Proceedings of Indian Geotechnical Conference
December 15-17, 2011, Kochi (Paper No. N-082.)

EVALUATION OF LIQUEFACTION POTENTIAL OF SOIL USING GENETIC PROGRAMMING

Sarat Kumar Das, Associate Professor, Civil Engineering Department, NIT, Rourkela, sarat@nitrkl.ac.in
Pradyut Kumar Muduli, Research Scholar, Civil Engineering Department, NIT, Rourkela, pradyut.muduli@gmail.com

ABSTRACT: Liquefaction of soil is one of the major causes for the significant damages to the buildings, lifeline systems and harbor facilities caused by the earthquakes. At present artificial intelligence techniques such as artificial neural network (ANN) and support vector machine (SVM) based models are found to be more efficient compared to statistical methods. The present study discusses about the evaluation of liquefaction potential of soil based on cone penetration test (CPT) data obtained after 1999 Chi-Chi, Taiwan, earthquake using evolutionary artificial intelligence technique known as genetic programming (GP). A comparative study is made among the existing three CPT based statistical methods and the developed GP model for prediction of liquefied and non-liquefied cases in terms of percentage success rate with respect to the field observations.

INTRODUCTION

Liquefaction of soil is one of the most disastrous seismic hazards. In the last century seismic hazard accounts around 30% of total casualties and 60% of the total property loss due to different natural hazards [1]. Soil liquefaction phenomena have been noticed in many historical earthquakes after first large scale observations of damage caused by liquefaction in the 1964 Niigata, Japan and 1964 Alaska, USA, earthquakes. Since 1964 a lot of work has been done to explain and evaluate the liquefaction hazard [2, 3, 4].

Simplified methods based on standard penetration test (SPT), cone penetration test (CPT) and shear wave velocity measurement test are most commonly used for the assessment of liquefaction potential of soils, due to difficulty in obtaining high quality undisturbed samples and cost involved therein. Simplified methods pioneered by Seed and Idris [2] mostly depend on a boundary curve which presents a limit sate and separates liquefaction cases from the non-liquefaction cases basing on field observations of soil in earthquakes at the sites where in situ data are available. Though SPT is most widely used soil exploration method, CPT is becoming more acceptable due to consistent, repeatable and identification of continuous soil profile. Hence, now a days CPT is being widely used for liquefaction susceptibility analysis of soil using various statistical and regression analysis techniques [5, 6]. Artificial intelligence techniques such as artificial neural network (ANN) [7, 8] and support vector machine (SVM) [9, 10] are found to be more efficient compared to statistical methods. However, the ANN has poor generalization, attributed to attainment of local minima during training and needs iterative learning steps to obtain better learning performances. The SVM has better generalization compared to ANN, but the parameters ‘C’ and insensitive loss function (ε) needs to be fine tuned by the user. Moreover these techniques will not produce an explicit relationship between the variables and thus the model obtained provides very little insight into the basic mechanism of the problem. The evolutionary artificial intelligence techniques based on the Darwinian theory of natural selection provide strong alternatives to the mentioned techniques. Genetic programming (GP) is one of such technique which can automatically select the system inputs to develop a model which fits well the input output relationship of the system. In the present study an attempt has been made to predict the liquefaction potential of soil based on CPT data obtained after Chi-Chi earthquake, Taiwan, 1999 [11] using GP. A comparative evaluation of the present study is made with three existing CPT based statistical methods for prediction of liquefied and non liquefied cases in terms of percentage success rate with respect to the field manifestations.

METHODOLOGY

In this paper liquefaction index based on field CPT data is predicted using genetic programming. The detailed methodology is presented as follows.

CPT-Based Method for Prediction of Liquefaction Index

The common deterministic methods used for the liquefaction potential evaluation of a site based on CPT data base follow the general stress based approach pioneered by Seed and Idriss [2]. These methods are based on determination of factor of safety (F_s) against the liquefaction occurrence and is defined as $F_s=CRR/CSR_{7.5}$, where, $CSR_{7.5}$ = cyclic stress ratio adjusted to the benchmark earthquake (moment magnitude, M_w=7.5) and *CRR* is cyclic resistance ratio. In deterministic approach liquefaction and non-liquefaction cases are predicted on the basis of corresponding $F_s \leq 1$ and $F_s > 1$ respectively. Due to model and parameter uncertainties there is always some probability that liquefaction can occur even if $F_s > 1$. In order to understand the meaning of calculated F_s within the framework of probability it is expressed in terms of probability of liquefaction (P_L) through Bayesian mapping function [4].

Juang *et al*. [8] while using ANN denoted liquefaction performance observation as liquefaction index (*LI*) and described in binary number. Thus *LI* = 1 if liquefaction is observed and *LI* = 0 if liquefaction is not observed. In the present study a model is developed using GP for the prediction of liquefaction potential of soil in terms of *LI* using CPT based liquefaction field performance dataset. The present method is compared with statistical based Robertson method [3], Olsen method [12] and Juang method [8] in terms of rate of successful prediction for liquefied and non liquefied cases.

Genetic Programming

Genetic Programming (GP) is a pattern recognition tool where the model is developed on the basis of adaptive learning over a number of cases of provided data, developed by Koza [13]. It mimics biological evolution of living organisms and makes use of principle of genetic algorithm (GA). In traditional regression analysis the user has to specify the structure of the model whereas in GP both structure and the parameters of the mathematical model are evolved automatically. It provides a descriptive solution in the form of tree structure or in the form of compact equation based on the provided data set. GP has been successfully implemented in various engineering problems but its use in solving geotechnical engineering problems is limited [14, 15]. A brief description about GP is presented as follows.

GP models /computer programs are composed of nodes and resemble to a tree structure and thus well known as GP tree. Nodes are the elements either from a functional set or terminal set. A functional set may include arithmetic operators (+, ×, ÷, or -), mathematical functions (sin(.), cos(.), or ln(.)), Boolean operators (AND, or OR), logical expressions (IF, or THEN) or any other user defined functions. The terminal set includes variables (like x_1, x_2, x_3, etc) or constants (like 2, 5, 9 etc) or both. The functions and terminals are randomly chosen to form a computer model/ program with a root node and the branches extending from each function nodes to end in terminal nodes as shown in Figure 1.

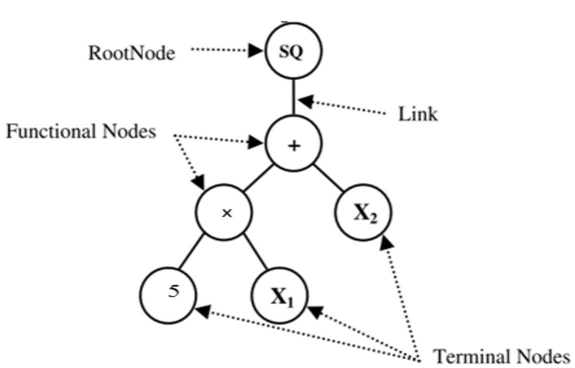

Fig. 1 Typical GP tree representing function $(5X_1+X_2)^2$

In the first stage of implementation of GP, a population of computer models is randomly created. Each individual in the population is measured by fitness criteria. The fitness criteria are calculated by the objective function and it determines quality of the individual in competing with rest of the population. At each generation a new population is generated by selecting individuals according to their fitness and implementing various evolutionary mechanisms like reproduction, crossover and mutation to the functions and terminals. The new population then replaces the existing population. This process is iterated until the termination criterion, which can be either a threshold fitness value or maximum number of generations, is satisfied. The GP model is developed using Matlab (Math Work Inc 2005).

RESULTS AND DISCUSSION

Development of the GP model for evaluation of liquefaction potential of soil

A devastating earthquake of moment magnitude (M_w), 7.6 struck Taiwan on 21st September 1999. The epicenter was at 23.87^0 N 120.75^0E which is near Chi-Chi of Taiwan. In this earthquake liquefaction of soil was the major cause of heavy damages [4]. In the present study post earthquake field observations and the CPT data collected, from Wufeng, Nantou, Yuanlin and Lunwei areas of Taiwan as per Ku *et al.* [11] are used. The database consists of total 134 cases, 46 out of them are liquefied cases and other 88 are non liquefied cases. Out of the above data 94 cases are randomly selected for training and remaining 40 data are used for testing the developed model. The data was normalized in the range 0 to 1 to avoid the dimensional effect of input parameters.

The model describes the relationship between the output, *LI*, and the inputs include soil and seismic parameters. The general form of proposed GP model can be presented as:

$$LI = \sum_{i=1}^{n} F\left[X, f(X), b_i\right] + b_0 \quad (1)$$

where F= the function created by the GP process referred herein as liquefaction index function, X = vector of input variables = { q_c, f_s, Pw, σ_v, σ_v', a_{max}/g} , q_c = measured cone tip resistance, f_s = sleeve resistance, Pw = the pore water pressure, σ_v = vertical total stress of soil at the depth studied, σ'_v = vertical effective stress of soil at the depth studied, a_{max}= peak horizontal ground surface acceleration, g = acceleration due to gravity, b_i is constant, f is a function defined by the user, n= the number of terms of target expression and b_0 is the bias.

In the present study a multigene GP is adopted where each individual consists of more than one gene, each of which is a traditional GP tree. Here function set used include: +, ×, ÷, -, sin(.), cos(.) and exp(.). In the GP procedure a number of potential models are evolved at random and each model is trained and tested using the training and testing cases respectively. The fitness of each model is determined by minimizing the mean square error (MSE) between the predicted and actual value of the output variable (*LI)* as the objective function,

$$f(\delta) = \frac{\sum_{i=1}^{n} (\delta)^2}{n} \tag{2}$$

where n = number of cases in the fitness group and δ is the difference between the predicted and actual value of *LI* as per the field data set. If the errors calculated by using Eq. 2 for all the models in the existing population do not satisfy the termination criteria, the evolution of new generation of population continues till the best model is developed.

The proposed model

The best model was obtained with population size of 2000 individuals at 150 generations with reproduction probability = 0.05, crossover probability = 0.85, mutation probability = 0.1, maximum number of genes = 2, maximum depth of GP tree = 4, and tournament selection. The developed models were analyzed with respect to engineering understanding of liquefaction potential evaluation of soil and after careful consideration of various alternatives the following expression was found to be most suitable for the prediction of liquefaction potential of soil in terms of *LI*.

$$\begin{aligned} LI &= 0.204\cos\left(7.159\,q_c\right) - 1.044q_c + 0.204\,\sigma_v'\left(a_{\max}/g\right) \\ &- \frac{0.562\,q_c\exp\left(-a_{\max}/g\right)}{4.818\,q_c - \exp\left(a_{\max}/g\right)} + 0.8376 \end{aligned} \tag{3}$$

Eq. 3 describes *LI* as a function of three main contributing variables q_c, σ'_v, and (a_{max}/g) out of various input variables as discussed earlier.

It is evident from the result presented in Table 1 that the proposed GP model was able to learn the complex relationship between the liquefaction index and its main contributing factors with a very high accuracy. It can be noted that the performances of GP for training and testing data are comparable and for *LI* within ±15% error limit the successful prediction values are 100% for training and 95% for testing. Similarly for both ±35% and ±50% error limit the success rates are 100% for training and 98% for testing showing good generalization of the developed model.

The probability of liquefaction/non-liquefaction of the total 134 cases as obtained using the statistical CPT based methods (Robertson, Olson and Juang) and the liquefaction index (*LI*) determined for the same 134 cases by the proposed GP model are evaluated and compared in Table 2. The assessed probability is used to judge whether the prediction of occurrence of liquefaction/non-liquefaction by a particular method is correct or not on the basis of the field manifestation as obtained from the database. In this study the success rate is measured based on three criteria from stringent to liberal (A to C) i.e. P_L = 0.85-1.0 is the most stringent consideration and in the range 0.5-1.0 is the least stringent consideration for liquefied cases and similarly for non liquefied cases a prediction is considered to be successful and most stringent if P_L in the range [0, 0.15]; if P_L is within the range 0 to 0.5 then considered to be least stringent criterion. In the proposed GP approach the evaluation of liquefaction potential is characterized by *LI* and the error limit for this analysis is varied from ±15% to ±50%. This means the result is considered to be successful when the difference of target output and the predicted value come within this range. Based on the comparison presented in Table 2 all the three statistical methods are quite comparable in accuracy whereas the Juang method is more accurate than other two statistical methods. But the GP approach of prediction of occurrence of liquefaction/non-liquefaction is better than all the statistical methods considered here. In the previous ANN [7, 8] and SVM [9, 10] based studies the results were discussed in terms of *LI* > 0.5 is taken as 1.0 and *LI* < 0.5 is taken as 0.0. Hence, in the present study the results are not compared with the previous ANN and SVM results.

The Eq. 3 can be used by professionals using a spreadsheet to evaluate the liquefaction potential of soil in terms of *LI* without going into complexities of model development using GP.

CONCLUSION

Case histories of soil liquefaction due to 1999 Chi-Chi earthquake is analyzed using evolutionary artificial intelligence technique, the genetic programming to predict the liquefaction potential of soil. The results are compared with the currently used statistical methods in terms of liquefaction index. Close rate of successful prediction for training and testing data shows good generalization capabilities of GP approach. The developed GP model is found to be more efficient compared to the statistical methods in separating liquefaction and non-liquefaction cases. However, it needs more study with new data sets of different liquefaction case histories to confirm or disprove the present findings.

Table 1 Results of GP model for training and testing data

	Case records of liquefaction/ non-liquefaction	GP Performance(in terms rate of successful prediction)		
		Within ±15% error limit	Within ±35% error limit	Within ±50% error limit
Training	94	100%	100%	100%
Testing	40	95%	98%	98%

Table 2 Comparison of developed GP model results with other statistical methods

Criterion For P_L	Juang Method No.of successful prediction	Juang Method Rate (%)	Olsen Method No.of successful prediction	Olsen Method Rate (%)	Robertson Method No. of successful prediction	Robertson Method Rate (%)	Error limit for *LI* (%)	GP model No. of successful prediction	GP model Rate (%)
Based on 46 Liquefied case									
A(P_L>0.85)	39	85	19	42	33	72	±15	45	98
B(P_L>0.65)	42	92	30	66	36	79	±35	46	100
C(P_L>0.5)	45	98	39	85	40	87	±50	46	100
Based on 88 Non liquefied cases									
A(P_L<0.15)	6	7	0	0	2	3	±15	87	99
B(P_L<0.35)	35	40	20	23	10	12	±35	87	99
C(P_L<0.5)	52	59	32	37	28	32	±50	87	99
Based on all 134 cases									
A	45	34	19	15	35	27	±15	132	96
B	77	58	50	38	46	35	±35	133	99
C	97	73	71	53	68	51	±50	133	99

REFERENCES

1. Das, S. K. and Muduli, P. K. (2011), Evaluation of liquefaction potential of soil using extreme learning machine, *Proc. of the 13th Intl. conference of the IACMAG*, May. 9-11, 2011, Melbourne, Austrailia, Vol.1, 548-552.
2. Seed, H. B. and Idriss, I. M. (1971), Simplified procedure for evaluating soil liquefaction potential, *Jl. of the Soil Mechanics and Foundations Division, ASCE*, 97(9), 1249-1273.
3. Robertson, P.K. and Wride , C. E. (1998), Evaluating cyclic liquefaction potential using cone penetration test, *Jl. of the Canadian Geotech.,* 35(3), 442-459.
4. Juang, C. H., Yuan, H., Lee, D. H. and Ku, C. S. (2002), Assessing CPT-based methods for liquefaction evaluation with emphasis on the cases from the Chi-Chi, Taiwan, earthquake, *Jl. of Soil Dynamics and Earthquake Engineering,* 22, 241-258.
5. Robertson, P.K. and Campanella, R. G. (1985), Liquefaction potential of sands using the CPT*, Jl. of Geotech. Engineering,* 111(3), 384-403.
6. Stark, T. D. and Olson, S. M. (1995), Liquefaction resistance using CPT and field case histories*, Jl. of Geotechnical Engineering,* 121(12), 856-869.
7. Goh, A. T. C. (1994), Seismic liquefaction potential assessed by neural networks, *Jl. of Geotechnical Engineering,* 120(9), 1467-1480.
8. Juang, C. H., Yuan, H., Lee, D. H. and Lin, P.S. (2003), Simplified Cone Penetration Test- based method for evaluating liquefaction resistance of soils, *Jl. of Geotech. and Geoenv. Engineering,* 129(1), 66-80.
9. Pal, M. (2006), Support vector machines-based modeling of seismic liquefaction potential, *Jl. for Numerical and Analytical Methods in Geomechanics,* 30, 983-996.
10. Goh, T.C. and Goh, S. H. (2007), Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data, *Jl. of Computers and Geomechanics,* 34, 410-421.
11. Ku, C. S., Lee, D. H. and Wu, J.H. (2004), Evaluation of soil liquefaction in the Chi-Chi, Twain earthquake using CPT. *Jl. of Soil Dynamics and Earthquake Engg.*, 24, 659-673.
12. Olsen, R. S. (1997), Cyclic liquefaction based on the cone penetration test, *Proc. NCEER Workshop on Evaluation of liquefaction resistance of soils. Tech. Rep. No. NCEER-97-0022,* Buffalo. NY: National Center for Earthquake Engineering Research, State University of New York at Buffalo, 225-276.
13. Koza, J. R. (1992), *Genetic programming: on the programming of computers by natural selection*, The MIT Press, Cambridge, Mass.
14. Javadi, A. A., Rezania, M. and Nezhad, M. M. (2006), Evaluation of liquefaction induced lateral displacements using genetic programming, *Jl. of Computers and Geotechnics,* 33, 222-233.
15. Rezania, M. and Javadi, A. A. (2007), A new genetic programming model for predicting settlement of shallow foundations, *Canadian Geotech. Jl.*, 44, 1462-1473.