Soils and Foundations 2014;54(2):233–242

The Japanese Geotechnical Society

Soils and Foundations

www.sciencedirect.com
journal homepage: www.elsevier.com/locate/sandf

Predicting pile dynamic capacity via application of an evolutionary algorithm

I. Alkroosh*, H. Nikraz

Department of Civil Engineering, Curtin University, Perth, Western Australia, Australia

Received 21 May 2013; received in revised form 11 October 2013; accepted 14 November 2013
Available online 3 April 2014

Abstract

This study presents the development of a new model obtained from the correlation of dynamic input and SPT data with pile capacity. An evolutionary algorithm, gene expression programming (GEP), was used for modelling the correlation. The data used for model development comprised 24 cases obtained from existing literature. The modelling was carried out by dividing the data into two sets: a training set for model calibration and a validation set for verifying the generalization capability of the model. The performance of the model was evaluated by comparing its predictions of pile capacity with experimental data and with predictions of pile capacity by two commonly used traditional methods and the artificial neural networks (ANNs) model. It was found that the model performs well with a coefficient of determination, mean, standard deviation and probability density at 50% equivalent to 0.94, 1.08, 0.14, and 1.05, respectively, for the training set, and 0.96, 0.95, 0.13, and 0.93, respectively, for the validation set. The low values of the calculated mean squared error and mean absolute error indicated that the model is accurate in predicting pile capacity. The results of comparison also showed that the model predicted pile capacity more accurately than traditional methods including the ANNs model.

Keywords: Pile; Dynamic capacity; Gene expression programming; Training; Validation

1. Introduction

Although it is common in design practice to predict pile capacity by static analysis, a pile driving formula or a dynamic formula is perhaps the most frequently used method for evaluating the capacity of driven piles, as described by Poulos and Davis (1980). Evaluation of pile dynamic capacity is considered useful as the main purpose of driving formulae is using the driving record of the pile to establish the safe working load for a pile, or to determine the driving requirements for a required working load (Ng et al., 2004).

Numerous researchers have proposed different procedures for evaluating pile capacity based on dynamic input. However, there are two approaches most commonly used dynamic formulae and wave equations. Despite the frequent and widespread use of these methods, their reliability is still questionable. Dynamic formulae have been investigated by researchers (e.g. Flaate, 1964; Housel, 1966), who concluded that pile capacities determined from dynamic formulae correlate poorly with static load test results and have a wider scatter when statistically compared. The *Manual for Design and Construction of Driven Pile Foundations* by Hannigan et al. (1996)

*Corresponding author.
Peer review under responsibility of The Japanese Geotechnical Society.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sandf.2014.02.013
0038-0806/

clearly characterizes unsatisfactory prediction in pile capacity by dynamic formulae. The wave equation analysis is also criticized by a number of researchers such as Svinkin and Woods (1998), who explained that this method does not take into account changes in soil properties after pile installation; thus the method apparently cannot predict reliable pile capacity for various elapsed times after driving has ceased.

The limited success of dynamic methods in achieving accurate evaluation of pile capacity can be attributed to the assumptions on which these methods are based along with an oversimplification of pile behaviour. Pile driving formulae assumes that the work done in forcing down the pile (*i.e.*, the product of the weight of the ram and the stroke) is equal to the product of the ultimate soil resistance. The main shortcoming of this assumption is that there is a difficulty in estimating the actual energy transmitted by the ram to the pile through the cap block, pile cap and cushion. Thus the energy losses in a real pile driving situation cannot be accounted for accurately (Coduto, 1994). The wave equation assumes that static resistance is a function of dynamic force and the velocity generated by hammer blows and damping coefficient. This assumption presents two difficulties: (1) the total resistance is time dependent and different variations in the method produce different results; (2) the dimensionless damping coefficient has a questionable correlation to soil type and needs to be calibrated for the specific pile, soil and site condition (Ng et al., 2004).

The complexity of pile behaviour and the presence of many involving factors have made it difficult to develop an accurate model based on traditional modelling procedures. Artificial intelligence techniques may be a better alternative, due to the capability of these techniques being able to deal with complex and highly nonlinear functions, and employing the considerable capacity of computers to perform enormously iterated work. A number of researchers (*e.g.* Chan et al.,1995; Teh et al., 1997; Abu-Kiefa, 1998; Das and Basudhar, 2006; Ardalan et al., 2009; Shahin, 2010; Ornek et al., 2012; Tarawneh, 2013) have successfully applied artificial neural networks (ANNs) for the modelling of pile behaviour. The modelling advantage of ANNs is their ability to capture the nonlinear and complex relationships between the targeted output and the factors affecting it, without having to assume *a priori* formulae describing this relationship. However, the main shortcoming of ANNs is the complexity of their network structure, as they represent the knowledge in terms of weight matrices together with biases that are not accessible to the user (Rezania and Javadi, 2007). In this regard, the genetic programming (GP) may present a better alternative. The main advantage of the GP over ANNs is the ability to provide the relationship between a set of inputs and the corresponding outputs in a simple mathematical form considered accessible to the user (Rezania and Javadi, 2007). Recently, the GP has been applied with success in solving engineering problems (*e.g.* Javadi et al., 2006; Rezania and Javadi, 2007; Alavi et al., 2011). In this paper, pile dynamic capacity has been correlated with SPT data and dynamic input using a developed version of genetic programming; that is, gene expression programming (GEP). Recently, GEP has been applied successfully in solving engineering problems (*e.g.* Cevic and Cabalar, 2009; Alkroosh and Nikraz, 2011a, 2011b; Gandomi, 2011; Gandomi and Alavi, 2012). The objectives of this paper are as follows:

- To investigate the feasibility of using GEP to correlate dynamic input data and SPT results with pile capacity;
- To evaluate the performance of the developed GEP model in training and validation sets by comparing its prediction of pile capacity with experimental data and with predictions of pile capacity by traditional methods along with the ANN model;
- To conduct a parametric study to evaluate the influence of the input variables on the performance of the model.

2. Overview of gene expression programming

GEP is an instance of an evolutionary algorithm from the field of evolutionary computation, invented by Ferreira (2001) as a global optimization algorithm. It has similarities to other evolutionary algorithms such as genetic algorithms (GAs), as well as other evolutionary automatic programming techniques such as genetic programming (GP). Similar to GAs, GEP uses the evolution of linear computer programs (individuals or chromosomes) of fixed length and likewise the GP the evolved programs are expressed in nonlinear forms of expression trees (ETs) of different sizes and shapes. However, GEP implements a different evolutionary computational method. The GEP distinguishes itself from GAs in that the evolved solutions are expressed in the form of parse trees of different sizes and structures and unlike GP, genetic variations are performed on chromosomes before they are translated into ETs.

The GEP chromosomes can be composed of single or multiple genes; each gene is encoded in a smaller sub-program. Every gene has a constant length and includes a head that contains functions (*e.g.* +,–) and terminals (*e.g.* d1, d2, which are the symbolic representation of the input variables), and a tail composed of terminals only. The genetic code represents a one-to-one relationship between the symbols of the chromosome, the functions or terminals. The process of information decoding from chromosomes to ETs is called translation; this is based on sets of rules that determine the spatial organization of the functions and terminals in the ETs and the type of interaction (link) between the sub-ETs (Ferreira, 2002). The principal terms of the GEP are described in the following subsections.

2.1. Initial population

In GEP, the search for a solution begins when a number of computer programs (individuals or chromosomes), referred to as the initial population, are randomly created from the set of functions and terminals defined by the user. Each program is expressed, evaluated and assigned ‘fitness’ according to how well it performs with regard to achieving the desired objective.