M. A. (2013). New Design Equations for Assessment of Load Carrying Capacity of Castellated Steel Beams: A Machine Learning Approach. *Neural Computing and Applications,* Vol. *23*(1), pp. 119-131.

Army, U. (1999). Guidelines on Ground Improvement for Structures and Facilities. *Report No. ETL 1110-1-185*, US Army Corps of Engineers, Department of the Army, Washington, DC, 109 p.

Auzins, N., and Southcott, P. H. (1999). Minimising Water Losses in Agriculture through the Application of Impact Rollers. *Proceedings of the 8th Australia New Zealand Conference on Geomechanics*, Hobart, Australia, pp. 651-657.

Avalle, D. L. (2004a). Ground Improvement Using the" Square" Impact Roller-Case Studies. *Proceedings of the 5th International Conference on Ground Improvement Techniques*, Kuala Lumpur, Malaysia, pp. 101-108.

Avalle, D. L. (2004b). Impact Rolling in the Spectrum of Compaction Techniques and Equipment. *Proceedings of the Earthworks Seminar, Australian Geomechanics Society*, Adelaide, Australia.

Avalle, D. L. (2004c). A Note on Specifications for the Use of the Impact Roller for Earthworks. *Proceedings of the Earthworks Seminar, Australian Geomechanics Society*, Adelaide, Australia.

Avalle, D. L. (2004d). Use of the Impact Roller to Reduce Agricultural Water Loss. *Proceedings of the 9th ANZ Conference on Geomechanics*, Auckland, New Zealand, pp. 513-518.

Avalle, D. L. (2006). Reducing Haul Road Maintenance Costs and Improving Tyre Wear Through the Use of Impact Rollers. *Proceedings of the Mining for Tyres Conference*, Perth, Australia, pp. 5.

Avalle, D. L. (2007a). Ground Vibrations During Impact Rolling. *Proceedings of the 10th Australia - New Zealand Conference (ANZ) on Geomechanics*, Brisbane, Australia, pp. 1-6.

Avalle, D. L. (2007b). Trials and Validation of Deep Compaction using the “Square” Impact Roller. *Proceedings of the AGS Symposium - Advances in Earthworks*, Sydney, Australia, pp. 63 - 70.

Avalle, D. L., and Carter, J. P. (2005). Evaluating the Improvement from Impact Rolling on Sand. *Proceedings of the 6th International Conference on Ground Improvement Techniques*, Coimbra, Portugal, pp. 8.

Avalle, D. L., and Grounds, R. (2004). Improving Pavement Subgrade with the "Sqaure" Impact Roller. *Proceedings of the 23rd Annual South African Transport Conference*, Pretoria, South Africa.

Avalle, D. L., and McKenzie, R. W. (2005). Ground Improvement of Landfill Site using the "Square" Impact Roller. *Australian Geomechanics*, Vol. *40*(4), pp. 15 - 21.

Avalle, D. L., Scott, B. T., and Jaksa, M. B. (2009). Ground Energy and Impact of Rolling Dynamic Compaction–Results from Research Test Site. *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering*, Alexandria, Egypt, pp. 2228-2231.

Avalle, D. L., and Young, G. (2004). Trial Programme and Recent Use of the Impact Roller in Sydney. *Proceedings of the Earthworks Seminar, Australian Geomechanics Society*, Adelaide, Australia.

Azamathulla, H. M., Guven, A., and Demir, Y. K. (2011). Linear Genetic Programming to Scour below Submerged Pipeline. *Ocean Engineering*, Vol. *38*(8), pp. 995-1000.

Babanajad, S. K., Gandomi, A. H., Mohammadzadeh S, D., and Alavi, A. H. (2013). Numerical Modeling of Concrete Strength Under Multiaxial Confinement Pressures Using Linear Genetic Programming. *Automation in Construction*, Vol. *36*, pp. 136-144.

Banimahd, M., Yasrobi, S. S., and Woodward, P. K. (2005). Artificial Neural Network for Stress–Strain Behavior of Sandy Soils: Knowledge Based Verification. *Computers and Geotechnics*, Vol. *32*(5), pp. 377-386.

Banzhaf, W., Francone, F. D., and Nordin, P. (1996). The Effect of Extensive Use of the Mutation Operator on Generalization in Genetic Programming Using Sparse Data Sets. In *Parallel Problem Solving from Nature—PPSN IV*, Springer, pp. 300-309.

Banzhaf, W., Nordin, P., Keller, R. E., and Francone, F. D. (1998). *Genetic Programming: An Introduction* (Vol. 270). Morgan Kaufmann, San Francisco, 479 p.

Barrett, A. J., and Wrench, B. P. (1984). Impact Rolling Trial on 'Collapsing' Aeolian Sand. *Proceedings of the 8th Regional Conference for Africa on Soil Mechanics and Foundation Engineering*, Harare, Zimbabwe, pp. 177-181.

Basheer, I. A. (2001). Empirical Modeling of the Compaction Curve of Cohesive Soils. *Canadian Geotechnical Journal*, Vol. *38*(1), pp. 29-45.

Basheer, I. A., and Hajmeer, M. (2000). Artificial Neural Networks: Fundamentals, Computing, Design, and Application. *Journal of microbiological methods*, Vol. *43*(1), pp. 3-31.

Baykasoğlu, A., Güllü, H., Çanakçı, H., and Özbakır, L. (2008). Prediction of Compressive and Tensile Strength of Limestone via Genetic Programming. *Expert Systems with Applications*, Vol. *35*(1-2), pp. 111-123.

Benson, C. H., Zhai, H., and Wang, X. (1994). Estimating Hydraulic Conductivity of Compacted Clay Liners. *Journal of Geotechnical Engineering*, Vol. *120*(2), pp. 366-387.

Berke, L., and Hajela, P. (1993). Application of Neural Nets in Structural Optimization. In *Optimization of Large Structural Systems*, Rozvany, G. I. N. (Ed.), Kluwer Academic Publishers, Dordrecht, The Nederlands, Vol. II, pp. 731-745.

Bierbaum, T. C., Lane, R. A., Piotto, J. L., and Treloar, T. J. (2010). *The Influence Zone of the Impact Roller*. Paper presented at the B.Eng.(Hons) Undergraduate Conference - The University of Adelaide, Adelaide, Australia.

Boroumand, A., and Baziar, M. H. (2005). *Determination of Compacted Clay Permeability by Artificial Neural Networks*. Paper presented at the 9th International Water Technology Conference (IWTC9), Sharm El-Sheikh, Egypt.

Bouazza, A., and Avalle, D. L. (2006). Effectiveness of Rolling Dynamic Compaction on an Old Waste Tip. *Proceedings of the 5th International Congress on Environmental Geotechnics (ICEG)*, Cardiff, UK, pp. 384 - 390.

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,* Vol. *38*(2), pp. 2-11.

Bradley, A., Crisp, A., Jiang, J., and Power, C. (2012). *Assessing the Effectiveness of Rolling Dynamic Compaction Using LS-DYNA.* Paper presented at the B.Eng.(Hons) Undergraduate Conference - The University of Adelaide, Adelaide, Australia.

Brameier, M., and Banzhaf, W. (2001). A Comparison of Linear Genetic Programming and Neural Networks in Medical Data Mining. *IEEE Transactions on Evolutionary Computation,* Vol. *5*(1), pp. 17-26.

Brameier, M. F., and Banzhaf, W. (2007). *Linear Genetic Programming*. Springer Science & Business Media, New York, USA, 320 p.

Breiman, L. (1994). Neural Networks: A Review from a Statistical Perspective: Comment *Statistical Science,* Vol. *9*(1), pp. 38-42.

Broons Hire. (2016). The Crushing and Compaction Specialists. Retrieved from http://www.broons.com/

Burke, L. I. (1991). Introduction to Artificial Neural Systems for Pattern Recognition. *Computers and Operations Research,* Vol. *18*(2), pp. 211-220.

Caudill, M. (1988). Neural Networks Primer, Part III. *AI Expert,* Vol. *3*(6), pp. 53-59.

Clegg, B., and Berrangé, A. R. (1971). The Development and Testing of An Impact Roller. *The Civil Engineer in South Africa,* Vol. *13*(3), pp. 65-73.

Clifford, J. M. (1976). *Impact Rolling and Construction Techniques*. Paper presented at the Australian Road Research Board Conference, Perth, Australia.

Clifford, J. M. (1978a). *Evaluation of Compaction Plant and Methods for the Construction of Earthworks in Southern Africa.* M.Sc. Thesis, University of Natal, Durban. 308 p.

Clifford, J. M. (1978b). The Impact Roller - Problem Solved. *The Civil Engineer in South Africa,* Vol. *20*(12), pp. 321-324.

Clifford, J. M., and Bowes, G. (1995). *Calculating the Energy Delivered by an Impact Roller*. Paper presented at the A trilogy of Papers for the September 1995 Lecture Tour and International Seminars to commemorate the tenth anniversary of the BH 1300 Standard Impact Roller, Adelaide, Australia.

Cybenko, G. (1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Systems,* Vol. *2*(4), pp. 303-314.

Dai, H., and MacBeth, C. (1997). Effects of Learning Parameters on Learning Procedure and Performance of a BPNN. *Neural Networks,* Vol. *10*(8), pp. 1505-1521.

Das, S. K., and Basudhar, P. K. (2007). Prediction of Hydraulic Conductivity of Clay Liners Using Artificial Neural Network. *Lowland technology international: the officisl journal of the International Association of Lowland Technology (IALT)/Institute of Lowland Technology, Saga University,* Vol. *9*(1), pp. 50-58.

Das, S. K., Samui, P., and Sabat, A. K. (2011). Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil. *Geotechnical and Geological Engineering,* Vol. *29*(3), pp. 329-342.

Dowla, F. U., and Rogers, L. L. (1995). *Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks*. MIT Press, Cambridge, Massachusetts, 241 p.

Erzin, Y., Gumaste, S. D., Gupta, A. K., and Singh, D. N. (2009). Artificial Neural Network (ANN) Models for Determining Hydraulic Conductivity of Compacted Fine-Grained Soils. *Canadian Geotechnical Journal,* Vol. *46*(8), pp. 955-968.

Fausett, L. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. Prentice-Hall, Inc., New Jersey, USA, 461 p.

Flood, I., and Kartam, N. (1994). Neural Networks in Civil Engineering. I: Principles and Understanding. *Journal of Computing in Civil Engineering,* Vol. *8*(2), pp. 131-148.

Francone, F. D. (2010). Discipulus TM with Notitia and Solution Analytics Owner's Manual, Register Machine Learning Technologies Inc., Littleton, CO, USA, 272 p.

Francone, F. D., and Deschaine, L. M. (2004). Extending the Boundaries of Design Optimization by Integrating Fast Optimization Techniques with Machine-Code-Based, Linear Genetic Programming. *Information Sciences,* Vol. *161*(3), pp. 99-120.

Fu, L. M. (2003). *Neural Networks in Computer Intelligence*. Tata McGraw-Hill Education Private Limited, New York, USA, 460 p.

Gandomi, A. H., Alavi, A. H., and Sahab, M. G. (2010a). New Formulation for Compressive Strength of CFRP Confined Concrete Cylinders Using Linear Genetic Programming. *Materials and Structures,* Vol. *43*(7), pp. 963-983.

Gandomi, A. H., Alavi, A. H., Sahab, M. G., and Arjmandi, P. (2010b). Formulation of Elastic Modulus of Concrete Using Linear Genetic Programming. *Journal of Mechanical Science and Technology,* Vol. *24*(6), pp. 1273-1278.

Gandomi, A. H., Mohammadzadeh S, D., Pérez-Ordóñez, J. L., and Alavi, A. H. (2014). Linear Genetic Programming for Shear Strength Prediction of Reinforced Concrete Beams without Stirrups. *Applied Soft Computing,* Vol. *19,* pp. 112-120.

Gandomi, A. H., Tabatabaei, S. M., Moradian, M. H., Radfar, A., and Alavi, A. H. (2011). A New Prediction Model for the Load Capacity of Castellated Steel Beams. *Journal of Constructional Steel Research,* Vol. *67*(7), pp. 1096-1105.

Garson, D. G. (1991). Interpreting Neural Network Connection Weights. *AI Expert,* Vol. *6*(7), pp. 47-51.

Goh, A. T. C. (1994). Seismic Liquefaction Potential Assessed by Neural Networks. *Journal of Geotechnical Engineering,* Vol. *120*(9), pp. 1467-1480.

Goh, A. T. C., Kulhawy, F. H., and Chua, C. G. (2005). Bayesian Neural Network Analysis of Undrained Side Resistance of Drilled Shafts. *Journal of Geotechnical and Geoenvironmental Engineering,* Vol. *131*(1), pp. 84-93.

Golbraikh, A., and Tropsha, A. (2002). Beware of q 2! *Journal of Molecular Graphics and Modelling,* Vol. *20*(4), pp. 269-276.

Günaydın, O. (2009). Estimation of Soil Compaction Parameters by Using Statistical Analyses and Artificial Neural Networks. *Environmental Geology,* Vol. *57*(1), pp. 203-215.

Hammerstrom, D. (1993). Working with Neural Networks. *IEEE spectrum,* Vol. *30*(7), pp. 46-53.

Hassoun, M. H. (1995). *Fundamentals of Artificial Neural Networks*. MIT press, London, England, 511 p.

Hausmann, M. R. (1990). *Engineering Principles of Ground Modification*. McGraw-Hill, New York, 632 p.

Hecht-Nielsen, R. (1988). Neurocomputing: Picking the Human Brain. *IEEE spectrum,* Vol. *25*(3), pp. 36-41.

Hecht-Nielsen, R. (1989). Theory of the Backpropagation Neural Network. *Proceedings of the International Joint Conference on Neural Networks (IJCNN),* San Diego, California, pp. 593-605.

Heshmati, A. A. R., Salehzade, H., Alavi, A. H., Gandomi, A. H., Badkobeh, A., and Ghasemi, A. (2008). On the Applicability of Linear Genetic Programming for the Formulation of Soil Classification. *American-Eurasian Journal of Agricultural and Environmental Science,* Vol. *4*(5), pp. 575-583.

Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural computation,* Vol. *9*(8), pp. 1735-1780.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, New York, USA, 232 p.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer Feedforward Networks are Universal Approximators. *Neural Networks,* Vol. *2*(5), pp. 359-366.

Isik, F., and Ozden, G. (2013). Estimating Compaction Parameters of Fine-and Coarse-Grained Soils by Means of Artificial Neural Networks. *Environmental Earth Sciences,* Vol. *69*(7), pp. 2287-2297.

Jaksa, M. B., Scott, B. T., Mentha, N. L., Symons, A. T., Pointon, S. M., Wrightson, P. T., and Syamsuddin, E. (2012). *Quantifying the Zone of Influence of the Impact Roller*. Paper presented at the International Symposium on Ground Improvement Brussels, Belgium.

Javadi, A. A., Rezania, M., and Nezhad, M. M. (2006). Evaluation of Liquefaction Induced Lateral Displacements Using Genetic Programming. *Computers and Geotechnics,* Vol. *33*(4–5), pp. 222-233.

Jeng, Y., and Strohm, W. E. (1976). *Prediction of the Shear Strength and Compaction Characteristics of Compacted Fine-Grained Cohesive Soils*. US Waterways Experiment Station, Soil and Pavement Laboratory, Vicksburg, MO, 85 p.

Johari, A., Habibagahi, G., and Ghahramani, A. (2006). Prediction of Soil–Water Characteristic Curve Using Genetic Programming. *Journal of Geotechnical and Geoenvironmental Engineering,* Vol. *132*(5), pp. 661-665.

Kelly, D. B. (2000). Deep In-Situ Ground Improvement Using High Energy Impact Compaction (Heic) Technology. *Proceedings of the International Conference on Geological and Geotechnical Engineering*, Melbourne, Australia, pp. 6.

Khotanzad, A., Afkhami-Rohani, R., Lu, T.-L., Abaye, A., Davis, M., and Maratukulam, D. J. (1997). ANNSTLF-a Neural-Network-Based Electric Load Forecasting System. *IEEE Transactions on Neural networks,* Vol. *8*(4), pp. 835-846.

Kim, K. (2010). *Numerical Simulation of Impact Rollers for Estimating the Influence Depth of Soil Compaction.* M.Sc. Thesis, Texas A&M University, Texas, USA. 144 p.

Kohonen, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological cybernetics,* Vol. *43*(1), pp. 59-69.

Koza, J. R. (1992). *Genetic Programming: on the Programming of Computers by Means of Natural Selection* (Vol. 1). MIT press, USA, 819 p.

Koza, J. R. (1994). Genetic Programming as a Means for Programming Computers by Natural Selection. *Statistics and computing,* Vol. *4*(2), pp. 87-112.

Koza, J. R., and Poli, R. (2005). Genetic Programming. In *Search Methodologies,* Burke, E. K. and Kendall, G. (Eds.), Springer, pp. 127-164.

Kuo, Y. L., Jaksa, M. B., Lyamin, A. V., and Kaggwa, W. S. (2009). Ann-Based Model for Predicting the Bearing Capacity of Strip Footing on Multi-Layered Cohesive Soil. *Computers and Geotechnics,* Vol. *36*(3), pp. 503-516.

Lachtermacher, G., and Fuller, J. D. (1994). Backpropagation in Hydrological Time Series Forecasting. *Stochastic and Statistical Methods in Hydrology and Environmental Engineering,* Vol. *10*(3), pp. 229-242.

Landpac International. (2016). Landpac Technologies: Ground Improvement and Verification Technologies and Solutions. Retrieved from http://www.landpac.com

Lukas, R. G. (1995). Geotechnical Engineering Circular No. 1: Dynamic Compaction. *Publication No. FHWA-SA-95-037,* Federal Highway Administration, US Department of Transportation, Washington DC, USA.

Lunne, T., Robertson, P. K., and Powell, J. J. M. (1997). *Cone Penetration Testing in Geotechnical Engineering Practice*. Blackie Academic and Professional, New York, USA, 352 p.

Maier, H. R. (1995). A Review of Artificial Neural Networks. *Report No. R 131*, Department of Civil and Environmental Engineering, The University of Adelaide, Adelaide, Australia, 96 p.

Maier, H. R., and Dandy, G. C. (1996). The Use of Artificial Neural Networks for the Prediction of Water Quality Parameters. *Water Resources Research,* Vol. *32*(4), pp. 1013-1022.

Maier, H. R., 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,* Vol. *13*(2), pp. 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 and Software,* Vol. *15*(1), pp. 101-124.

Maier, H. R., and Dandy, G. C. (2001). Neural Network Based Modelling of Environmental Variables: A Systematic Approach. *Mathematical and Computer Modelling,* Vol. *33*(6), pp. 669-682.

Maier, H. R., Jain, A., Dandy, G. C., and Sudheer, K. P. (2010). Methods Used for the Development of Neural Networks for the Prediction of Water Resource Variables in River Systems: Current Status and Future Directions. *Environmental Modelling & Software,* Vol. *25*(8), pp. 891-909.

Maren, A. J., Harston, C. T., and Pap, R. M. (1990). *Handbook of Neural Computing Applications*. Academic Press, San Diego, California, 470 p.

Masters, T. (1993). *Practical Neural Network Recipes in C ++*. Academic Press, San Diego, California, USA, 493 p.

McCann, K., and Dix, S. (2007). *Engineered Impact Compaction of Un-Engineered Fills*. Paper presented at the Earthworks Symposium, Sydney, Australia.

Mehr, A. D., Kahya, E., and Olyaie, E. (2013). Streamflow Prediction Using Linear Genetic Programming in Comparison with a Neuro-Wavelet Technique. *Journal of Hydrology,* Vol. *505*, pp. 240-249.

Mehr, A. D., Kahya, E., and Yerdelen, C. (2014). Linear Genetic Programming Application for Successive-Station Monthly Streamflow Prediction. *Computers and Geosciences,* Vol. *70*, pp. 63-72.

Mentha, N. L., Pointon, S. M., Symons, A. T., and Wrightson, P. T. (2011). *The Effectiveness of the Impact Roller*. Paper presented at the B.Eng.(Hons) Undergraduate Conference - The University of Adelaide, Adelaide, Australia.

Minns, A. W., and Hall, M. J. (1996). Artificial Neural Networks As Rainfall-Runoff Models. *Hydrological Sciences Journal,* Vol. *41*(3), pp. 399-417.

Mitchell, J. M., and Jardine, F. M. (2002). Guide to Ground Treatment. *Report No. CIRIA C573*, Construction Industry Research and Information Association, London, UK, 260 p.

Mitchell, T. M. (1997). Does Machine Learning Really Work? *AI magazine*, Vol. *18*(3), pp. 11.

Mousavi, S. M., Alavi, A. H., Mollahasani, A., and Gandomi, A. H. (2011). A Hybrid Computational Approach to Formulate Soil Deformation Moduli Obtained from PLT. *Engineering Geology*, Vol. *123*(4), pp. 324-332.

Munfakh, G. A., and Wyllie, D. C. (2000). Ground Improvement Engineering-Issues and Selection. *Proceedings of the ISRM International Symposium*, Melbourne, Australia.

Naderi, N., Roshani, P., Samani, M. Z., and Tutunchian, M. A. (2012). Application of Genetic Programming for Estimation of Soil Compaction Parameters. *Applied Mechanics and Materials*, Vol. *147*, pp. 70-74.

Najjar, Y. M., and Basheer, I. A. (1996). Utilizing Computational Neural Networks for Evaluating the Permeability of Compacted Clay Liners. *Geotechnical and Geological Engineering*, Vol. *14*(3), pp. 193-212.

Najjar, Y. M., Basheer, I. A., and Naouss, W. A. (1996). On the Identification of Compaction Characteristics by Neuronets. *Computers and Geotechnics*, Vol. *18*(3), pp. 167-187.

Nash, T. R. (2010). The Effectiveness of an Impact Roller on Alluvial Sandy Clays. *Australian Geomechanics*, Vol. *45*(4), pp. 105.

Neusciences. (2000). Neuframe Version 4.0. , Neusciences Corporation, Southampton, Hampshire.

Noor, S., and Singh, A. (2012). Use of Genetic Programming to Evaluate Proctor Properties of Compacted Soils. *International Journal of Latest Trends in Engineering and Technology*, Vol. *1*(4), pp. 17-21.

Nordin, P. (1994). A Compiling Genetic Programming System that Directly Manipulates the Machine Code. In *Advances in genetic programming*, Kinnear, K. E. (Ed.), MIT Press, London, UK, Vol. 1, pp. 311-331.

Nordin, P., Banzhaf, W., and Francone, F. D. (1999). Efficient Evolution of Machine Code for Cisc Architectures Using Instruction Blocks and Homologous Crossover. In *Advances in genetic programming*, Spector, L., Langdon, W. B., O'Reilly, U.-M. and Angeline, P. J. (Eds.), The MIT Press, London, England, Vol. 3, pp. 275-299.

Oltean, M., and Grosan, C. (2003). A Comparison of Several Linear Genetic Programming Techniques. *Complex Systems*, Vol. *14*(4), pp. 285-314.

Onoda, T. (1995). Neural Network Information Criterion for the Optimal Number of Hidden Units. *Proceedings of the IEEE International Conference on Neural Networks*, Perth, Australia, pp. 275-280.

Orchant, C. J., Kulhawy, F. H., and Trautmann, C. H. (1988). Critical Evaluation of In-Situ Test Methods and Their Variability. *Report No. EL-5507*, Electric Power Research Institute, Palo Alto, California, 310 p.

Parisi, R., Di Claudio, E. D., Orlandi, G., and Rao, B. D. (1996). A Generalized Learning Paradigm Exploiting the Structure of Feedforward Neural Networks. *IEEE Transactions on Neural networks*, Vol. *7*(6), pp. 1450-1460.

Phear, A. G., and Harris, S. J. (2008). Contributions to Géotechnique 1948–2008: Ground Improvement. *Géotechnique*, Vol. *58*(5), pp. 399-404.

Pinard, M. I. (1999). Innovative Developments in Compaction Technology Using High Energy Impact Compactors. *Proceedings of the 8th Australia New Zealand Conference on Geomechanics*, Hobart, Australia, pp. 775-781.

Pinard, M. I., and Ookeditse, S. (1988). Evaluation of High Energy Impact Compaction Techniques for Minimising Construction Water Requirements in Semi Arid Regions. *Proceedings of the 14th Australian Road Research Board (ARRB) Conference*, Canberra, Australia.

Poli, R., Langdon, W. B., McPhee, N. F., and Koza, J. R. (2007). Genetic Programming: An Introductory Tutorial and a Survey of Techniques and Applications. *Tutorial No. CES-475*, University of Essex, UK, 112 p.

Poli, R., Langdon, W. B., McPhee, N. F., and Koza, J. R. (2008). *A Field Guide to Genetic Programming*. Lulu.com, San Francisco, USA, 250 p.

Pooya Nejad, F., Jaksa, M. B., Kakhi, M., and McCabe, B. A. (2009). Prediction of Pile Settlement Using Artificial Neural Networks Based on Standard Penetration Test Data. *Computers and Geotechnics,* Vol. *36*(7), pp. 1125-1133.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016a). Application of Artificial Neural Networks for Predicting the Impact of Rolling Dynamic Compaction Using Dynamic Cone Penetrometer Test Results. Journal of *Rock Mechanics and Geotechnical Engineering*, In Press.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016b). A Genetic Programming Approach for Predicting the Effectiveness of Rolling Dynamic Compaction Using Cone Penetration Test Data. *ICE-Ground Improvement*, Under Review.

Ranasinghe, R. A. T. M., Jaksa, M. B., Kuo, Y. L., and Pooya Nejad, F. (2016c). Prediction of the Effectiveness of Rolling Dynamic Compaction Using Artificial Neural Networks and Cone Penetration Test Data. *Soils and Foundations*, Submitted for Review.

Rashed, A., Bazaz, J. B., and Alavi, A. H. (2012). Nonlinear Modeling of Soil Deformation Modulus Through Lgp-Based Interpretation of Pressuremeter Test Results. *Engineering applications of artificial intelligence*, Vol. *25*(7), pp. 1437-1449.

Rezania, M., and Javadi, A. A. (2007). A New Genetic Programming Model for Predicting Settlement of Shallow Foundations. *Canadian Geotechnical Journal*, Vol. *44*(12), pp. 1462-1473.

Ripley, B. D. (1994). Neural Networks and Related Methods for Classification. *Journal of the Royal Statistical Society. Series B (Methodological),* Vol. *56*(3), pp. 409-456.

Robertson, P. K. (1990). Soil Classification Using the Cone Penetration Test. *Canadian Geotechnical Journal,* Vol. *27*(1), pp. 151-158.

Rogers, L. L., and Dowla, F. U. (1994). Optimization of Groundwater Remediation Using Artificial Neural Networks with Parallel Solute Transport Modeling. *Water Resources Research,* Vol. *30*(2), pp. 457-481.

Rojas, R. (1996). *Neural Networks: A Systematic Introduction*. Springer Science & Business Media, Berlin, Germany, 502 p.

Roy, P. P., and Roy, K. (2008). On Some Aspects of Variable Selection for Partial Least Squares Regression Models. *QSAR & Combinatorial Science,* Vol. *27*(3), pp. 302-313.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning Representations by Backpropagating Errors. In *Cognitive Modeling,* Polk, T. A. and Seifert, C. M. (Eds.), The MIT Press, London, UK, pp. 213-221.

Salchenberger, L. M., Cinar, E., and Lash, N. A. (1992). Neural Networks: A New Tool for Predicting Thrift Failures. *Decision Sciences,* Vol. *23*(4), pp. 899-916.

Sarle, W. S. (1994). Neural Networks and Statistical Models. *Proceedings of the 19th Annual SAS Users Group International Conference*, Dallas, Texas, pp. 1538-1550.

Scott, B., and Suto, K. (2007). Case Sudy of Ground Improvement at an Industrial Estate Containing Uncontrolled Fill. *Proceedings of the 10th Australia - New Zealand Conference on Geomechanics*, Brisbane, Australia, pp. 1 - 6.

Scott, B. T., and Jaksa, M. B. (2012). Mining Applications and Case Studies of Rolling Dynamic Compaction. *Proceedings of the 11th Australia-New Zealand (ANZ) Conference on Geomechanics*, Melbourne, Australia, pp. 961-966.

Scott, B. T., and Jaksa, M. B. (2014). Evaluating Rolling Dynamic Compaction of Fill Using CPT. *Proceedings of the 3rd International Symposium on Cone Penetration Testing* Las Vegas, Nevada, USA, pp. 941-948.

Scott, B. T., Jaksa, M. B., and Kuo, Y. L. (2012). *Use of Proctor Compaction Testing for Deep Fill Construction Using Impact Rollers*. Paper presented at the International Conference on Ground Improvement and Ground Control (ICGI 2012), Wollongong, Australia.

Selle, B., and Muttil, N. (2011). Testing the Structure of a Hydrological Model Using Genetic Programming. *Journal of Hydrology*, Vol. *397*(1–2), pp. 1-9.

Sette, S., and Boullart, L. (2001). Genetic Programming: Principles and Applications. *Engineering applications of artificial intelligence*, Vol. *14*(6), pp. 727-736.

Shahin, M. A. (2003). *Use of Artificial Neural Networks for Predicting Settlement of Shallow Foundations on Cohesionless Soils.* Ph.D. Thesis, University of Adelaide, Adelaide, Australia. 317 p.

Shahin, M. A. (2010). Intelligent Computing for Modeling Axial Capacity of Pile Foundations. *Canadian Geotechnical Journal*, Vol. *47*(2), pp. 230-243.

Shahin, M. A. (2016). State-of-the-Art Review of Some Artificial Intelligence Applications in Pile Foundations. *Geoscience Frontiers*, Vol. *7*(1), pp. 33-44.

Shahin, M. A., and Jaksa, M. B. (2006). Pullout Capacity of Small Ground Anchors by Direct Cone Penetration Test Methods and Neural Networks. *Canadian Geotechnical Journal*, Vol. *43*(6), pp. 626-637.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2005a). Neural Network Based Stochastic Design Charts for Settlement Prediction. *Canadian Geotechnical Journal*, Vol. *42*(1), pp. 110-120.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2008). State of the Art of Artificial Neural Networks in Geotechnical Engineering. *Electronic Journal of Geotechnical Engineering*, Vol. *8*, pp. 1-26.

Shahin, M. A., Jaksa, M. B., and Maier, H. R. (2009). Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications. *Advances in Artificial Neural Systems,* Vol. *2009*, pp. 1-9.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2000). Predicting the Settlement of Shallow Foundations on Cohesionless Soils Using Back-Propagation Neural Networks. *Report No. R 167*, Department of Civil and Environmental Engineering, University of Adelaide, Adelaide, Australia, 19 p.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2002). Predicting Settlement of Shallow Foundations using Neural Networks. *Journal of Geotechnical and Geoenvironmental Engineering,* Vol. *128*(9), pp. 785-793.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2003). Settlement Prediction of Shallow Foundations on Granular Soils Using B-Spline Neurofuzzy Models. *Computers and Geotechnics,* Vol. *30*(8), pp. 637-647.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2004). Data Division for Developing Neural Networks Applied to Geotechnical Engineering. *Journal of Computing in Civil Engineering,* Vol. *18*(2), pp. 105-114.

Shahin, M. A., Maier, H. R., and Jaksa, M. B. (2005b). Investigation into the Robustness of Artificial Neural Networks for a Case Study in Civil Engineering. *Proceedings of the 16th International Congress on Modelling and Simulation*, Melbourne, Australia, pp. 79-83.

Sinha, S. K., and Wang, M. C. (2008). Artificial Neural Network Prediction Models for Soil Compaction and Permeability. *Geotechnical and Geological Engineering*, Vol. *26*(1), pp. 47-64.

Smith, G. N. (1986). *Probability and Statistics in Civil Engineering*. Collins, London, England, 240 p.

Smith, M. (1993). *Neural Networks for Statistical Modeling*. Thomson Learning, New York, USA, 235 p.

Standards Association of Australia (1997). Determination of the Penetration Resistance of a Soil - 9kg Dynamic Cone Penetrometer Test. In *Methods of Testing Soils for Engineering Purposes - Soil Strength and Consolidation Tests, AS 1289.6.3.2*, Sydney, Australia.

Standards Association of Australia (1999). Determination of the Static Cone Penetration Resistance of a Soil - Field Test Using a Mechanical and Electrical Cone or Friction-Cone Penetrometer. In *Methods of Testing Soils for Engineering Purposes - Soil Strength and Consolidation Tests, AS 1289.6.5.1*, Sydney, Australia.

Standards Association of Australia (2003a). Determination of the Dry Density/Moisture Content Relation of a Soil Using Modified Compactive Effort. In *Methods of Testing Soils for Engineering Purposes - Soil Compaction and Density Tests, AS 1289.5.2.1*, Sydney, Australia.

Standards Association of Australia (2003b). Determination of the Dry Density/Moisture Content Relation of a Soil Using Standard Compactive Effort. In *Methods of Testing Soils for Engineering Purposes - Soil Compaction and Density Tests, AS 1289.5.1.1*, Sydney, Australia.

Standards Association of Australia (2004a). Determination of the Field Density of a Soil - Sand Replacement Method Using a Sand-Cone Pouring Apparatus. In *Methods of Testing Soils for Engineering Purposes - Soil Compaction and Density Tests, AS 1289.5.3.1*, Sydney, Australia.

Standards Association of Australia (2004b). Determination of the Penetration Resistance of a Soil - Standard Penetration Test (SPT). In *Methods of Testing Soils for Engineering Purposes - Soil Strength and Consolidation Tests, AS 1289.6.3.1*, Sydney, Australia.

Standards Association of Australia (2007). Determination of Field Density and Field Moisture Content of a Soil Using a Nuclear Surface Moisturedensity Gauge - Direct Transmission Mode. In *Methods of Testing Soils for Engineering Purposes - Soil Compaction and Density Tests, AS 1289.5.8.1*, Sydney, Australia.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. Vol. *36*(2), pp. 111-147.

Sulewska, M. J. (2010a). Neural Modelling of Compactibility Characteristics of Cohesionless Soil. *Computer Assisted Mechanics and Engineering Sciences,* Vol. *17*(1), pp. 27-40.

Sulewska, M. J. (2010b). Prediction Models for Minimum and Maximum Dry Density of Non-Cohesive Soils. *Polish Journal of Environmental Studies,* Vol. *19*(4), pp. 797-804.

Swingler, K. (1996). *Applying Neural Networks: A Practical Guide*. Morgan Kaufmann, San Francisco, USA, 319 p.

Tang, Z., de Almeida, C., and Fishwick, P. A. (1991). Time Series Forecasting Using Neural Networks vs. Box-Jenkins Methodology. *Simulation,* Vol. *57*(5), pp. 303-310.

Taskiran, T. (2010). Prediction of California Bearing Ratio (CBR) of Fine Grained Soils by AI Methods. *Advances in Engineering Software,* Vol. *41*(6), pp. 886-892.

Terashi, M., and Juran, I. (2000). Ground Improvement - State of the Art. *Proceedings of the International Conference on Geological and Geotechnical Engineering*, Melbourne, Australia, pp. 461 - 519.

Tokar, A. S., and Johnson, P. A. (1999). Rainfall-Runoff Modeling Using Artificial Neural Networks. *Journal of Hydrologic Engineering,* Vol. *4*(3), pp. 232-239.

Torres, R. d. S., Falcão, A. X., Gonçalves, M. A., Papa, J. P., Zhang, B., Fan, W., and Fox, E. A. (2009). A Genetic Programming Framework for Content-Based Image Retrieval. *Pattern Recognition,* Vol. *42*(2), pp. 283-292.

Twomey, J. M., and Smith, A. E. (1997). Validation and Verification. In *Artificial neural networks for civil engineers: Fundamentals and applications*, Kartam, N., Flood, I. and Garrett, J. H. (Eds.), American Society of Civil Engineers, New York, USA, pp. 44-64.

Walker, M. (2001). Introduction to Genetic Programming, University of Montana, Missoula, USA, 9 p.

Wang, M., and Huang, C. (1984). Soil Compaction and Permeability Prediction Models. *Journal of Environmental Engineering,* Vol. *110*(6), pp. 1063-1083.

Warner, B., and Misra, M. (1996). Understang of Neural Networks as Statistical Tools. *The American Statistician,* Vol. *50*(4), pp. 284-293.

Weigend, A. S., Huberman, B. A., and Rumelhart, D. E. (1990). Predicting the Future: A Connectionist Approach. *International journal of neural systems,* Vol. *1*(3), pp. 193-209.

White, H. (1989). Learning in Artificial Neural Networks: A Statistical Perspective. *Neural computation,* Vol. *1*(4), pp. 425-464.

Wythoff, B. J. (1993). Backpropagation Neural Networks: A Tutorial. *Chemometrics and Intelligent Laboratory Systems,* Vol. *18*(2), pp. 115-155.

Zupan, J., and Gasteiger, J. (1993). *Neural Networks for Chemists: An Introduction*. John Wiley & Sons, Inc., New York, USA, 305 p.

Appendix A

C Code for the Selected Optimal Linear Genetic Programming Model Based on Cone Penetrometer Test Data

The selected optimal LGP program based on CPT data is represented in C code as follows:

Note that input 000, 001, 002 and 003 represent the depth of measurement (m), initial cone tip resistance (MPa) and sleeve friction (kPa) prior to compaction and the number of roller passes, respectively.

```
float DiscipulusCFunction(float v[])
{
long double f[8];
long double tmp = 0;
int cflag = 0;
f[0]=f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=f[7]=0;

L0:  f[0]+=Input001;
L1:  f[0]*=0.1595308780670166f;
L2:  f[0]=sin(f[0]);
L3:  f[1]+=f[0];
L4:  f[1]+=f[0];
L5:  f[0]-=f[0];
L6:  f[0]=cos(f[0]);
L7:  f[2]-=f[0];
L8:  f[0]/=0.6593866348266602f;
L9:  f[0]-=-0.8144187927246094f;
L10: f[0]*=Input003;
L11: f[0]-=Input002;
L12: f[0]+=f[2];
L13: f[2]*=f[0];
L14: f[0]/=f[0];
L15: f[0]+=f[1];
L16: f[0]*=-0.5910544395446777f;
L17: f[0]+=f[2];
L18: f[2]+=f[0];
L19: f[0]=fabs(f[0]);
L20: f[3]+=f[0];
L21: f[0]=cos(f[0]);
L22: f[0]*=0.4784109592437744f;
L23: f[3]+=f[0];
L24: f[0]/=f[0];
L25: f[2]+=f[0];
L26: f[0]*=Input000;
L27: f[2]-=f[0];
L28: f[0]+=-0.6615190505981445f;
L29: f[0]*=-0.3511595726013184f;
L30: f[3]*=f[0];
L31: f[0]+=Input001;
L32: f[1]+=f[0];
L33: f[2]*=f[0];
L34: f[0]*=f[2];
L35: f[0]+=f[1];
L36: f[0]=C_F2XM1;
L37: f[3]*=f[0];
L38: f[0]-=f[0];
L39: f[0]-=f[2];
```

```
L40:    f[0]=sin(f[0]);
L41:    f[3]-=f[0];
L42:    f[0]*=f[0];
L43:    f[3]+=f[0];
L44:    f[1]-=f[0];
L45:    f[0]-=Input002;
L46:    f[0]*=0.2955143451690674f;
L47:    f[1]-=f[0];
L48:    f[0]*=f[2];
L49:    f[0]+=Input001;
L50:    f[0]+=f[0];
L51:    f[3]-=f[0];
L52:    f[0]+=Input001;
L53:    f[0]+=f[0];
L54:    f[0]*=Input000;
L55:    f[0]-=f[1];
L56:    f[0]=sin(f[0]);
L57:    f[0]*=f[0];
L58:    f[0]+=Input001;
L59:    f[0]/=0.8695814609527588f;
L60:    f[0]+=Input001;
L61:    f[0]*=Input003;
L62:    f[0]+=Input001;
L63:    f[0]-=Input000;
L64:    f[0]+=Input001;
L65:    f[0]-=Input002;
L66:    f[0]=fabs(f[0]);
L67:    f[3]+=f[0];
L68:    f[0]-=0.2995789051055908f;
L69:    f[0]+=f[1];
L70:    f[0]+=-0.6615190505981445f;
L71:    f[0]*=-0.3288925886154175f;
L72:    f[0]*=0.2570122480392456f;
L73:    f[3]*=f[0];
L74:    f[0]-=f[3];
L75:    f[0]+=Input001;
L76:
if (!_finite(f[0])) f[0]=0;
return f[0];
}
```

INTENTIONALLY BLANK

Appendix B

C Code for the Selected Optimal Linear Genetic Programming Model Based on Dynamic Cone Penetrometer Test Data

The selected optimal LGP program based on DCP data is represented in C code as follows:

Note that input 000, 001, 002, 003 and 004 represent the soil type, average depth (m), initial number of roller passes, initial DCP (blows/300 mm) and final number of roller passes, respectively. The soil type variable uses a numerical representation, where 1, 2, 3 and 4 are assigned to Sand–Clay, Clay–Silt, Sand–None and Sand–Gravel, respectively.

```
float DiscipulusCFunction(float v[])
{
long double f[8];
long double tmp = 0;
int cflag = 0;
f[0]=f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=f[7]=0;

L0:  f[0]-=Input003;
L1:  f[0]+=Input000;
L2:  f[0]/=-0.0910041332244873f;
L3:  f[0]=fabs(f[0]);
L4:  f[0]-=1.048232078552246f;
L5:  f[0]=fabs(f[0]);
L6:  f[0]*=Input001;
L7:  f[0]=sqrt(f[0]);
L8:  f[0]=sin(f[0]);
L9:  f[0]/=-0.7297487258911133f;
L10: f[0]*=-0.2360081672668457f;
L11: f[0]+=0.1756083965301514f;
L12: f[0]-=Input002;
L13: f[1]-=f[0];
L14: f[0]*=f[0];
L15: f[0]+=f[0];
L16: f[0]+=Input002;
L17: f[0]=cos(f[0]);
L18: f[0]=cos(f[0]);
L19: f[0]+=1.501374244689941f;
L20: f[0]+=Input003;
L21: f[0]/=0.00262165069580078f;
L22: f[0]=sin(f[0]);
L23: f[0]=fabs(f[0]);
L24: f[0]+=1.987620830535889f;
L25: f[1]*=f[0];
L26: f[0]-=f[0];
L27: f[0]-=0.4281637668609619f;
L28: f[0]/=-1.427085638046265f;
L29: f[0]*=Input004;
L30: f[1]-=f[0];
L31: f[0]-=f[1];
L32: f[0]/=f[0];
L33: f[0]+=Input003;
```

```
L34:    f[1]/=f[0];
L35:    f[0]+=f[0];
L36:    f[0]+=Input000;
L37:    f[0]*=Input004;
L38:    f[0]+=-0.5786151885986328f;
L39:    f[1]-=f[0];
L40:    f[0]+=f[0];
L41:    f[0]-=Input004;
L42:    f[0]*=Input003;
L43:    f[0]+=f[1];
L44:    f[1]+=f[0];
L45:    f[0]+=Input001;
L46:    f[0]-=Input002;
L47:    f[0]*=f[1];
L48:    f[0]*=Input003;
L49:    f[0]-=Input004;
L50:    f[0]+=f[0];
L51:    f[1]-=f[0];
L52:    f[0]*=Input003;
L53:    f[0]*=-1.427085638046265f;
L54:    f[0]-=Input000;
L55:    f[0]=sin(f[0]);
L56:    f[1]+=f[0];
L57:    f[0]+=Input000;
L58:    f[1]+=f[0];
L59:    f[0]-=Input002;
L60:    f[1]+=f[0];
L61:    f[0]-=f[0];
L62:    f[0]+=0.1756083965301514f;
L63:    f[0]*=f[0];
L64:    f[1]*=f[0];
L65:    f[0]-=f[1];
L66:    f[0]+=Input003;
L67:    f[0]*=0.7790718078613281f;
L68:    f[0]=sin(f[0]);
L69:    f[0]*=0.9177978038787842f;
L70:    f[0]*=0.9177978038787842f;
L71:
return f[0];
}
```

INTENTIONALLY BLANK

Appendix C

International, Peer-Reviewed Conference Paper

A Joint initiative of the AUSTRALIAN GEOMECHANICS SOCIETY and
NEW ZEALAND GEOTECHNICAL SOCIETY

11th Australia and New Zealand Young Geotechnical Professionals Conference

11YGPC

Queenstown, New Zealand, October 2016

PUBLISHED BY:

The New Zealand Geotechnical Society
PO Box 12241, Wellington, NZ

Proceedings of the 11th Australia and New Zealand Geotechnical Symposium.
Eds. Neeson, F.C; Lacey, D.; Buxton, D.; Storie, L.; Queenstown (2016)
ISBN 978-0-473-37653-6

Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC
Queenstown, New Zealand
25 – 28 October 2016

The 11YGPC Organising Committee would like to thank the following reviewers for the arduous task of undertaking reviews of papers presented in these Proceedings

TECHNICAL REVIEWERS

Gavin Alexander	David Airey
Kevin Anderson	Derek Arnott
Hayden Bowen	Tom Bowling
Andrew Campbell	Robert Day
Guy Cassidy	Mahdi Disfani
C Y Chin	Jim Hambleton
Stephen Crawford	Jinsong Huang
Fernando Della Pasqua	Richard Kaser
Tony Fairclough	Hadi Khabbaz
Debra Fellows	Burt Look
Eleni Gkeli	Peter Mitchell
Sally Hargraves	Mark Orr
Campbell Keepa	Derek Pennington
Andrew Langbein	Rick Piovesan
Don McFarlane	Mizanur Rahman
Rebecca McMahon	Diacho Sheng
Andy Mott	Ian Shipway
Charlie Price	Adrian Smith
Stuart Read	Mogana Sundaram
Ross Roberts	Abbas Taheri
Nick Rogers	Ramtin Tajeddin
Greg Saul	Megan Walske
Andrew Stiles	David Williams

The Graphic Design assistance provided by Karryn Muschamp is gratefully acknowledged.

Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC
Queenstown, New Zealand
25 – 28 October 2016

Application of Artificial Intelligence Techniques for Rolling Dynamic Compaction

R. A. T. M. Ranasinghe[1] and M. B. Jaksa[2]

[1]School of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia, 5005; email: tharanga.ranasinghe@adelaide.edu.au
[2]School of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia, 5005; PH (+61) 8 8313 4314; FAX (+61) 8 8313 4359; email: mark.jaksa@adelaide.edu.au

ABSTRACT

Rolling dynamic compaction (RDC), involving non-circular modules towed behind a tractor, is now widespread and accepted among many other soil compaction methods. However, to date, there is no accurate method to reliably predict the increase in soil strength after the application of a given number of passes of RDC. This paper presents the application of artificial intelligence (AI) techniques in the form of artificial neural networks (ANNs) and genetic programming (GP) for a priori prediction of the density improvement by means of RDC in a range of ground conditions. These AI-based models are developed by using in situ soil test data, specifically cone penetration test (CPT) and dynamic cone penetration (DCP) test data obtained from several ground improvement projects that employed the 4-sided, 8-tonne 'impact roller'. The predictions of ANN- and GP-based models are compared with the corresponding actual values and they show strong correlations ($r > 0.8$). Additionally, the robustness of the optimal models is investigated in a parametric study and it is observed that the model predictions are in a good agreement with the expected behaviour of RDC.

Keywords: rolling dynamic compaction, artificial neural networks, genetic programming

1 INTRODUCTION

Impact rolling, generically known as rolling dynamic compaction (RDC), involves heavy (6–12 tonnes) non-circular modules (3-, 4- and 5-sided), which rotate about their corners and fall to the ground when drawn behind a tractor. The square (4-sided) impact rolling module, which is the focus of this study, is shown in Figure 1. RDC is now widespread globally in the construction industry, as this technique provides an alternative to the traditional approaches of ground improvement, with superior compaction capabilities. As such, RDC is effective in that it has a greater influence depth – more than 1 m beneath the ground surface and sometimes as deep as 3 m in some soils (Avalle and Carter, 2005) – compared to conventional static and vibratory compaction, where the influence depths are generally less than 0.5 m (Clifford, 1976). As a result, thicker lifts, in excess of 0.5 m, can be employed, as compared to traditional compaction lifts of approximately 0.3 m, which enhances RDC's cost effectiveness. Moreover, RDC traverses the ground at a speed of 9–12 km/h, which is far more efficient than traditional compaction using a vibratory roller, which travels at a speed of 4 km/h (Pinard, 1999). As a consequence of these improved capabilities, RDC is utilised in many applications worldwide, particularly, (i) in the civil construction industry for in situ densification and subgrade proof-rolling, (ii) in the agricultural sector mainly for the improvement of existing water storages, channels and embankments, (iii) in the mining industry for the construction of tailing dams, rock rubblisation in open cut mine waste tips and the compaction of capping over waste rocks.

Figure 1. The 4-sided 'impact roller' towed behind a tractor

However, to date, there is no reliable method available to predict the effectiveness of RDC in advance. As a result, RDC is often adopted based on experience from previous work undertaken in similar soils and site conditions. In addition, field trials are usually undertaken prior to site works to ascertain the operational parameters, especially the optimal number of roller passes required to achieve the desired percentage of maximum dry density. Therefore, this study aims to develop an accurate and robust tool for predicting the performance of RDC in a range of ground conditions. This research makes use of artificial intelligence (AI) in the form of artificial neural networks (ANNs) and genetic programming (GP), which have also been proven successful in the broader geotechnical engineering context [e.g. (Alavi et al., 2013; Shahin et al., 2005)]. The developed models are validated against a set of unseen data and, in addition, ANN-and GP-based models are compared with each other over a range of performance measures. Additionally, a parametric study is carried out to assess the robustness of the optimal models. It is important to note that these are the very first models that permit a priori prediction of ground improvement as a result of RDC.

2 DATABASE

The data used in this study have been obtained from the results of several field trials undertaken by Broons (SA) Hire, an Australian company operating a range of ground improvement technologies, including RDC. Attention is given to the 4-sided, 8 tonne 'impact roller' (BH-1300). The database is comprised of in situ strength data in the form of cone penetration test (CPT) and dynamic cone penetrometer (DCP) test results with respect to the number of roller passes. In total, the database contains 1,755 records from 91 CPT soundings and 2,048 DCP records from 12 field projects.

In this study, two sets of AI models (i.e. two ANNs and two GP models) have been developed incorporating each of the CPT and DCP datasets for model calibration and validation. The models are developed to predict the degree of density improvement of the ground with respect to the number of roller passes. Therefore, a single output variable of cone tip resistance after compaction, q_{cf} (MPa) is adopted for the CPT models, whilst the average DCP blow count per 300 mm is incorporated in the DCP models. However, it is important that the input variables of these models effectively address the significant factors that influence soil behaviour as a consequence of RDC. Hence, the input and output variables involved in the model development and their ranges are presented in Table 1.

Table 1. Input/output variable ranges used in model development

Data	Model variables		Range
CPT	Input	Depth of measurement, D (m)	0.2 – 4.0
		Cone tip resistance prior to compaction, q_{ci} q□ci□ (MPa)	0.19 – 50.65
		Sleeve friction prior to compaction, f_{si} (kPa)	1.67 – 473.86
		No. of Roller Passes, P	5 – 40
	Output	Cone tip resistance after compaction, q_{cf} (MPa)	0.17 – 50.36
DCP	Input	Soil type	Sand–Clay, Clay–Silt, Sand–None, Sand–Gravel
		Average depth, D (m)	0.15 – 1.95
		Initial no. of roller passes, P_i	0 – 50
		Initial DCP count (blows/300 mm)	3 – 65
		Final no. of roller passes, P_f	2 – 60
	Output	Final DCP count (blows/300 mm)	2 – 84

3 ARTIFICIAL NEURAL NETWORKS (ANNS)

The ANN concept has emerged from knowledge of the functionality of the human brain and nervous system. ANNs are a data-driven approach and unlike the statistical modelling, they do not require prior knowledge of the underlying relationships among the variables. Besides, the functional relationships between inputs and outputs are learnt by the ANNs from a set of example data.

3.1 Development of ANN models

ANN modelling is carried out using the PC-based software, *Neuframe* version 4.0 (Neuframe, 2000). As described in §2, the model incorporating CPT data consists of 4 inputs along with a single output, whilst the model involving DCP data consists of 5 inputs and a single output. In this study, the cross-

Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC
Queenstown, New Zealand
25 – 28 October 2016

validation technique (Stone, 1974) is used as the stopping criterion, which involves the division of the available dataset into three subsets: training, testing and validation. The training subset is used for model calibration, which involves optimisation of the connection weights in the network. With the testing set, model performance is periodically assessed during training, whereas the network is validated against the independent validation set once the model has been optimised. Data division is carried out in such a way that the training and validation sets contain 80% and 20% of the total data, respectively. The training set is further divided into two subsets; 80% for training and 20% for testing.

In this study, multi-layer perceptron (MLP) models are developed with the use of the back-propagation error method (Rumelhart et al., 1986). The selected MLP architecture is comprised of 3 layers: the input layer, one hidden layer and the output layer. The number of nodes in the input and output layers represent the number of model inputs and outputs and thus, the CPT- and DCP-based models consist of 4 and 5 nodes, respectively in the input layer, whilst both models contain a single node in the output layer. However, the optimisation of the nodes in the hidden layer is crucial so that the structure is neither too complex nor too simple, but appropriate enough to capture accurately the non-linearity of the relationships between the input and output parameters. Therefore, a stepwise, trial-and-error approach is adopted to achieve the optimal network architecture, where a number of ANN models are trained, beginning from the simplest form with a single hidden layer node model and successively increasing the number of nodes. As suggested by Caudill (1988), $2n + 1$ is the upper limit of hidden nodes for a network to map any continuous function, with n being the number of input nodes. Accordingly, the CPT- and DCP-based models are tested to a maximum number of hidden nodes of 9 and 11, respectively. The ANNs, with each trial number of hidden nodes, are initially trained having assigned the default software values to the internal parameters (i.e. learning rate = 0.2, momentum term = 0.8) and the sigmoidal transfer function is used for both the hidden and output layers. However, after determining the best topology, the network with the optimal number of hidden nodes is retrained with different combinations of internal parameters, specifically with different learning rates and momentum terms. After model training, model performance is assessed in terms of the coefficient of correlation (*R*), root mean square error (*RMSE*) and mean absolute error (*MAE*).

4 GENETIC PROGRAMMING (GP)

Genetic programming (GP) is an evolutionary computational approach inspired from biological evolution based on Darwinian Theory. The technique was first introduced by Koza (1992) as an extension to the genetic algorithm (GA). In GP, the computer programs are called individuals, which have been made up of a set of functions, variables and constants. In accordance with a GP algorithm, these individuals are automatically evolved into a solution for a given particular problem. In this study, a particular, robust variant of GP, namely linear genetic programming (LGP), is used. In contrast to the conventional tree-based GP approach, the data flow of evolved programs in LGP has a more general, register-based graph representation at the functional level. Additionally, the programs in LGP evolve either in an imperative programming language (e.g. C, C++) (Brameier and Banzhaf, 2007) or directly in machine language (Nordin, 1994). In brief, the LGP algorithm (Brameier and Banzhaf, 2007) involves:

i. *Initialising a population of randomly generated programs and evaluation of their fitness value;*
ii. *Running a tournament and selection of the winning programs;*
iii. In this step, 4 randomly selected programs are subjected to tournament selection, where two programs are selected as the winners based on their fitness.
iv. *Transforming the winner programs;*
v. The two selected winner programs are copied and transformed probabilistically into offsprings subjected to genetic operations; i.e. crossover and mutation.
vi. *Replacing the tournament losing programs with the offspring programs; and*
vii. *Repeating Steps 2 to 4 until the termination or convergence criteria are satisfied.*

4.1 Development of LGP models

LGP modelling utilises the same subsets of CPT and DCP data that are employed in the ANN model development, with the intention of conducting a fair comparison between the model performances. The commercially available software *Discipulus* version 5.2 (Francone, 2010) is used for the necessary computations. The selection of control parameters is considered to be vital in LGP modelling, since it has a direct impact on the model's generalisation capacity. In this modelling, the

control parameters are defined in accordance with the recommended values by previous similar LGP applications and also depending on the observations from preliminary runs. As presented in Table 2, several different parameter combinations, in terms of population size, number of demes and crossover rate, are investigated, whilst most of the other minor parameters are set to the software default values.

Table 2. Parameter settings used for the LGP model development

Parameter	Settings
Function set	+, –, ×, /, Absolute, Square Root, Trigonometric (*sin, cos*), Exponential
Population size	500, 1,000, 2,000, 5,000, 7,500, 10,000
Number of demes	10, 20
Program size	Initial = 80 bytes, Maximum = 512 bytes
Mutation frequency	95%
Crossover frequency	50%, 90%

A relatively large number of LGP projects are carried out, where all of the above listed combinations of parameters are tested and replicated 5 times for each setting, in order to address different random initial conditions. Each LGP project is made up of a series of runs, which begins in short and successively progresses in length during the course of a project. Within a run, large numbers of genetic programs are evolved and they are monitored for minimum error, as this study uses the mean square error (*MSE*) as the fitness function. A LGP run is terminated after a reasonable number of generations have evolved without improvement in terms of *MSE*. Finally, a LGP project is terminated given a reasonable time to evolve into an accurate model and when no further improvement in model performance is likely to occur.

5 RESULTS AND DISCUSSION

The following sections summarise the results of the optimal models, along with details of parametric study.

5.1 Performance analysis

The developed ANN- and LGP-based models are evaluated and compared using the performance measures as discussed above; i.e. *R*, *RMSE* and *MAE* with respect to each of the 3 data subsets. In regards to the ANN modelling, with respect to both the CPT and DCP datasets, the models with 4 nodes in the hidden layer were found to be the optimal. The results of these optimal models, along with the optimal LGP results, are presented in Table 3.

Table 3. Comparison of the performance statistics of the optimal ANN- and LGP-based models

Data	Model	*R*			*RMSE*+			*MAE*+		
		T	S	V	T	S	V	T	S	V
CPT	ANN	0.87	0.87	0.86	4.19	4.33	4.16	2.89	3.03	2.93
	LGP	0.87	0.88	0.87	4.05	4.08	4.03	2.72	2.73	2.71
DCP	ANN	0.85	0.83	0.79	6.45	6.52	7.54	4.88	4.74	5.59
	LGP	0.84	0.87	0.81	6.22	5.35	6.80	4.18	3.70	4.74

T: Training, S: Testing, V: Validation
+ For CPT models units is MPa and for DCP models units is blows/300 mm

As can be observed, the performances of the obtained optimal models, based on the CPT and DCP datasets, are very good given the strong correlation coefficient [i.e. $R > 0.8$ (Smith, 1993)] and relatively low error values (i.e. *RMSE* and *MAE*). Therefore, it can be concluded that both the ANN and LGP models developed in this study have the capability of predicting the effectiveness of RDC to a high degree of accuracy. Although both the ANN and LGP models exhibit similar performance, it is also apparent that the LGP approach slightly outperforms the ANN technique.

5.2 Parametric study

In order to further explore the accuracy and the validity of the selected optimal models, a parametric study is carried out. The model's generalisation capability is examined so that the model behaviour conforms to the known physical behaviour of the system. This involves investigating the model response to a new set of unseen synthetic input data, where only a single input parameter is varied at

Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC
Queenstown, New Zealand
25 – 28 October 2016

a time, while the others are kept constant at a pre-defined value. It is essential that the input variables are varied between the lower and upper bounds of the training dataset, since the ANN- and LGP-based models perform best when interpolating rather than extrapolating.

For the ANN and LGP models incorporating CPT data, the output variable, q_{cf}, is examined, while the input variables of q_{ci}, f_{si}, P and D are varied. For instance, as illustrated in Figures 2 and 3, the predicted value of q_{cf} is studied for both the ANN and LGP models, respectively, by varying the input variables, q_{ci} = 2, 5 MPa, f_{si} = 50, 100 kPa and P = 10, 20, 30, 40. As can be seen, q_{cf} consistently improves as the number of passes, P, increases from 10 to 40, while q_{ci} and f_{si} remain constant at 2 or 5 MPa and 50 or 100 kPa, respectively. This indicates that the soil strength at a particular location is improved beyond the initial strength with successive roller passes. However, when comparing Figure 2(a) and (b) or Figure 3(a) and (b), it is observed that q_{cf} improves only marginally as f_{si} increases from 50 to 100 kPa, while q_{ci} remains constant either at 2 or 5 MPa. Hence, this suggests that f_{si} is less influential on q_{cf}. Nevertheless, it is evident from the parametric study that the distinct non-linear relationship between q_{cf} and P has been appropriately captured by these models.

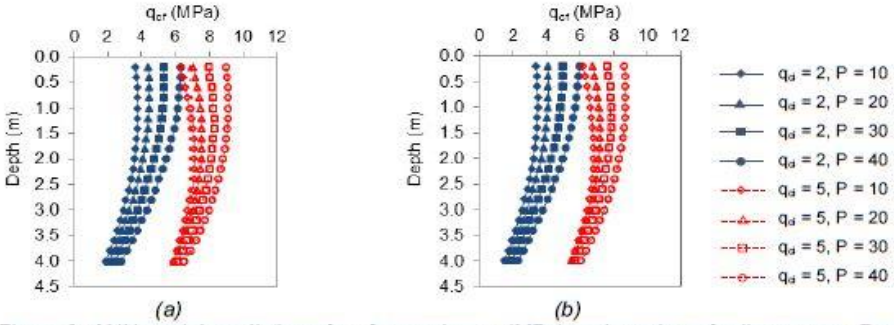


Figure 2. ANN model prediction of q_{cf} for varying q_{ci} (MPa) and number of roller passes, P when: (a) f_{si} = 50 kPa; and (b) f_{si} = 100 kPa

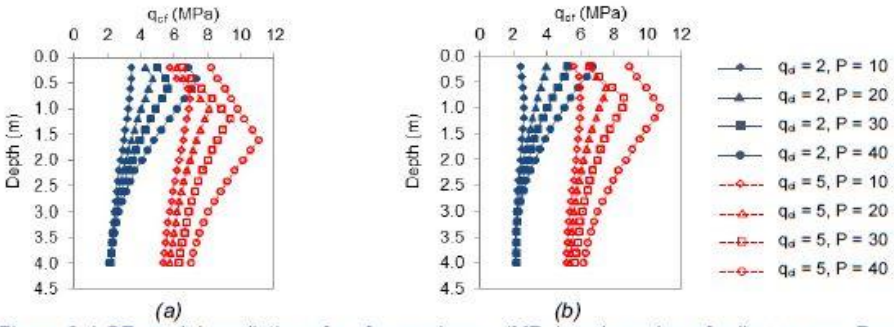


Figure 3. LGP model prediction of q_{cf} for varying q_{ci} (MPa) and number of roller passes, P when: (a) f_{si} = 50 kPa; and (b) f_{si} = 100 kPa

In a similar manner, the optimal ANN- and LGP-based models for DCP data are also investigated in a parametric study. Here, the post-compaction condition of the ground, represented by the final DCP blow count, is predicted from the optimal ANN and LGP models for a given initial DCP blow count (i.e. 5 and 15 blows/300 mm), in each of the different soil types (i.e. Sand–Clay, Clay–Silt, Sand–None and Sand–Gravel) and different numbers of roller passes (i.e. 5, 10, 15, 20, 30, 40). The resulting ANN and LGP model predictions are presented in Figures 4 and 5, respectively. It is observed that the final DCP blow count increases with increasing number of roller passes for a given initial DCP blow count in each of the soil types. It is evident from these results that the ground is significantly improved as a consequence of RDC. Moreover, this confirms the model predictions from the parametric study are in good agreement with the expected behaviour of RDC compaction. Therefore, it can be concluded that the optimal ANN and LGP models are robust when predicting the effectiveness of RDC and thus, they can be used with confidence.

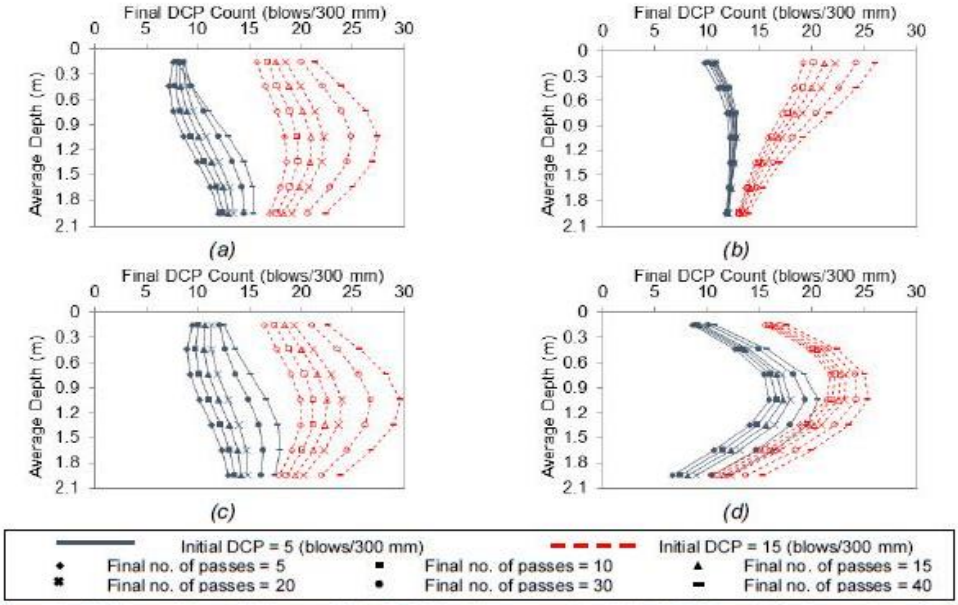


Figure 4. ANN model predictions of final DCP with respect to initial DCP and final number of roller passes in (a) Sand–Clay (b) Clay–Silt (c) Sand–None (d) Sand–Gravel

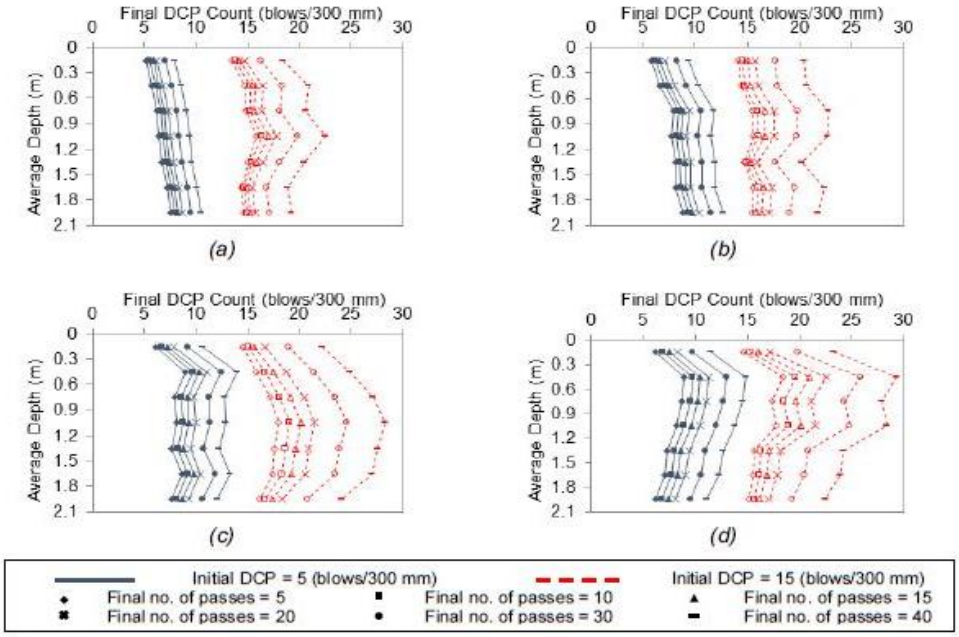


Figure 5. LGP model predictions of final DCP with respect to initial DCP and final number of roller passes in (a) Sand–Clay (b) Clay–Silt (c) Sand–None (d) Sand–Gravel

Proceedings of the 11th ANZ Young Geotechnical Professionals Conference – 11YGPC
Queenstown, New Zealand
25 – 28 October 2016

6 CONCLUSION

This paper investigates the effectiveness of rolling dynamic compaction (RDC) on different soil types and presents novel and unique predictive models based on artificial intelligence (AI) techniques in the form of artificial neural networks (ANNs) and linear genetic programming (LGP). These models incorporate an extensive database of ground density data in terms of cone penetration test (CPT) and dynamic cone penetrometer (DCP) test results associated with the Broons 4-sided, 8 tonne 'impact roller'. A total of 4 models have been developed: two involving ANNs – one for the CPT and the other for the DCP; and two LGP models – again, one for the CPT and the other for the DCP. The resulting optimal ANN- and LGP-based models yield high accuracy of model predictions, with a high coefficient of correlation ($R > 0.8$) and with lower error values, i.e. root mean square error (*RMSE*) and mean absolute error (*MAE*) when validated against a set of unseen data. The results indicate that the LGP-based models slightly outperform their ANN counterparts and overall produce slightly more accurate predictions. In addition, a parametric study has been carried out to assess the generalisation ability and robustness of these optimal models. The results of the parametric study demonstrate that the response of the models agrees well with the expected physical relationships among the input and output parameters. Therefore, it can be concluded that the developed models are both robust and reliable when forecasting the performance of RDC in various ground conditions. The models developed in this study are intended to provide initial predictions for planning purposes and may replace, or at the very least augment, the necessity for field trials prior to full-scale construction, which in turn contributes to significant cost savings.

7 ACKNOWLEDGEMENTS

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP120101761). The authors wish to acknowledge Mr. Stuart Bowes at Broons Hire (SA) Pty Ltd for his kind assistance and continuing support, especially in providing access to the in situ test results upon which the numerical models are based. The authors are also grateful to colleagues Drs. Yin Lik Kuo and Fereydoon Pooya Nejad and Mr. Brendan Scott for their contribution to this work.

REFERENCES

Alavi, A. H., Gandomi, A. H., Mollahasani, A., Bazaz, J. B. and Talatahari, S. (2013). Linear and tree-based genetic programming for solving geotechnical engineering problems. *Metaheuristics in Water, Geotechnical and Transport Engineering*, 289-310.

Avalle, D. L. and Carter, J. P. (2005). *Evaluating the improvement from impact rolling on sand*. Paper presented at the 6th International Conference on Ground Improvement Techniques, Coimbra, Portugal; .

Brameier, M. F. and Banzhaf, W. (2007). *Linear genetic programming*: Springer Science & Business Media.

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

Clifford, J. M. (1976). *Impact rolling and construction techniques*. Paper presented at the Australian Road Research Board (ARRB) Conference, Perth, Australia; .

Francone, F. (2010). Discipulus TM with Notitia and solution analytics owner's manual. Register Machine Learning Technologies Inc., Littleton, CO, USA.

Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection* (Vol. 1): MIT press.

Neuframe. (2000). Neurosciences. Neuframe Version 4.0. . Southampton, Hampshire: Neurosciences Corp.

Nordin, P. (1994). A compiling genetic programming system that directly manipulates the machine code *Advances in genetic programming* (Vol. 1 pp. 311-331). MIT press, USA: Kenneth E. and Jr. Kinnear Editors.

Pinard, M. I. (1999). *Innovative developments in compaction technology using high energy impact compactors*. Paper presented at the 8th Australia New Zealand Conference on Geomechanics: Consolidating Knowledge, Hobart, Australia.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). Learning representations by backpropagating errors. *Nature, 323*(1), 533-536.

Shahin, M. A., Jaksa, M. B. and Maier, H. R. (2005). Neural network based stochastic design charts for settlement prediction. *Canadian Geotechnical Journal, 42*(1), 110-120.

Smith, M. (1993). *Neural networks for statistical modeling*. New York: Thomson Learning.

Stone, M. (1974). *Cross-Validatory Choice and Assessment of Statistical Predictions*. 36(2), 111-147.