International Journal of Earthquake Engineering and Hazard Mitigation (IREHM), Vol. 1, N. 1
ISSN 2282-7226 September 2013

Hybridizing Genetic Programming with Orthogonal Least Squares for Modeling of Soil Liquefaction

Amir H. Gandomi[1], Amir H. Alavi[2]

Abstract – *Precise estimation of the strain energy density required to trigger soil liquefaction, denoted as capacity energy, has been the focus of many studies. The main objective of this paper is to develop a robust prediction model for the soil capacity energy using a novel hybrid technique coupling genetic programming with orthogonal least squares, called GP/OLS. The proposed model was developed upon experimental results collected through a literature review. A traditional genetic programming analysis was performed to benchmark the GP/OLS model. The predictions made by the derived model were found to be more accurate than those provided by the genetic programming and other existing models. A subsequent parametric study was carried out and the trends of the results were confirmed via some previous laboratory studies.*

Keywords: *Genetic Programming, Orthogonal Least Square, Soil Liquefaction, Capacity Energy, Formulation*

I. Introduction

Liquefaction is the phenomenon of vanishing intergranular stresses as a material response to some loading paths. These loading paths can be isochoric paths. In practical cases, the isochoric situation plays an important role since it corresponds to an undrained loading. Liquefaction commonly occurs in loose and saturated sandy soils. The presence of water appears to be necessary only to allow easy verification of the isochoric condition [1].

The liquefaction phenomenon can be caused by seismic shaking, nonseismic vibration or waved-induced shear stresses. During this phenomenon, soil immediately loses most of its strength leading to extreme deformations, flow of water and suspension of sediment [2]. According to Green [3], the available liquefaction evaluation procedures are categorized into three main groups: (1) the stress-based procedures [4], [5], (2) the strain-based procedure [6], and 3) the energy-based procedures. The focus of this study is on the energy-based procedures.

Further details about the stress and strain-based methods can be found in [7]. The energy method is based on obtaining the amount of total strain energy at the onset of liquefaction from laboratory testing or field recorded data. As it is known, hysteresis loops can be generated from the stress and strain time histories in a typical cyclic laboratory test. The strain energy (capacity energy) for each cycle of loading is equivalent to the area inside the hysteresis loop.

If this strain energy is lower than the strain energy imparted by earthquake to the sand layer during the seismic design event, the soil liquefies [8], [9].

Numerous studies have focused on deriving energy-based models for the liquefaction potential evaluation.

Green et al. [7] developed an energy-based model with only one calibration parameter on the basis of the stress controlled cyclic triaxial test data. Several relationships were proposed between the soil capacity energy, amplitude of the applied shear strain and soil initial parameters on the basis of a series of laboratory cyclic torsional shear tests [2],[10], or centrifuge tests [11], [12]. All the proposed relationships have a similar form as they were thoroughly derived by performing a multiple linear regression (MLR) analysis [9], [13].

Recently, soft computing techniques such as artificial neural networks (ANNs) have been employed to assess the soil liquefaction resistance with emphasis on energy-based hypotheses. Chen et al. [14] presented a seismic wave energy-based method with back-propagation neural networks to assess the liquefaction probability. Baziar and Jafarian [13] used ANNs to extract a correlation between soils initial parameters and the strain energy required to trigger liquefaction in sands. Alavi and Gandomi [9] proposed generalized relationships to predict the strain energy density required to trigger liquefaction using novel soft computing techniques, namely linear genetic programming and multi expression programming. This study proposes a new hybrid soft computing method that couples genetic programming (GP) and orthogonal least squares (OLS), called GP/OLS to evaluate the liquefaction capacity of sands and silty sands. The main purpose is to utilize the GP/OLS technique to generate a linear-in-parameters soil liquefaction prediction model. The results obtained by the derived model were further compared with those of the traditional GP and other existing models.

Manuscript received and revised August 2013, accepted September 2013

A comprehensive database of laboratory tests on different sands and silty sands was employed to develop the proposed correlations.

II. Genetic Programming

GP is an extension of genetic algorithms (GA) where the solutions are computer programs rather than fixed length binary strings [15]. GP automatically generates computer models based on the rules of natural genetic evolution. While GA creates a string of numbers to represent the solution, the GP solutions are computer programs represented as tree structures and expressed in a functional programming language. Recently, GP and its variants have received notable attentions from geotechnical and earthquake engineering community [16]-[19].

GP optimizes a population of computer programs according to a fitness landscape. The fitness of each program in the population is assessed using a fitness function [20], [21]. Each program evolved by GP is a structured tree as demonstrated in Fig. 1. The tree model is converted into equation y by reading it from left to right.

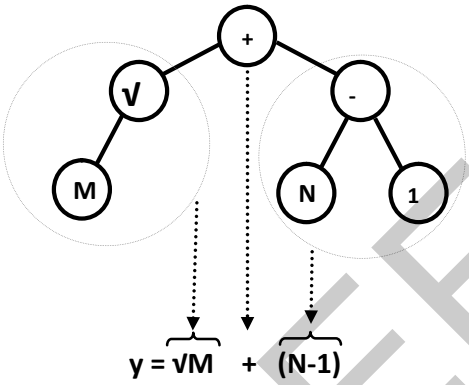


Fig. 1. Tree representation of a GP model

The creation of the initial population is a blind random search for solutions in the large space of possible solutions.

Once a population of models has been randomly created, the GP algorithm evaluates the individuals, selects individuals for reproduction, generates new individuals by mutation, crossover, and direct reproduction, and finally creates the new generation in all iterations [15], [21]. During the crossover procedure, a point on a branch of each solution (program) is randomly selected and the set of terminals and/or functions from each program are then swapped to create two new programs.

The evolutionary process continues by evaluating the fitness of the new population and starting a new round of reproduction and crossover. During the mutation process, the GP algorithm occasionally selects a function or terminal at random from a model and mutates it [20], [21].

II.1. Genetic Programming for Linear-in-Parameters Models

In general, GP creates not only nonlinear models but also linear-in-parameters models. In order to avoid parameter models, the parameters must be removed from the set of terminals. That is, it contains only variables: T = (x_0(k),..., x_i(k)}, where $x_i(k)$ denotes the i^{th} repressor variable. Hence, a population member represents only F_i nonlinear functions [20], [22]. The parameters are assigned to the model after "extracting" the F_i function terms from the tree, and determined using a least square (LS) algorithm. A simple technique for the decomposition of the tree into function terms can be used. The subtrees, representing the F_i function terms were determined by decomposing the tree starting from the root as far as reaching nonlinear nodes (nodes which are not "+" or "-") [20]. As can be seen in Fig. 2, the root node is a "+" operator; therefore, it is possible to decompose the tree into two subtrees of "A" and "B".

The root node of the "A" tree is anew a linear operator; therefore, it can be decomposed into "C" and "D" trees. As the root node of the "B" tree is a nonlinear node (/), it cannot be decomposed. The root nodes of "C" and "D" trees are also nonlinear. Consequently, the final decomposition procedure results in three subtrees: "B", "C", and "D".

According to the results of the decomposition, it is possible to assign parameters to the functional terms represented by the obtained subtrees. The resulted linear-in-parameters model for this example is y: $p_0 + p_1(x_2 + x_1)/x_0 + p_2x_0 + p_3x_1$ [20].

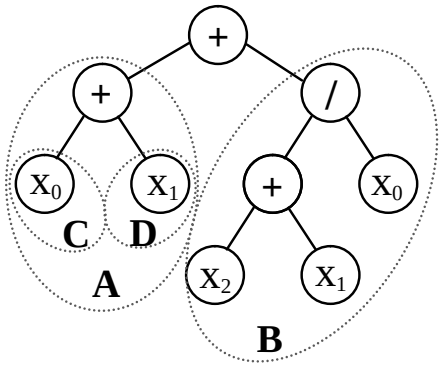


Fig. 2. Decomposition of a tree to function terms [24]

GP can be used for selecting from special model classes, such as a polynomial model. To achieve it, the set of operators must be restricted and some simple syntactic rules must be introduced. For instance, if the set of operators is defined as F= {×, +} and there is a syntactic rule that exchanges the internal nodes that are below a "×"-type internal nodes to "×"- type nodes, GP will generate only polynomial models [15], [20], [23].

II.2. Orthogonal Least Squares Algorithm (OLS)

Refers to the example for the dimension and the

position.

The great advantage of using linear-in-parameter models is that the LS method can be used for identifying the model parameters, which is much less computationally demanding than other nonlinear optimization algorithms, because the optimal $p = [p_1,..., p_m]^T$ parameter vector can be analytically calculated (Gandomi et al. 2010) [20]:

$$p = (U^{-1}U)^T U_y \tag{1}$$

in which $y = [y(1),..., y(N)]^T$ is the measured output vector and the U regression matrix is:

$$U = \begin{pmatrix} U_1(x(1)) & \cdots & U_M(x(1)) \\ \vdots & \ddots & \vdots \\ U_1(x(N)) & \cdots & U_M(x(N)) \end{pmatrix} \tag{2}$$

The OLS algorithm [24], [25] (Billings et al., 1988; Chen et al., 1989) is an effective algorithm for determining which terms are significant in a linear-in-parameters model.

The OLS technique introduces the error reduction ratio (err), which is a measure of the decrease in the variance of output by a given term [20].

The matrix form corresponding to the linear-in-parameters model is:

$$y = U_p + e \tag{3}$$

where the U is the regression matrix, p is the parameter vector, and e is the error vector. The OLS method transforms the columns of the U matrix into a set of orthogonal basis vectors to inspect the individual contributions of each term [20], [26].

It is assumed in the OLS algorithm that the regression matrix U can be orthogonally decomposed as U = WA, where A is a M by M upper triangular matrix (i.e., $A_{ij} = 0$ if i > j) and W is a N by M matrix with orthogonal columns in the sense that WTW = D is a diagonal matrix (N is the length of the y vector and M is the number of repressors).

After this decomposition, the OLS auxiliary parameter vector g can be calculate as [20]:

$$g = D^{-1}W^T y \tag{4}$$

where g_i represents the corresponding element of the OLS solution vector. The output variance $(y^T y)/N$ can be described as:

$$y^T y = \sum_{i=1}^{M} g_i^2 w_i^T w_i + e^T e \tag{5}$$

Therefore, the error reduction ratio $[err]_i$ of the U_i term can be expressed as:

$$[err]_i = \frac{g_i^2 w_i^T w}{y^T y} \tag{6}$$

This ratio offers a simple mean for order and selects the model terms of a linear-in-parameters model on the basis of their contribution to the performance of the model [20].

II.3. Hybrid Genetic Programming-Orthogonal Least Squares Algorithm (GP/OLS)

Madár et al. [27], [28] combined GP and OLS to make a hybrid algorithm with better efficiency. It was shown that introducing this strategy into the GP process results in more robust and interpretable models [27].

The main feature of this hybrid approach is to transform the trees to simpler trees which are more transparent, but their accuracies are close to the original trees. In this coupled algorithm, GP generates a lot of potential solutions in the form of a tree structure during the GP operation. These trees may have better and worse terms (subtrees) that contribute more or less to the accuracy of the model represented by the tree. OLS is used to estimate the contribution of the branches of the tree to the accuracy of the model, whereas, using the OLS, one can select the less significant terms in a linear regression problem [20]. According to this strategy, terms (subtrees) having the smallest error reduction ratio are eliminated from the tree [22]. This "tree pruning" approach is realized in every fitness evaluation before the calculation of the fitness values of the trees [20]. Since GP works with the tree structure, the further goal is to preserve the original structure of the trees as far as it possible. The GP/OLS method always guarantees that the elimination of one or more function terms of the model can be done by pruning the corresponding subtrees, so there is no need for structural rearrangement of the tree after this operation. The way the GP/OLS method works on its basis is simply demonstrated in Fig. 3. Assume that the function which must be identified is $y(x) = 0.8u(x-1)^2 + 1.2y(x-1) - 0.9y(x-2) - 0.2$.

As can be seen in Fig. 3, the GP algorithm found a solution with four terms: $u(x-1)^2$, $y(x-1)$, $y(x-2)$, $u(x-1) \times u(x-2)$. Based on the OLS algorithm, the subtree with the least error reduction ratio ($F_4 = u(x-1) \times u(x-2)$) is eliminated from the tree. Subsequently, the error reduction ratios and mean square error values (and model parameters) are calculated again. The new model (after pruning) may have a higher mean square error but it obviously has a more adequate structure [20].

III. Model Development & Analysis

This study considers the feasibility of using the GP/OLS and the traditional tree-based GP approaches to obtain meaningful relationships between the level of energy required for liquefaction of sands and the influencing parameters.

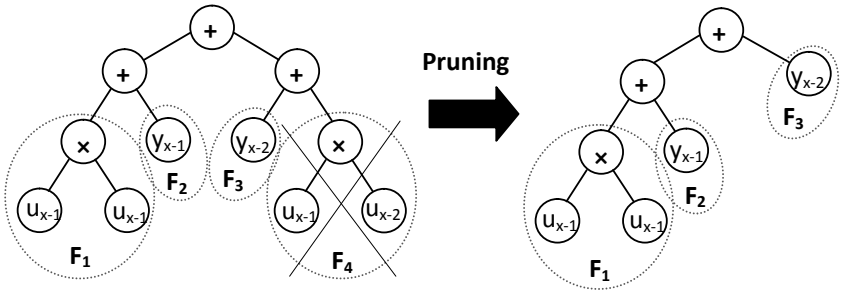


Fig. 3. Pruning of a tree with OLS [20]

Two of the most widely used parameters in the available energy-based pore pressure build-up models (e.g. [2], [12]) for the liquefaction assessment were considered as the input parameters. Hence, the formulation of liquefaction capacity energy will be as follows:

$$Log(W) = f(\sigma'_{mean}, D_r) \quad (7)$$

where:

W: Measured strain energy density required for triggering liquefaction (capacity energy). This capacity energy is the accumulative area of stress–strain loops up to the liquefaction triggering.

σ'_{mean}: Soil initial effective mean confining pressure

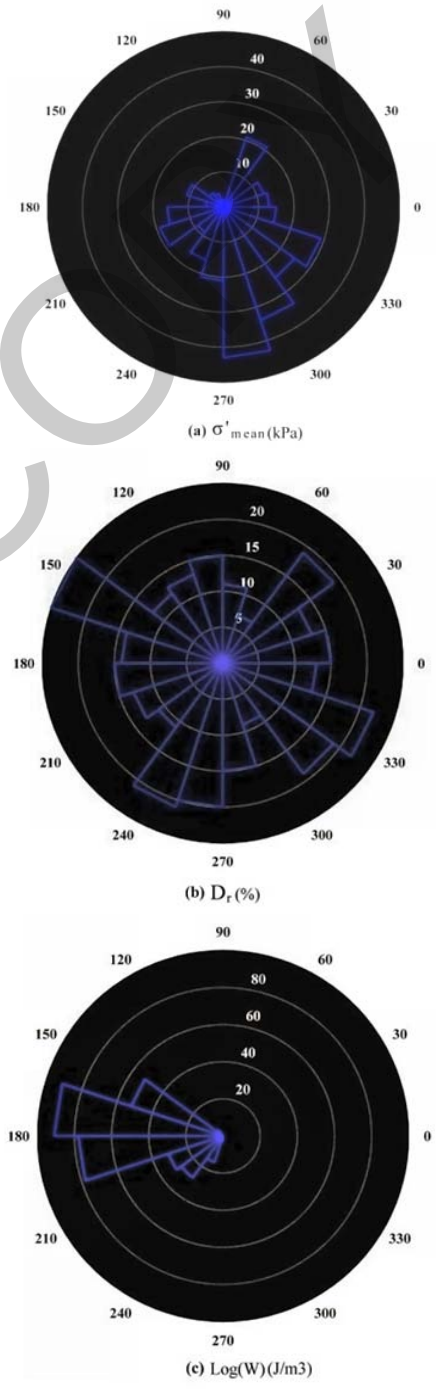
D_r: Initial relative density after consolidation

The significant influence of σ'_{mean} and D_r in determining W is well understood [2], [12], [29]. Correlation coefficient (R), root mean square error (RMSE) and mean absolute percent error (MAPE) were used to evaluate the capabilities of the proposed correlations.

III.1. Experimental Database

A comprehensive database of previously published cyclic tests was used for the development of the proposed correlations. A major part of this database has been used by Baziar and Jafarian [13] for the ANN-based modeling of the strain energy. The database consists of 237 cyclic triaxial [3], [8], 61 cyclic torsional shear [2], [30], 6 cyclic simple shear (VELACS project) [31], and 18 liquefaction triggering centrifuge [11] tests data. These are a total of 322 cyclic triaxial, torsional, simple shear and centrifuge ground level liquefaction element tests on Monterey, Yatesville, Reid Bedford, LSFD, LSI-30, Toyoura and Nevada 40% clean and silty sands. The database includes the measurements of several variables such as σ'_{mean} (kPa), D_r (%) and W (J/m^3).

To visualize samples distribution, the data are presented by angle histogram plots (Figures 4(a)–(c)). The ranges and statistics of different input and output parameters involved in the model development are given in Table I. Furthermore, the table contains data of some element tests under random loading. The criteria for liquefaction triggering (failure) is initial liquefaction (r_u = 1) or double amplitude of strain of 5% (ε_{DA} = 5%), whichever occurs first [9], [13].



Figs. 4. The angle histograms of input and output variables

III.2. Building GP and GP/OLS Models

The available database was used for the training and testing the GP and GP/OLS prediction correlations.

There are various parameters involved in the GP and GP/OLS algorithm. The parameter selection affects the generalization capability of the models. These parameters were selected based on some previously suggested values [20], [28], [32], [33] and also after trial and error approach. The parameter settings are shown in Table II. For the GP and GP/OLS analyses, the data sets were randomly divided into the training and testing subsets.

In order to obtain a consistent data division, several combinations of the training and testing sets were considered. The selection was such that the maximum, minimum, mean and standard deviation of parameters were consistent in training and testing data sets [13]. Out of the 322 data, 242 data (75%) are used as training and 80 data (25%) for the testing of the generalization capability of the GP and GP/OLS-based correlations.

For better results, both the input and the output variables are normalized in this study. The formula used to normalize the variables to a range of [U, L] is as follows [21]:

$$x_{normalized} = ax + b \tag{8}$$

where:

$$a = \frac{U - L}{x_{max} - x_{min}} \tag{9}$$

$$b = U - ax_{max} \tag{10}$$

in which x_{max} and x_{min} are the maximum and minimum values of the variable. In this study, U = 0.9 and L = 0.1.

The GP and GP/OLS-based equations were obtained by converting the related expression tree into a mathematical form recording the nodes from left to right in each layer of the expression tree, from root layer down to the deepest one to form the string. The best GP and GP/OLS-based formulas were chosen on the basis of a multi-objective strategy as below [9]:

i. Involving all input variables.
ii. Providing the best fitness value on the training set of data.

The formulation of the strain energy density required for triggering liquefaction, W (J/m^3), for the best result by the traditional GP algorithm is as given below:

$$Log(W)_{GP} = \left[1.1759\left(\sigma'_{mean} + D_r + -0.4147\left(\sigma'_{mean} + D_r\right)^2\right)\right]^4 - 0.5854 \tag{11}$$

where σ'_{mean} and D_r, respectively, denote the soil initial effective mean confining pressure and initial relative density after consolidation. A comparison of GP predicted against experimental liquefaction capacity energy is shown in Fig. 5.

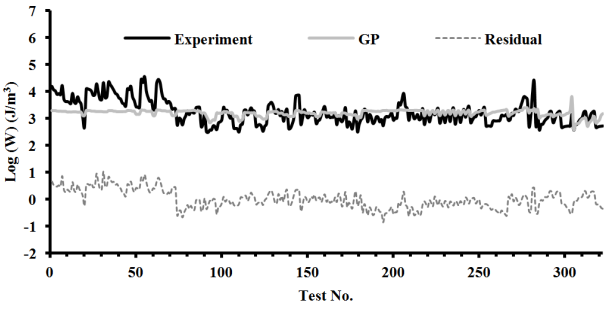


Fig. 5. Traditional GP predicted versus experimental liquefaction capacity energy

In this figure, residual is the difference between the experimental and predicted capacity energy.

TABLE I
THE VARIABLES USED IN MODEL DEVELOPMENT

Parameters	Range	Minimum	Maximum
Inputs			
σ'_{mean} (kPa)	266.2	27.8	294
D_r (%)	149.6	-44.5	105.1
Output			
$Log(W)$(J/m^3)	2.07	2.47	4.53

TABLE II
PARAMETER SETTINGS FOR THE GP AND GP/OLS MODELS

Parameter	Settings
Function set	+, -, ×, /
Population size	500-1000
Maximum tree depth	64
Maximum number of evaluated individuals	250
Generation	50
Type of selection	roulette-wheel
Type of mutation	point-mutation
Type of crossover	one-point (2 parents)
Type of replacement	elitist
Probability of crossover	0.5
Probability of mutation	0.5
Probability of changing terminal–non-terminal nodes (vice versa) during mutation	0.25

III.3. GP/OLS Model for Capacity Energy

The prediction equation of W (J/m^3) for the best result by the GP/OLS algorithm is given as follows:

$$Log(W)_{GP/OLS} = -0.000023384\sigma'^2_{mean} + 0.0112\sigma'_{mean} + 0.0072D_r + 2.0627 \tag{12}$$

Fig. 6 presents a comparison of the liquefaction capacity energy predicted by GP/OLS versus experimental values.

III.4. Comparison of The Capacity Energy Models

Performance statistics of the formulas obtained by GP and GP/OLS and other well-know MLR-based equations found in the literature [2], [12] are summarized in Table III. It was not possible to evaluate some other energy-based pore pressure build-up models by the present database, since they needed calibration parameters.

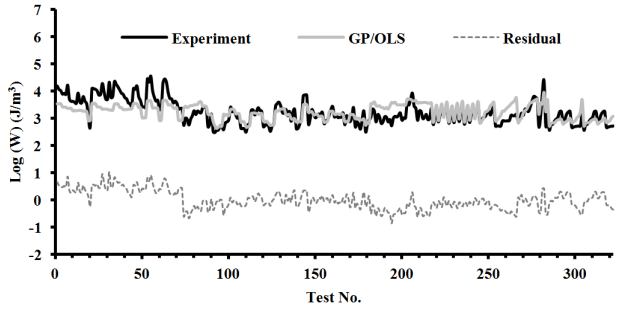


Fig. 6. GP/OLS predicted and experimental liquefaction capacity energy

Additionally, the relationships proposed for the calibration parameters in some of those models yielded very approximate results for the present database. However, comparing the performance of the proposed relationships, it can be seen from Table III that the GP/OLS model has produced better results than the traditional GP model on the training, testing and whole of data. It can be observed that the GP and GP/OLS-based formulas provide considerably better results compared to the existing regression models. The GP/OLS model provides other significant advantages. Because of tree pruning process, the GP/OLS-based equation is really short, very simple and can be used for routine design practice via hand calculations, especially in comparison with the traditional GP.

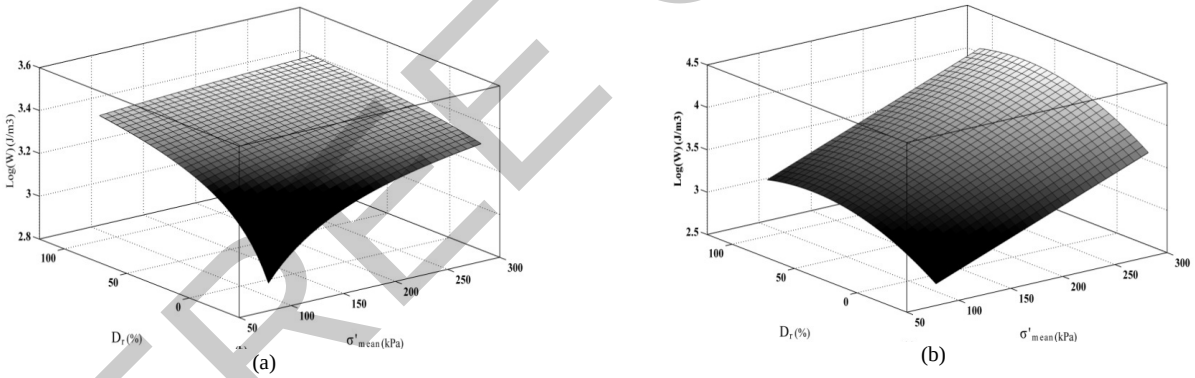
III.5. Parametric Study

For further verification of the GP and GP/OLS-based models, a parametric study was performed in this study. In Figs. 7(a) and (b), three dimensional surface graphs of σ'_{mean}, D_r and W are presented. The sensitivity of W prediction to σ'_{mean} and D_r can be determined referring to these figures. The results of the parametric study for these models indicate that the increases in σ'_{mean} and D_r cause amplification in the capacity energy of sands.

These results are in close agreement with the results of laboratory studies carried out by other researchers (e.g., [2],[34],[35].

TABLE III
PERFORMANCE OF ENERGY-BASED MODELS FOR LIQUEFACTION ASSESSMENT

	Train			Test			All		
Model	R	RMSE	MAPE	R	RMSE	MAPE	R	RMSE	MAPE
Liang (1995)	0.387	0.491	0.359	0.337	0.574	0.436	0.383	0.512	0.376
Dief and Figueroa (2001)	0.444	0.492	0.360	0.486	0.572	0.430	0.427	0.519	0.381
GP, Eq. (13)	0.595	0.397	0.319	0.642	0.405	0.308	0.605	0.400	0.317
GP/OLS, Eq. (14)	0.604	0.355	0.288	0.665	0.342	0.280	0.618	0.352	0.286



Figs. 7. Three-dimensional surface maps of capacity energy and related parameters in : (a) GP model (b) GP/OLS model

IV. Conclusions & Future Directions

In this paper, a combined GP and OLS algorithm, called GP/OLS, was employed to assess the amount of strain energy required up to liquefaction triggering. A reliable database of previously published cyclic tests results was used to develop the correlations. The most widely used parameters in the energy-based pore pressure build-up models (σ´mean and Dr) were considered as the input parameters. In order to assess the capability of the GP/OLS model, a new traditional tree-based GP model was developed. Due to high nonlinearity in liquefaction development, the proposed nonlinear GP/OLS-based correlation produces considerably better outcomes over the existing linear energy-based pore pressure build-up models.

The GP/OLS model also outperforms the proposed GP model. The results of the parametric study were confirmed with the results of experimental studies presented by other researchers. Further research can be focused on both the problem domain and the computing one. As more data become available, including those for other case histories, the same models can be improved to make more accurate predictions for a wider range. Since fuzzy logic can provide a systematic method to deal with imprecise and incomplete information, the process of developing a hybrid fuzzy-GP or GP/OLS model for such problems can be a suitable topic for further studies.

References

[1] Darve, F. (1996). Liquefaction phenomenon of granular materials and constitutive instability. *Engineering Computation*, 13(7), 5-28.

[2] Liang, L. (1995). *Development of an energy method for evaluating the liquefaction potential of a soil deposit*. PhD dissertation, Department of Civil Engineering, Case Western Reserve University, Cleveland, OH.

[3] Green, R. A. (2001). *Energy-based evaluation and remediation of liquefiable soils*. PhD dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

[4] Seed, H. B., & Idriss, I. M. (1971). Simplified procedure for evaluating soil liquefaction potential. *Journal of the Soil Mechanics and Foundations Division*, 97(SM8), 1249–1274.

[5] Whitman, R. V. (1971). Resistance of soil to liquefaction and settlement. *Soils and Foundations*, 11(4), 59–68.

[6] Dobry, R., Ladd, R. S., Yokel, F. Y., Chung, R. M., & Powell, D. (1982). *Prediction of pore water pressure build-up and liquefaction of sands during earthquakes by the cyclic strain method*. Building science series, Washington, DC: National Bureau of Standards, US Department of Commerce, US Governmental Printing Office, Vol. 138.

[7] Green, R. A., Mitchell, J. K., & Polito, C. P. (2000). Energy-based excess pore pressure generation model for cohesionless soils. In: John Booker memorial symposium (pp. 16–17), Sydney, NSW, Australia.

[8] Ostadan, F., Deng, N., & Arango, I. (1996). Energy-Based Method for Liquefaction Potential Evaluation, Phase I . Feasibility Study. U.S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Building and Fire Research Laboratory.

[9] Alavi A.H., Gandomi A.H., "Energy-Based Models for Assessment of Soil Liquefaction." *Geoscience Frontiers*, 3(4): 541-555, 2012.

[10] Figueroa, J. L., Saada, A. S., Liang, L., & Dahisaria, M. N. (1994). Evaluation of soil liquefaction by energy principles. *Journal Geotechnical Engineering ASCE*, 20(9),1554–1569.

[11] Dief, H. M. (2000). *Evaluating the liquefaction potential of soils by the energy method in the centrifuge*. PhD Dissertation, Department of Civil Engineering, Case Western Reserve University, Cleveland, OH.

[12] Dief, H. M., & Figueroa, J. L. (2001). Liquefaction assessment by the energy method through centrifuge modeling. In X. Zeng (Ed.), NSF international workshop on earthquake simulation in Geotechnical Engineering, CWRU, Cleveland, OH.

[13] Baziar, M. H., & Jafarian, Y. (2007). Assessment of liquefaction triggering using strain energy concept and ANN model: Capacity Energy. *Journal of Soil Dynamics and Earthquake Engineering*, 27(12),1056–72.

[14] Chen, Y. R., Hsieh, S. C., Chen, J. W., & Shih, C. C. (2005). Energy-based probabilistic evaluation of soil liquefaction. *Journal of Soil Dynamics and Earthquake Engineering*, 25(1):55–68.

[15] Koza, J. (1992). Genetic programming, on the programming of computers by means of natural selection. Cambridge (MA): MIT Press.

[16] Yang X.S., Gandomi A.H., Talatahari S., Alavi A.H., "*Metaheuristics in Water Resources, Geotechnical and Transportation Engineering*", 496 pages, Elsevier,Waltham, MA, USA, 2012.

[17] Gandomi A.H., Yang X.S., Talatahari S., Alavi A.H., "Metaheuristic Applications in Structures and Infrastructures", 430 pages, Elsevier,Waltham, MA, USA, 2013.

[18] Gandomi A.H., Alavi A.H., "Multi-Stage Genetic Programming: A New Strategy to Nonlinear System Modeling." *Information Sciences, Elsevier*, 181(23): 5227-5239, 2011.

[19] Gandomi A.H., Alavi A.H., "A New Multi-Gene Genetic Programming Approach to Nonlinear System Modeling. Part II: Geotechnical and Earthquake Engineering Problems." *Neural Computing & Applications, Springer*, 21:189-201, 2012.

[20] Gandomi A.H., Alavi A.H., Arjmandi P., Aghaeifar A., Seyednoor, M., "Genetic Programming and Orthogonal Least Squares: A Hybrid Approach to Modeling the Compressive Strength of CFRP-Confined Concrete Cylinders." *Journal of Mechanics of Materials and Structures*, Mathematical Sciences, UC Berkeley, 5(5), 735–753, 2010.

[21] Alavi, A.H., Gandomi, A.H., Mollahasani, A., (2011). A Genetic Programming-Based Approach for Performance Characteristics Assessment of Stabilized Soil. Variants of Evolutionary Algorithms for Real-World Applications, Springer-Verlag, Berlin, Chapter 9, 343-375.

[22] Pearson, R. (2003). Selecting nonlinear model structures for computer control. *Journal of Process Control*, 13(1), 1–26.

[23] Madár, J., Abonyi, J., & Szeifert, F. (2004). Genetic programming for system identification, In *Intelligent Systems Design and Applications* (ISDA 2004) Conference, Budapest, Hungary.

[24] Billings, S., Korenberg, M., & Chen, S. (1988). Identification of nonlinear outputaffine systems using an orthogonal least-squares algorithm, *International Journal of Systems Science*, 19(8), 1559–1568.

[25] Chen, S., Billings, S., & Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification, *International Journal of Control*, 50(5), 1873–1896.

[26] Cao, H., Yu, J., Kang, L., & Chen, Y. (1999). The kinetic evolutionary modelling of complex systems of chemical reactions", *Computers and Chemical Engineering*, 23(1), 143–151.

[27] Madár, J., Abonyi, J., & Szeifert, F. (2005). Genetic Programming for the Identification of Nonlinear Input-Output Models", *Industrial & Engineering Chemistry Research*, 44(9), 3178-3186.

[28] Madár, J., Abonyi, J., & Szeifert, F. (2005). Genetic Programming for the Identification of Nonlinear Input-Output Models, white paper, Available online at: http://www.fmt.vein.hu/softcomp/gp/ie049626e.pdf

[29] Nemat-Nasser, S., & Shokooh, A. (1979). A unified approach to densification and liquefaction of coheshionless sand in cyclic shearing. *Canadian Geotechnical Journal*,16(4), 659–678.

[30] Towhata, I., & Ishihara, K. (1985). Shear work and pore water pressure in undrained shear. *Soils and Foundations*, 25(3), 73–84.

[31] Arulmoli, K., Muraleetharan, K. K., Hosain, M. M., & Fruth L. S. (1992). VELACS laboratory testing program - soil data report. The Earth Technology Corporation, Irvine, Calif. *Report to the National Science Foundation*, Washington, DC.

[32] Mousavi S.M., Alavi A.H., Gandomi A.H., Mollahasani A., "Formulation of Soil Angle of Shearing Resistance Using a Hybrid GP and OLS Method." *Engineering with Computers*, 29(1), 37-53, 2013.

[33] Mousavi S.M., Gandomi A.H., Alavi A.H., Vesalimahmood M., "Modeling of Compressive Strength of HPC Mixes Using a Combined Algorithm of Genetic Programming and Orthogonal Least Squares." *Structural Engineering and Mechanics*, 36(2), 225-241, 2010.

[34] Lee, K. L., & Seed, H. B. (1967). Cyclic stress conditions causing liquefaction of sand. *Journal of the Soil Mechanics and Foundations Division*, 93(SMl), 47–70.

[35] Polito, C. P., & Martin, J. R. (2001). Effects of non-plastic fines on the liquefaction resistance of sands *Journal of Geotechnical and Geoenvironmental Engineering*, 127(5), 408–415.

Authors' information

[1]Department of Civil Engineering, University of Akron, Akron, OH 44325, USA.

[2]Department of Civil and Environmental Engineering, Engineering Building, Michigan State University, East Lansing, MI, 48824, USA.

