Journal of Rock Mechanics and Geotechnical Engineering 11 (2019) 815–823

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.rockgeotech.org

Full Length Article

Genetic programming for predictions of effectiveness of rolling dynamic compaction with dynamic cone penetrometer test results

R.A.T.M. Ranasinghe*, M.B. Jaksa, F. Pooya Nejad, Y.L. Kuo

School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, 5005, Australia

ARTICLE INFO

Article history:
Received 20 May 2018
Received in revised form
4 September 2018
Accepted 23 October 2018
Available online 26 March 2019

Keywords:
Ground improvement
Rolling dynamic compaction (RDC)
Linear genetic programming (LGP)
Dynamic cone penetrometer (DCP) test

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. For this, in this context, a prediction model is developed based on linear genetic programming (LGP), which is one of the common approaches in application of artificial intelligence for nonlinear forecasting. The model is based on in situ density-related data in terms of dynamic cone penetrometer (DCP) results obtained from several projects that have employed the 4-sided, 8-t impact roller (BH-1300). It is shown that the model is accurate and reliable over a range of soil types. Furthermore, a series of parametric studies confirms its robustness in generalizing data. In addition, the results of the comparative study indicate that the optimal LGP model has a better predictive performance than the existing artificial neural network (ANN) model developed earlier by the authors.

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 widely used globally. This technique involves towing a heavy (6–12 t), non-circular (3-, 4- and 5-sided) module behind a tractor, where the module rotates about its corners as it is drawn forward (Avalle, 2004). 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 in 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 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 surface. 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 9–12 km/h, whereas the traditional compaction rollers travel at 4 km/h speed (Pinard, 1999). Furthermore, RDC can also treat thicker lifts, i.e. in excess of 0.5 m whilst, with the conventional compaction, the lift thickness is usually limited between 0.2 m 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), agricultural sector (Avalle, 2004), and mining applications (Scott and Jaksa, 2012).

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

* Corresponding author.
E-mail address: tara.ranasinghe@gmail.com (R.A.T.M. Ranasinghe).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

https://doi.org/10.1016/j.jrmge.2018.10.007

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 basically 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 it 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. Similar to artificial neural networks (ANNs), genetic programming (GP) can be considered as another alternative approach to conventional methods, due to its ability to approximate any linear/nonlinear relationship among a set of observed input and output data in the absence of prior knowledge of the underlying mechanisms of the system. Recently, the authors suggested an approach based on ANNs (Ranasinghe et al., 2017a, 2019), 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, where LGP models were developed using CPT data (Ranasinghe et al., 2017b), have suggested that LGP outperforms ANNs, in addition to other benefits.

The developed LGP-based model in this study utilizes a reliable database consisting of DCP results obtained from several ground improvement projects, associated with the Broons 4-sided, 8-t 'impact roller'. Since this dataset has also been employed previously in the ANN model developed recently by the authors (Ranasinghe et al., 2017a), a comparative study is conducted between the results obtained herein and 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.

2. Linear genetic programming (LGP)

GP is an evolutionary computational approach of nonlinear modeling, where the computer programs evolve automatically to optimize a solution towards a pre-defined goal. This 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. The GAs are often recognized 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 are dynamically varied during evolution and are usually executable without post-processing (Koza, 1992). Moreover, GAs are generally used for parameter optimization, 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 is 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. LGP and graph-based GP (Banzhaf et al., 1998; Poli et al., 2007; Alavi et al., 2013). In the present study, emphasis is placed on 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, 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 to 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., 2010). 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 as the interpretation of internal program representation. However, in AIMGP, the individuals are directly executable by the processor, which 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 utilized in this study.

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

(1) Initializing a population of randomly generated programs and evaluating their fitness;
(2) Running a tournament and selecting the winning programs;
(3) Transforming the winner programs into offsprings probabilistically subjected to genetic operations, i.e. crossover and mutation;
(4) Replacing the tournament losing programs with the offspring programs; and
(5) Repeating Steps 2–4 until the termination or convergence criteria are satisfied.

3. Methodology

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

3.1. Database, data analysis and data pre-processing

This study utilizes a comprehensive database that comprises in situ strength data in the form of DCP test results. The DCP (ASTM D6951-03, 2003) is one of the most commonly used in situ test methods, which provides an indication of soil strength in terms of rate of penetration (blow/mm) (Mousavi et al., 2018). 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-t 'impact roller' (BH-1300), which is operated by the Australian company Broons (SA) Hire. In total, the database contains 2048 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 from 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, and 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 behavior 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 (blow/300 mm), whilst the single output variable is described by the final DCP count (blow/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 numbers of roller passes corresponding to the initial and final densities 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 generalized 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, four distinct soil types are characterized 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 the LGP-based model development and their statistics are presented in Table 1, with Fig. 1 summarizing the histograms of the model input and output variables.

Prior to model development, the entire dataset is subdivided into two subsets, i.e. 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 generalization capabilities. Since the optimal model is evaluated with respect to an unseen dataset, 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., 2017a), the same three 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 (1310 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 three 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 2, the subsets effectively represent the same population by the similar values of mean, standard deviation, minimum, maximum and range.

3.2. LGP-based modeling 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 modeling, since it has a direct impact on the model's generalization capacity. Therefore, in this study, the control parameters are defined in accordance with the suggested values from similar LGP applications (Gandomi et al., 2010, 2014; Rashed et al., 2012; Alavi et al., 2013; Babanajad et al., 2013; Alavi and Sadrossadat, 2016) and also from observations obtained from preliminary runs. As presented in Table 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.

In this study, several LGP projects are carried out including only the arithmetic functions (+, −, ×, /). Furthermore, in order to permit the evolution of highly nonlinear models, inclusion of mathematical functions (sin, cos, exponential, absolute and

Table 1
Descriptive statistics of the dataset used in LGP model development.

Data type	Variable	Statistical parameters					
		Mean	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
Input	Average depth, D (m)	0.82	0.52	0.15	1.95	0.33	−0.74
	Initial number of roller passes	7.89	10.65	0	50	1.63	2.06
	Initial DCP count (blow/300 mm)	16.41	10.69	3	65	1.51	2.64
	Final number of roller passes	21.13	16.25	2	60	1.01	−0.03
Output	Final DCP count (blow/300 mm)	18.14	11.25	2	84	1.42	2.94

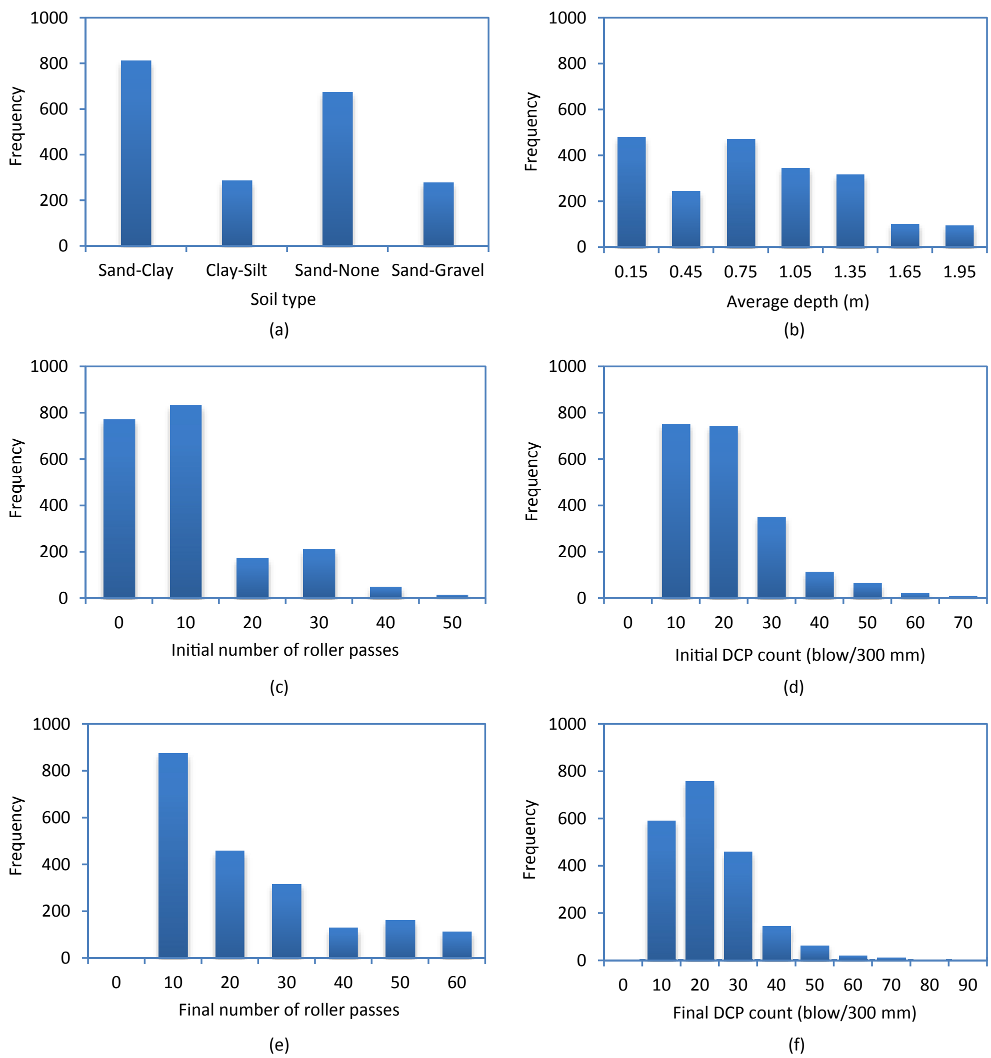


Fig. 1. Histograms of the model variables used in LGP model development: (a) soil type, (b) average depth, (c) initial number of roller passes, (d) initial DCP count, (e) final number of roller passes, and (f) final DCP count.

square root), in addition to basic arithmetic operators, is also considered. This study applies the mean square error (MSE) as the fitness measure, which can be defined as

$$MSE = \frac{\sum_{i=1}^{n}(h_i - t_i)^2}{n} \tag{1}$$

where h_i and t_i are the actual and the model predicted output values for the ith sample, respectively; and n denotes the number of samples.

The population size parameter is regulated at several different levels, i.e. 500, 1000, 2000, 2500 and 5000. 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

Table 2
Statistical properties of the data subsets.

Data type	Model variable	Data subset	Statistical parameters				
			Mean	Standard deviation	Minimum	Maximum	Range
Input	Average depth (m)	Training	0.81	0.51	0.15	1.95	1.8
		Testing	0.82	0.51	0.15	1.95	1.8
		Validation	0.83	0.52	0.15	1.95	1.8
	Initial number of roller passes	Training	7.69	10.61	0	50	50
		Testing	7.65	10.44	0	50	50
		Validation	8.71	10.93	0	50	50
	Initial DCP count (blow/300 mm)	Training	16.57	10.86	3	65	62
		Testing	15.88	10.64	3	59	56
		Validation	16.31	10.2	3	61	58
	Final number of roller passes	Training	21.14	16.25	2	60	58
		Testing	21.16	16.49	2	60	58
		Validation	21.08	16.11	2	60	58
Output	Final DCP count (blow/300 mm)	Training	18.3	11.29	2	84	82
		Testing	17.8	10.81	2	73	71
		Validation	17.93	11.47	3	75	72

Table 3
Parameter setting used in LGP modeling.

Parameter	Settings
Function set	+, −, ×, /, absolute, square root, trigonometric (sin, cos), exponential
Population size	500, 1000, 2000, 5000
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%

these genetic operations are considered at two levels of 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, i.e. 128 and 256 bytes.

Likewise, in this study, many numbers of LGP projects are carried out, where all 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 h 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 three 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).

4. Results and discussion

The following sections summarize 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 (blow/300 mm) is presented in computing code in the C language in Appendix A.

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 three data subsets (training, testing and validation), are presented in Table 4. The selected model's performance and reliability are assessed based on the criteria suggested by Smith (1986), as presented in the following.

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

(1) $|R| \geq 0.8$, there exists a strong correlation;
(2) $0.2 < |R| < 0.8$, there exists a correlation; and
(3) $|R| \leq 0.2$, there exists a weak correlation between the two variables.

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 criteria are 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. This implies that the model predicts the target values accurately and also incorporates a generalization capability.

Fig. 2 presents the 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 two standard error (SE) envelops of 0.5–2 times the

Table 4
Performance statistics of the selected optimal LGP model.

Data subset	Performance criteria		
	R	RMSE (blow/300 mm)	MAE (blow/300 mm)
Training	0.84	6.22	4.18
Testing	0.87	5.35	3.7
Validation	0.81	6.8	4.74

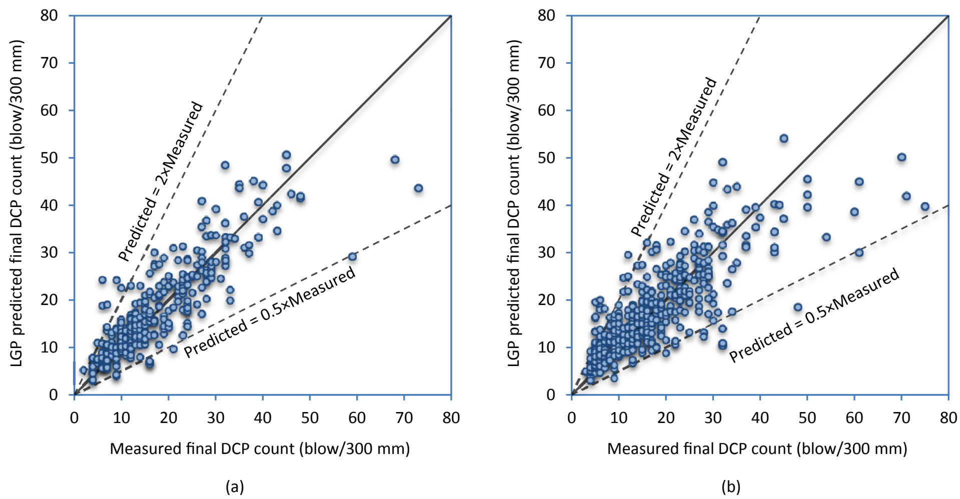


Fig. 2. Measured versus LGP predicted final DCP counts with respect to (a) testing dataset and (b) validation dataset.

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.

In order to investigate further the model's performance, the LGP predictions and the measured data, with respect to the validation dataset, are assessed subject to several additional measures, as suggested in the literature. Table 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 those obtained from the ANN-based model recently developed by the authors (Ranasinghe et al., 2017a). The performance indices in terms of R, RMSE and MAE are again adopted and the results are presented in Table 6. As can be seen, the results for both models are very promising, as indicated by the strong coefficient of correlation ($R \geq 0.8$), along with the relatively low error values with respect to all three data subsets. However, it is also evident that the LGP predictions are slightly superior to those from the ANN model.

Further to the above assessment, the predictions of final DCP count from the LGP and ANN models are compared graphically with the measured final DCP count values, with respect to the testing and validation datasets, and the resulting histograms are shown in Fig. 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 generalization 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.

Table 6
Comparison of the performance statistics of LGP and ANN models.

Data subset	Performance criteria					
	R		RMSE (blow/300 mm)		MAE (blow/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	3.7	4.74
Validation	0.81	0.79	6.8	7.54	4.74	5.59

4.2. Parametric study

Although the LGP model yields satisfactory performance in terms of the measures as discussed above, it is essential to investigate the model's behavior 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 dataset which lies within the range that model is trained against, in order to examine whether the results conform to the known physical behavior of the system. Each of the model input variables is varied successively, while maintaining all other input

Table 5
Additional performance measures of the LGP model for the validation dataset.

Item	Formula	Source	Condition	Result
1	$k = \dfrac{\sum_{i=1}^{n}(h_i t_i)}{\sum_{i=1}^{n} h_i^2}$	Golbraikh and Tropsha (2002)	$0.85 < k < 1.15$	0.92
2	$k' = \dfrac{\sum_{i=1}^{n}(h_i t_i)}{\sum_{i=1}^{n} t_i^2}$	Golbraikh and Tropsha (2002)	$0.85 < k' < 1.15$	0.98
3	$R_0^2 = 1 - \dfrac{\sum_{i=1}^{n}(t_i - h_i^0)^2}{\sum_{i=1}^{n}(t_i - \overline{t_i})^2}$, where $h_i^0 = kt_i$	Roy and Roy (2008)	Should be close to 1	0.97
4	$R_0'^2 = 1 - \dfrac{\sum_{i=1}^{n}(h_i - t_i^0)^2}{\sum_{i=1}^{n}(h_i - \overline{h_i})^2}$, where $t_i^0 = k' h_i$	Roy and Roy (2008)	Should be close to 1	1
5	$m = \dfrac{R^2 - R_0^2}{R^2}$	Golbraikh and Tropsha (2002)	$m < 0.1$	−0.49
6	$n = \dfrac{R^2 - R_0'^2}{R^2}$	Golbraikh and Tropsha (2002)	$n < 0.1$	−0.54

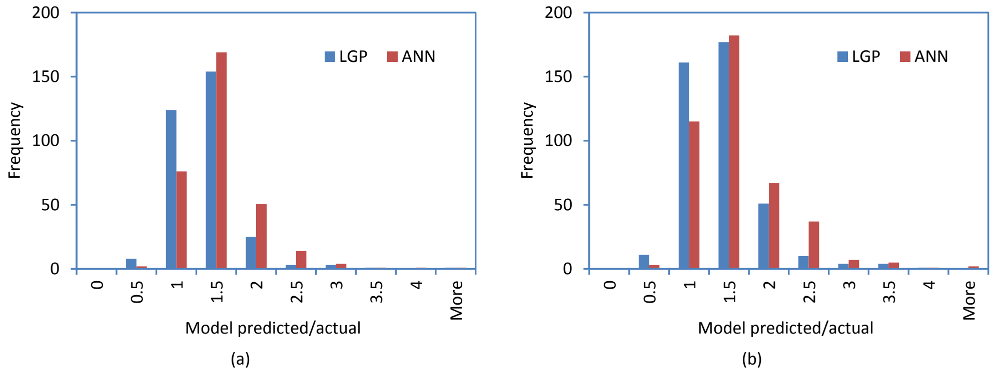


Fig. 3. Frequency histograms for model predicted and measured final DCP counts with respect to (a) testing data and (b) validation data.

variables at a pre-defined value. In this case, the model output of final DCP count (blow/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 blow/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 Fig. 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 behavior demonstrates that the ground is significantly improved with the application of RDC, as evidenced by the increasing DCP count, which is consistent with the behavior observed in 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).

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 utilized in a number of applications (Gandomi et al., 2010; 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, and the 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 in percentage) that each input variable appears, in the 30 best 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). In addition, the average and maximum

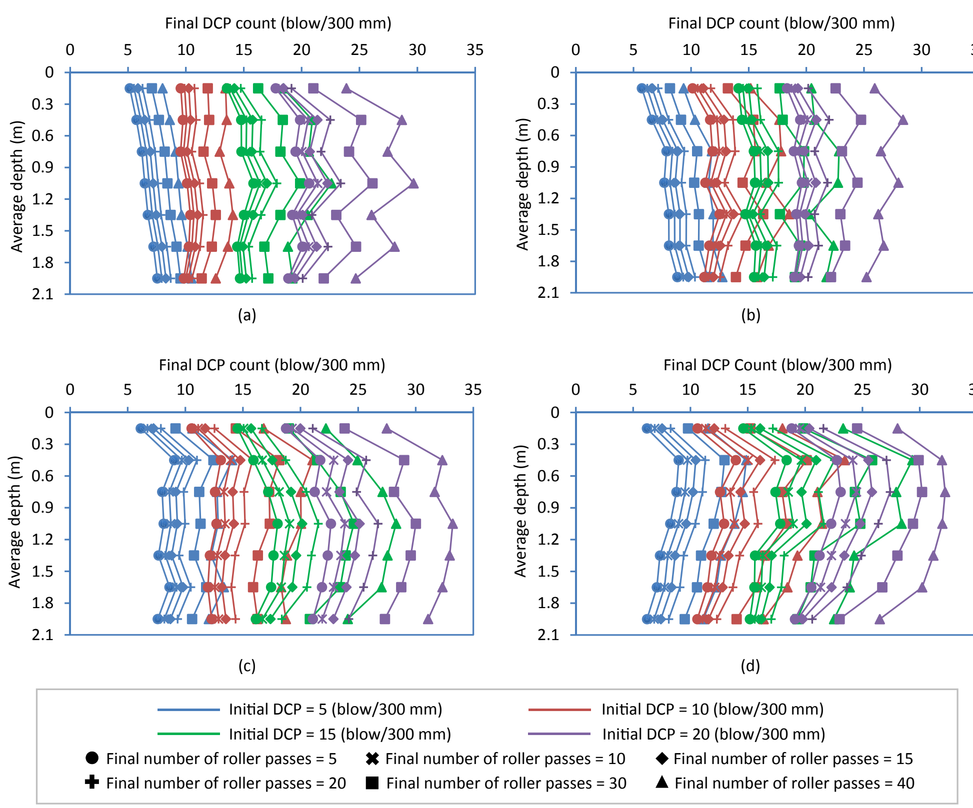
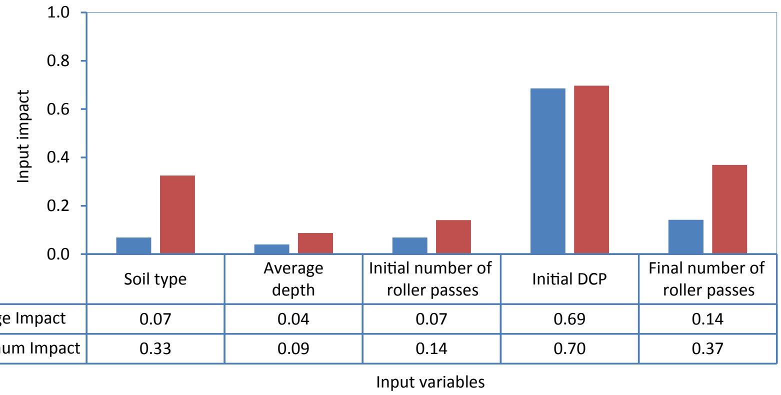


Fig. 4. Variation of final DCP count with respect to initial DCP count and final number of roller passes in (a) sand-clay, (b) clay–silt, (c) sand-none, and (d) sand-gravel.



	Soil type	Average depth	Initial number of roller passes	Initial DCP	Final number of roller passes
Average Impact	0.07	0.04	0.07	0.69	0.14
Maximum Impact	0.33	0.09	0.14	0.70	0.37

Fig. 5. Contribution of the input variables to optimal LGP model predictions.

impact values with respect to each of the input variables, which describes the average and maximum effects that result from the removal of each corresponding input variable from the 30 best programs, are also calculated and the resulting histogram is presented in Fig. 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.

A similar sensitivity analysis has been conducted during the ANN modeling phase, as described by Ranasinghe et al. (2017a). However, in relation to the ANN models, the input variables of soil type and initial DCP count were found to be the most important, whilst the relative importance of the input variables – average depth, initial number of roller passes and final number of roller passes – was reduced in turn. Although the relative importance of each variable is inconsistent when comparing the ANN and LGP models, it is evident that the selected input variables have a substantial effect on the model predictions.

5. Summary and conclusions

This paper presents a new approach based on GP for the predictions of the efficacy of RDC, which is considered to be an alternative to conventional soil compaction technology. A particular variant of GP, i.e. LGP, is used to develop the model. A comprehensive database consisting of in situ density data in terms of DCP test results is utilized, associated with various ground improvement project records involving the Broons' 4-sided, 8-t 'impact roller'. In model development, five input variables, i.e. soil type, average depth (m), initial number of roller passes, initial DCP count (blow/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 (blow/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 RMSE of 6.8 (blow/300 mm), and an MAE of 4.74 (blow/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 generalization 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 to the final model predictions. The results indicate that the input variables utilized in this study are significant with respect to the predictions of the final DCP count resulting from the application of RDC. The optimal LGP program is included in a C code in order to disseminate the model and facilitate its use in practice.

The LGP model presented in this study 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 predesign 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 incorporate a limited number of soil types. Moreover, this study has focused solely on the 4-sided, 8-t impact roller, and as such, the developed predictive models are valid only for that specific RDC module (BH-1300).

Conflicts of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project No. 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrmge.2018.10.007.

References

Alavi AH, Gandomi AH, Mollahasani A, Bazaz JB, Talatahari S. Linear and tree-based genetic programming for solving geotechnical engineering problems. In: Metaheuristics in water, geotechnical and transport engineering. 1st ed. London: Elsevier; 2013.

Alavi AH, Gandomi AH. Energy-based numerical models for assessment of soil liquefaction. Geoscience Frontiers 2012;3(4):541–55.

Alavi AH, Sadrossadat E. New design equations for estimation of ultimate bearing capacity of shallow foundations resting on rock masses. Geoscience Frontiers 2016;7(1):91–9.
ASTM D6951-03. Standard test method for use of dynamic cone penetrometer in shallow pavement applications. West Conshohocken, PA, USA: American Society for Testing and Materials (ASTM); 2003.
Avalle DL, Carter JP. Evaluating the improvement from impact rolling on sand. In: Proceedings of the 6th international conference on ground improvement techniques. Coimbra, Portugal; 2005.
Avalle DL. Reducing haul road maintenance costs and improving tyre wear through the use of impact rollers. In: Mining for tyres conference. Perth Australia; 2006.
Avalle DL. Use of the impact roller to reduce agricultural water loss. In: Proceedings of the 9th ANZ conference on Geomechanics. Auckland Australia; 2004. p. 513–8.
Babanajad SK, Gandomi AH, Mohammadzadeh SD, Alavi AH. Numerical modeling of concrete strength under multiaxial confinement pressures using linear genetic programming. Automation in Construction 2013;36:136–44.
Banzhaf W, Nordin P, Keller RE, Francone FD. Genetic programming: an introduction on the automatic evolution of computer programs and its application. San Francisco, USA: Morgan Kaufmann; 1998.
Brameier M, Banzhaf W. A comparison of linear genetic programming and neural networks in medical data mining. IEEE Transactions on Evolutionary Computation 2001;5(1):17–26.
Brameier MF, Banzhaf W. Linear genetic programming. New York, USA: Springer Science & Business Media; 2007.
Clegg B, Berrangé AR. The development and testing of an impact roller. The Civil Engineer in South Africa 1971;13:65–73.
Clifford JM. Impact rolling and construction techniques. In: Australian road research board conference; 1976. p. 21–9.
Francone FD, Deschaine LM. Extending the boundaries of design optimization by integrating fast optimization techniques with machine-code-based, linear genetic programming. Information Sciences 2004;161(3–4):99–120.
Francone FD. Discipulus TM with Notitia and solution analytics owner's manual. Littleton, CO, USA: Register Machine Learning Technologies Inc.; 2010.
Gandomi AH, Alavi AH, Sahab MG. New formulation for compressive strength of CFRP confined concrete cylinders using linear genetic programming. Materials and Structures 2010;43(7):963–83.
Gandomi AH, Mohammadzadeh SD, Pérez-Ordóñez JL, Alavi AH. Linear genetic programming for shear strength prediction of reinforced concrete beams without stirrups. Applied Soft Computing 2014;19:112–20.
Golbraikh A, Tropsha A. Beware of q2! Journal of Molecular Graphics and Modelling 2002;20(4):269–76.
Holland JH. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. Oxford, UK: University of Michigan Press; 1975.
Jaksa MB, Scott BT, Mentha NL, Symons AT, Pointon SM, Wrightson PT, Syamsuddin E. Quantifying the zone of influence of the impact roller. In: ISSMGE-TC 211 international symposium on ground improvement; 2012. p. 41–52.
Koza JR. Genetic programming: on the programming of computers by means of natural selection. Cambridge, UK: The MIT Press; 1992.
Mousavi SH, Gabr MA, Borden RH. Correlation of dynamic cone penetrometer index to proof roller test to assess subgrade soils stabilization criterion. International Journal of Geotechnical Engineering 2018;12(3):284–92.
Mousavi SM, Alavi AH, Mollahasani A, Gandomi AH. A hybrid computational approach to formulate soil deformation moduli obtained from PLT. Engineering Geology 2011;123(4):324–32.
Nordin P. A compiling genetic programming system that directly manipulates the machine code. Advances in Genetic Programming 1994;1:311–31.
Pinard MI. Innovative developments in compaction technology using high energy impact compactors. In: Proceedings of the 8th Australia New Zealand conference on Geomechanics: consolidating knowledge. Hobart, Australia: Australian Geomechanics Society; 1999. p. 775–81.
Poli R, Langdon WB, Mcphee NF, Koza JR. Genetic programming: an introductory tutorial and a survey of techniques and applications. Technical Report CES-475. University of Essex; 2007.
Ranasinghe RATM, Jaksa MB, Kuo YL, Pooya Nejad F. 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 2017a;9(2):340–9.
Ranasinghe RATM, Jaksa MB, Pooya Nejad F, Kuo YL. Predicting the effectiveness of rolling dynamic compaction using genetic programming. Proceedings of the Institution of Civil Engineers Ground Improvement 2017b;170(4):193–207.
Ranasinghe RATM, Jaksa MB, Pooya Nejad F, Kuo YL. Prediction of the effectiveness of rolling dynamic compaction using artificial neural networks and cone penetration test data. Chinese Journal of Rock Mechanics and Engineering 2019;38(1):153–70.
Rashed A, Bazaz JB, Alavi AH. Nonlinear modeling of soil deformation modulus through LGP-based interpretation of pressuremeter test results. Engineering Applications of Artificial Intelligence 2012;25(7):1437–49.
Rezania M, Javadi AA. A new genetic programming model for predicting settlement of shallow foundations. Canadian Geotechnical Journal 2007;44(12):1462–73.
Roy PP, Roy K. On some aspects of variable selection for partial least squares regression models. QSAR Combinatorial Science 2008;27(3):302–13.
Scott BT, Jaksa MB. Mining applications and case studies of rolling dynamic compaction. In: Proceedings of the Australia-New Zealand (ANZ) conference on Geomechanics. Melbourne, Australia; 2012. p. 961–6.
Selle B, Muttil N. Testing the structure of a hydrological model using genetic programming. Journal of Hydrology 2011;397(1–2):1–9.
Smith GN. Probability and statistics in civil engineering. London: Collins; 1986.
Torres RDS, Falcão AX, Gonçalves MA, Papa JP, Zhang B, Fan W, Fox EA. A genetic programming framework for content-based image retrieval. Pattern Recognition 2009;42(2):283–92.

R.A.T.M. Ranasinghe is currently pursuing doctoral degree in civil engineering in the School of Civil, Environmental and Mining Engineering at the University of Adelaide, Australia. She completed the bachelor of the science of engineering (Honors) degree in civil engineering with a first class in 2011 from the University of Moratuwa, Sri Lanka. She is currently conducting a research in developing a robust predictive tool to forecast the performance of rolling dynamic compaction using artificial intelligence (AI) techniques. Her research interests include ground improvement, AI techniques, and computer applications in geotechnical engineering. She is a member of the Institution of Engineers Sri Lanka (IESL) and Australian Geomechanics Society (AGS).