Genetic programming for modelling of soil – structure interactions

Yeruva Ramana Reddy

Sr. Geotechnical Engineer & Department of Civil Engineering

Indian Institute of Technology, India

yramanareddyiit@gmail.com

Abstract— *The primary objective of this study is to discuss the modeling features and interpretation of genetic programming as well as its applicability in geotechnical engineering. The soil-structure systems are endless in terms of the solid medium. To deal with this geometric infinite, many solutions have been devised to reduce the system dimensions. When nonlinearities are included, analyzing an accurate soil-structure system is prohibitively time-consuming [1]. In addition, the boundaries stated are usually only used for basic geometries. For complicated engineering challenges, adopting innovative data-based solutions has attracted numerous significant research efforts in recent years. Using optimized neural networks to solve the soil structure issue is explored in this research as an essential branch of data-based methods. It has been shown that artificial intelligence (AI) can solve numerous geotechnical engineering issues that are above the computing capabilities of classical mathematics and conventional procedures [1] [2]. For geotechnical engineers, genetic programming (GP) is a fascinating AI method used to solve many problems. Recently, GP, drawing inspiration from human evolution, has shown success in modeling various geotechnical engineering issues and outperformed established techniques in terms of accuracy.*

Keywords: Genetic Programming, soil-structure interaction, Artificial intelligence, Artificial Neural Network (ANN)

I. INTRODUCTION

Imprecise physical processes linked with creating geotechnical materials (such as soil and rock) lead to uncertainty in their behavior. It is difficult to predict how these materials behave because of this ambiguity. Structural components that interact with them are subject to the same rules [2] Regarding the structural components employed to transmit superstructure loads, pile foundations are prone to material uncertainty and modeling complexity. The capacity of artificial intelligence (AI) to forecast the complicated behavior of materials is superior to that of previous methods [2]. Because of this, artificial intelligence (AI) has become a popular and handy tool in geotechnical engineering. There are several examples of artificial intelligence (AI) approaches, including Genetic Programming (GP) and Artificial Neural networks (ANN). GP and ANN applications in estimating the bearing capacity of pile foundations are reviewed in this research.

This work's primary goal is to offer a short explanation of GP approaches and a literature assessment of their use in soil-structure interaction modeling. Genetic programming is a kind of AI-derived from genetic algorithms and inspired by biological evolution. Evolving algorithms are used to find a computer program that can execute a particular computing job [3,4]. This method is an approach to problem-solving that is not specific to any area. It involves the evolution of computer programs that are made of operations and interfaces to solve, or nearly solve, issues by creating a structured description of the information. The structural depiction mimics live creatures' development and natural genetic operations[4]. The primary benefit of GPs over ANNs is their capacity to present the connection between a set of input data and their related outputs in a straightforward mathematical form that anybody can understand.

The first thing to simulate GP is to create a random number of computational methods (also called chromosomes). A user-defined collection of functions and terminals is used to populate the system at launch time. This tree-like framework consists of a root of the tree and branch offices of operational nodes and terminals, which are the essential building elements of the general-purpose GP models [5.]. User-defined contractors, parameters, and parameters are all representations of operations and terminals in GP that are used to do any arithmetic or trigonometric calculation. To begin analysis in GP, a collection of functions representing the issue or data must first be determined. Each person in the population is given a score based on their ability to adapt to their surroundings. The objective function determines the fitness requirements, which measures how well an individual does compare to the rest of the population [7].

Reproduction, hybridization, and mutation are applied to a fraction of the computer simulations to generate a new population. Reproduction refers to transferring a computer simulation from an original population into a new demographic without making any changes; crossover refers to the process of genetically recombining randomly selected sections of two computational methods. Mutation refers to replacing a randomly picked functional or port component with others from the same functionality or terminus set. In the end, the current population will be supplanted by the new one [8]. An acceptable mistake or a maximum number of iterations might be used as a termination condition for this process of evolution. Finally, the optimal computerized framework is designed by GP to use the selected objective functions.

II. RESEARCH PROBLEM

The main problem that will be solved by this research is to explore how a genetic programming model is applied in soil-structure interactions. The failure loads, deformations, and flows are some of the issues that this geotechnical engineering model will attempt to address. Experimentation, theoretical modeling, or relying on prior knowledge are the most common methods for resolving these issues. Genetic programming (GP), a cutting-edge AI technology, and its applicability in geotechnical engineering are discussed in this work. It has been shown that artificial intelligence (AI) can solve numerous geotechnical engineering issues that are above the computing capabilities of classical mathematics and conventional procedures [9].[10,11] Recently, GP, drawing inspiration from human evolution, has shown success in modeling various geotechnical engineering issues and outperformed established techniques in terms of accuracy. Geotechnical modeling and formulation in GP are detailed and explained in depth in this chapter, which also presents and discusses an overview of the most successful GP applications.

III. LITERATURE REVIEW

A. *Genetic Programming*

GP is a more recent approach for analyzing stacks of intelligence. As a result, there are fewer parameters in Genetic algorithms (GAs) for each category compared to ANN. The Darwinian principle of "survival of the fittest" informs GP, an evolutionary algorithm approach. It is possible to develop an optimally-structured solution using machine learning techniques without any prior assumptions about how the answer is constructed. GP is an example of this sort of machine learning technique." It is possible to characterize GP programs by utilizing syntax trees, the nodes of which primarily consist of functional and terminal components [10].

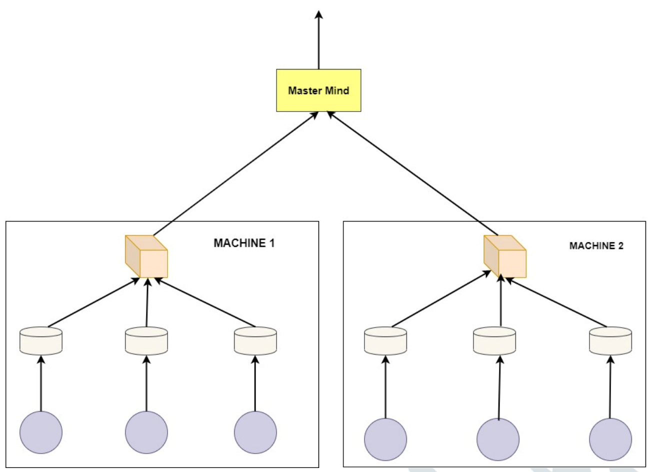

Fig i: Distributed genetic algorithm with parallel components

B. *Interactions between the structure and the soil*

The coil support layer's strength decreases as water content rises. Mineralogy, fabric, and pore water all influence soil geotechnical qualities, but the interdependence of these variables make it impossible to isolate their effects using typical statistical approaches. Capillary action and changes in the water's elevation and water penetration through the coating are the primary causes of the drop in the California Bearing Ratio after immersion (CBRimm) index. To accurately assess the long-term health of the soil, it is vital to account for the impact of water [10]. As a result, immersing the soil sample in water for four days is a critical stage in the method.

C. *Genetic programming-based forecasting of soil-structure interactions*

A database from soil measurements serves as the foundation for genetic programming modeling: heterogeneous buildings and marly formations around them. Soil samples might come from areas with various geotechnical issues, most of which were caused by the soil's behavior. Fundamental regression analysis is performed before genetic programming is used to create a new model. This research sets out to determine if the input and output variables in the constructed model have a linear or non-linear association. A model is considered to have high levels of correlation when its coefficient of determination (R2) is very near to 1[10]. Correlation coefficients of 0.66–0.77 are found between the input variables and the CBRimm index, indicating a relationship between the model parameters and this index[10].

D. *Using a Genetic Algorithm*

Generic search algorithms based on natural evolution are among the most widely used. GA has gained much attention in engineering design optimization in the last several decades. As far back as the 1960s, a team of biologists attempted to replicate the process that occurs naturally in nature in a computer program [11]. Genetic algorithms (GAs) are any population-based method that employs selection, crossovers, and mutations across chromosomes to discover the best solution. In practice, a person's chromosome or genotype is referred to as their membership in the population, depending on whether it is a binary or real-valued function string. Several Genetic algorithms (GAs) have been used in optimization research after Barricelli [11].

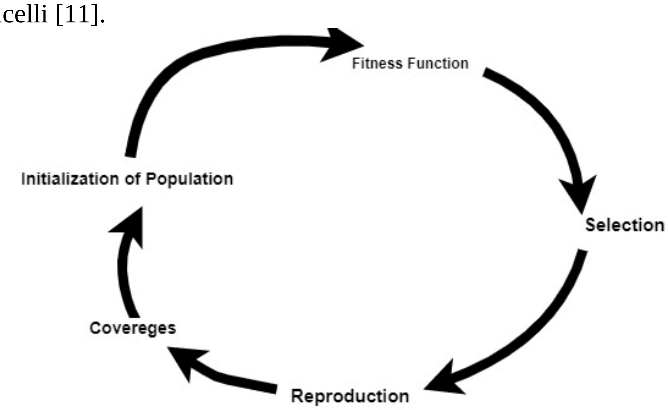

Fig ii: Genetic Algorithm cycle

E. *The implementation of multi-gene genetic programming*

Modified Genetic Programming (MGGP) is a genetic programming and parameter estimation technique hybrid. As evidence of MGGP's capabilities, it is often used to construct challenging geotechnical engineering issues. Problems studied include piles' evaluation of undrained lateral load capacity, drilled shafts' settling around tunnels, and soil liquefaction's alpha factor for undrained side resistance[12]. Validity is assessed for a subset of test outcomes that are not drawn from training data. MGGP's precision, efficiency, and enormous potential are shown numerically. MGGP, unlike artificial neural networks and other soft cognitive computing, gives a constitutive predictive model. For pre-design, MGG-based solutions are beneficial.

An introduction of genetic programming (GP) and multi-gene genetic programming (MGGP) is presented using the Genetic Algorithm (GA). Darwin's "Survival of the Fittest" hypothesis is the foundation of genetic algorithms. Genetic algorithms are built on this principle. Scientists use genetic algorithms to explore big issue spaces and identify the best answers utilizing the natural development of the research process. Even though GA and GP have certain similarities, some distinctions in the answers that the two algorithms produce are worth discussing. The Genetic Algorithm (GA) generates solutions in binary form, while the Genetic Algorithm (GP) uses a tree-based structure with variable depth [12]. The GP structures and GA parameters optimization is another point of distinction. Regarding the usefulness of the two algorithms, one may claim that the genetic algorithms can solve difficulties if the person knows what the answer previously acquired by genetic programming looks like. This method may help formulate a solution.

Using a random mix of items from the function and terminal sets depicted in the diagrams shows that the algorithm generates the population in an unplanned manner. The half-and-half method is used to produce models of homogeneous shape and size. Boolean operators and non-linear functions (such as Sin and Cosine) are included in the function set, as can arithmetic operators (+ and –) and even Boolean functions (such as Log and Sqrt). Process input variables and random constants [12] make up the terminal set.

In maintaining a certain level of genetic diversity within the population, individuals are first screened using Genetic Programming (GP) and then reproduced using a mutation operator, followed by another crossing to produce a new generation of individuals better suited to the problem at hand. To produce the growing population of the following generation, individuals improved by genetic algorithms go through an elimination stage in which the optimal solution based on their performance parameters is retained. The fitness function compares the expected and measured values to determine performance. It has been decided to use the Mean Absolute Error (MAE) as an indication to verify the model.

Despite its simplicity, the MGGP model is a stochastic exploration for addressing a challenging optimization issue. Using poorly correlated parameters does not indicate that evolutionary algorithms adequately address all geotechnical issues. When there aren't enough resources available, learning and performance suffer [16].

IV. SIGNIFICANCE

The U.S. construction has benefited significantly from using artificial neural networks (ANNs) to simulate loads settling during building projects. The selection of relevant factors is critical to constructing accurate prediction models based on soil qualities. Including all variables in a model enhances the model's complexity but does not improve its predictive power [15]. Furthermore, comparable results are obtained by varying a variety of factors. As a result, researchers are always searching for low-complexity predictive models with good predictive power. For example, the feature selection (F.S.) technique reduces the number of features/variables while simultaneously maximizing the prediction accuracy of a model[16]. The most widely acknowledged method for assessing the stability of a rock-socketed pile is to conduct a pile load test. However, the test takes a long time and is quite expensive. In the case of strata that are all of the same thickness, the pile analysis is nave. There are several empirical, semi-experimental, theoretical, and experimental approaches for simulating the unique response of a site to site-specific load settling in circumstances when strata are spatially variable[17]. The Artificial Neural Network (ANN) method is one alternative to typical computing approaches.

V. FUTURE IN THE U.S.

The use of A.I. in construction can assist stakeholders in realizing value across the lifecycle of a project. This process includes design, bidding, and budgeting; design and planning; operational and asset management; and business strategy reinvention. The field of construction management research has increasingly used artificial neural networks (ANNs) in recent decades [17] due to their superior performance in complicated situations. The use of artificial neural networks (ANNs) in U.S. construction projects is only going to grow. Using Building Information Modeling (BIM), architects, engineers, and construction managers may more effectively plan and design buildings and construct and maintain them. The simulations must consider the plans for the architectural, engineering, structural, electrical, and plumbing (MEP) departments, as well as the order of their separate operations in terms of planning and designing the building of a project [18]. TBM building has grown popular in metropolitan areas with a significant population because of their little impact on the environment and rapid construction rates [19]. A critical criterion for designing and implementing earth pressure balance shields (EPBs) is the maximum surface settlement (MSS), calculated before tunnelling. For U.S. building technology, artificial intelligence (A.I.) technologies will provide an alternate strategy for dealing with very complicated issues that cannot be modelled in mathematics [20]. The accurate estimates of ground settlement mean that engineers and academics have gone to considerable efforts to predict the effects of tunnelling, using both empirical and analytical methods.

VI. CONCLUSION

This paper discussed how piles driven in cohesive soils may be modelled using an artificial neural network technique suggested in this research. Neural networks are used to predict the movement of the ground caused by tunnelling projects, to evaluate how effective two or more trenches are at mitigating ground vibration and evaluate how effective geofoam-filled trenches are at reducing the movement of the ground caused by tunnelling, and to assess the effects of railway traffic on free-field vibrations. Many variables contribute to the difficulty of analyzing pile load-settlement behaviour. This research proposes a novel method for modelling the load-settlement behaviour of pile foundations buried in sand and mixed soils using artificial neural networks (ANNs) (subjected to axial loads). ANN models are created for several types of piling, such as those buried in sandy or mixed soil and behind a layer of cohesive soil. It seems that ANN models can adequately anticipate the complicated nonlinear behaviour of pile load-settlement with a high degree of accuracy, as shown by the findings. In the training and testing sets used to create ANN models, the coefficients of correlation have high values near unity, according to the statistical analysis.

REFERENCES

[1] L. V. Fausett, Fundamentals of neural networks: Architectures, algorithms and applications. Englewood Cliffs, NJ: Prentice-Hall International, 1994.

[2] Öztaş, M. Pala, E. Özbay, E. Kanca, N. Çagˇlar, and M. A. Bhatti, "Predicting the compressive strength and slump of high strength concrete using neural network," *Construction and Building Materials*, vol. 20, no. 9, pp. 769–775, Nov. 2006. Available: https://doi.org/10.1016/j.conbuildmat.2005.01.054

[3] R. E. Keller, W. Banzhaf, P. Nordin, and F. D. Francone, Genetic Programming : An Introduction : On the Automatic Evolution of Computer Programs and Its Applications (The Morgan Kaufmann Series in Artificial Intelligence). Morgan Kaufmann Publishers, 1997.

[4] G. L. M. Leung, A. W. G. Wong, and Y. H. Wang, "Prediction of resilient modulus of compacted saprolitic soils by CBR approach for road pavement subgrade: a re-examination," *International Journal of Pavement Engineering*, vol. 14, no. 4, pp. 403–417, Apr. 2013. Available: https://doi.org/10.1080/10298436.2012.727993

[5] H. Gandomi, A. H. Alavi, M. R. Mirzahosseini, and F. M. Nejad, "Nonlinear Genetic-Based Models for Prediction of Flow Number of Asphalt Mixtures," *Journal of Materials in Civil Engineering*, vol. 23, no. 3, pp. 248–263, Mar. 2011. Available: https://doi.org/10.1061/(asce)mt.1943-5533.0000154

[6] S. Aljahdali and A. Sheta, "Evolving Software Effort Estimation Models Using Multigene Symbolic Regression Genetic Programming," *International Journal of Advanced Research in Artificial Intelligence*, vol. 2, no. 12, 2013. Available: https://doi.org/10.14569/ijarai.2013.021207

[7] Yildirim and O. Gunaydin, "Estimation of California bearing ratio by using soft computing systems," *Expert Systems with Applications*, vol. 38, no. 5, pp. 6381–6391, May. 2011. Available: https://doi.org/10.1016/j.eswa.2010.12.054

[8] H. Cheng and S. Yang, "Multi-population Genetic Algorithms with Immigrants Scheme for Dynamic Shortest Path Routing Problems in Mobile Ad Hoc Networks," in *Applications of Evolutionary Computation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 562–571. Available: https://doi.org/10.1007/978-3-642-12239-2_58.

[9] R. M. Ebeling, "Methods of evaluating the stability and safety of gravity earth retaining structures founded on rock," Dissertation, Virginia Polytechnic Institute and State University, 1989. Accessed: http://hdl.handle.net/10919/54351

[10] R. M. Ebeling, R. L. Mosher, K. Abraham, and J. F. Peters, "Soil-Structure Interaction Study of Red River Lock and Dam No. 1 Subjected to Sediment Loading," *Defense Technical Information Center, Fort Belvoir*, VA, Sep. 1993. https://doi.org/10.21236/ada270213

[11] M. Pala, N. Caglar, M. Elmas, A. Cevik, and M. Saribiyik, "Dynamic soil–structure interaction analysis of buildings by neural networks," *Construction and Building Materials*, vol. 22, no. 3, pp. 330–342, Mar. 2008. Available: https://doi.org/10.1016/j.conbuildmat.2006.08.015

[12] X. Lu, B. Chen, P. Li, and Y. Chen, "Numerical Analysis of Tall Buildings Considering Dynamic Soil-Structure Interaction," *Journal of Asian Architecture and Building Engineering*, vol. 2, no. 1, pp. 1–8, May. 2003. Available: https://doi.org/10.3130/jaabe.2.1

[13] J. Penzien, C.-H. Chen, Y.-J. Lee, H.-N. Hsiau, and I. Katayama, "Soil-Structure Interaction Analysis Using Multiple Input Motions," in *Developments in Dynamic Soil-Structure Interaction*. Dordrecht: Springer Netherlands, 1993, pp. 167–178. Available: https://doi.org/10.1007/978-94-011-1755-5_8

[14] J. J. Grefenstette, Proceedings of the First International Conference on Genetic Algorithms and their Applications. Lawrence Erlbaum, 1985.

[15] S. Pal, G. Wije Wathugala, and S. Kundu, "Calibration of a constitutive model using genetic algorithms," *Computers and Geotechnics*, vol. 19, no. 4, pp. 325–348, Jan. 1996. Available: https://doi.org/10.1016/s0266-352x(96)00006-7

[16] C. Pereira, J. Maranha, and A. Brito, "Advanced constitutive model calibration using genetic algorithms," in *Numerical Methods in Geotechnical Engineering*. CRC Press, 2014, pp. 485–490.. Available: https://doi.org/10.1201/b17017-87

[17] C.Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, pp. 79–82, 2005. Available: https://doi.org/10.3354/cr030079.

[18] H. Gandomi and A. H. Alavi, "A new multi-gene genetic programming approach to non-linear system modeling. Part II: geotechnical and earthquake engineering problems," *Neural Computing and Applications*, vol. 21, no. 1, pp. 189–201, Aug. 2011. Available: https://doi.org/10.1007/s00521-011-0735-y.

[19] D. Liu, H. Zhang, and S. Hu, "Neural networks: Algorithms and applications," *Neurocomputing*, vol. 71, no. 4-6, pp. 471–473, Jan. 2008. Available: https://doi.org/10.1016/j.neucom.2007.09.001.

[20] H. Ardalan, A. Eslami, and N. Nariman-Zadeh, "Piles shaft capacity from CPT and CPTu data by polynomial neural networks and genetic algorithms," *Computers and Geotechnics*, vol. 36, no. 4, pp. 616–625, May. 2009. Available: https://doi.org/10.1016/j.compgeo.2008.09.003