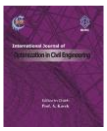
INTERNATIONAL JOURNAL OF OPTIMIZATION IN CIVIL ENGINEERING
Int. J. Optim. Civil Eng., 2016; 6(3):385-403



AN OPTIMIZED NEURO-FUZZY GROUP METHOD OF DATA HANDLING SYSTEM BASED ON GRAVITATIONAL SEARCH ALGORITHM FOR EVALUATION OF LATERAL GROUND DISPLACEMENTS

M. Goharriz and S.M. Marandi[*, †]
Department of Civil Engineering, Shahid Bahonar University of kerman, Kerman, Iran

ABSTRACT

During an earthquake, significant damage can result due to instability of the soil in the area affected by internal seismic waves. A liquefaction-induced lateral ground displacement has been a very damaging type of ground failure during past strong earthquakes. In this study, neuro-fuzzy group method of data handling (NF-GMDH) is utilized for assessment of lateral displacement in both ground slope and free face conditions. The NF-GMDH approach is improved using gravitational search algorithm (GSA). Estimation of the lateral ground displacements requires characterization of the field conditions, principally seismological, topographical and geotechnical parameters. The comprehensive database was used for development of the model obtained from different earthquakes. Contributions of the variables influencing the lateral ground displacement are evaluated through a sensitivity analysis. Performance of the NF-GMDH-GSA models are compared with those obtained from gene-expression programming (GEP) approach, and empirical equations in terms of error indicators parameters and the advantages of the proposed models over the conventional method are discussed. The results showed that the models presented in this research may serve as reliable tools to predict lateral ground displacement. It is clear that a precise correlation is easier to be used in the routine geotechnical projects compared with the field measurement techniques.

Keywords: earthquake; liquefaction; lateral ground displacement; NF-GMDH; GSA; GEP: empirical equation.

Received: 20 November 2015; Accepted: 10 January 2016

[*]Corresponding author: Civil Engineering Department, Shahid Bahonar University of kerman, Kerman, Iran
[†]E-mail address: marandi@uk.ac.ir (S.M. Marandi)

1. INTRODUCTION

Liquefaction is known as one of the major causes of ground failure due to earthquake. Assessment of liquefaction potential and determination of liquefaction induced lateral spreading are complex geotechnical engineering problems and have attracted considerable attention of geotechnical researchers in the past three decades [1].

During strong earthquakes (1964 Niigata, 1964 Alaska, 1971 San Fernando, 1983 Nihonkai Chubu, 1989 Loma Prieta, 1990 Luzon, 1995 Hyogoken-Nambu, 1999 Izmit, etc.), liquefaction-induced lateral spreads caused heavy and tremendous damages to civil engineering structures and facilities such as buried pipeline networks, pile foundations and quay walls.

Due to a large number of factors, the determination of liquefaction-induced lateral displacement is a complex geotechnical engineering problem. Several methods have been developed to predict lateral ground displacements using analytical [2], laboratory [3, 4], and finite element methods [5]. However, these methods have not been able to estimate lateral ground displacements caused by liquefaction with a good accuracy.

Assessment of liquefaction potential and determination of liquefaction induced lateral ground displacements have been considered by many researchers [1, 6, 7, and 8]. Empirical models based on case histories have remained the more popular methods [8, 9, 10, 11, 12, 13, and 14]. For assessment of liquefaction induced lateral ground displacements, empirical correlations and multi-linear regression models were introduced [15]. With eight major earthquakes database, which happened between 1906 and 1987 in the Japan and U.S.A., and using multiple linear regression analyses (MLR), three regression equations were developed for free-face, ground slope, and combination of these two models [6]. On a different note, Zhang et al. introduced a "Lateral Displacement Index (LDI)" calculated by integration of maximum shear strain over potentially liquefiable layers based on empirical correlation on a cumulative shear strain model, and then used it in a couple of simple correlations for free-face and sloping ground cases [16]. In other researches, a different cumulative strain model used to arrive at LDI [17, 18].

The progress of advanced computational methods for problems analysis has necessitated the accurate determination and estimation of lateral ground displacement. In recent years, new aspects of modeling, optimization, and problem solving have been evolved in light of the pervasive development in computational software and hardware. These aspects of software engineering are referred to as soft computing based methods such as artificial intelligence, which is a powerful tool for multivariate and nonlinear modeling. In case of complicated problems, experimentalists prefer these trial approaches rather than analytical optimization. A large number of researchers applied artificial intelligence (AI) models in various fields of geotechnical engineering such as stress-strain modeling of soil [19], slope stability [20], shallow foundations [21], and liquefaction [22].

In the past years, the GMDH networks provided successful evaluations in various field of geotechnical engineering sciences such as prediction of the scour depth around hydraulic structures and estimation of the S_u-N_{SPT} correlation [23]. Application of the GMDH networks yielded relatively precise estimations than those obtained using empirical equations based on regressive models. The main concern of the GMDH network is to present analytical solutions for various problems within a feed forward network in the form of quadratic polynomial

whose weighting coefficients are obtained using regression method [23].

Structure of the GMDH network has been improved based on multi-stage fuzzy decision rule as neuro-fuzzy GMDH to obtain physical insights of problems with high degree of complexity. The NF-GMDH networks have been successfully applied to the different problems such as grinding characteristics, forecasting the unreliable mobile communication, prediction of longitudinal dispersion in water networks, prediction of discharge flow in compound channels, local scour depth at group pier under waves and current flow, and maximum scour depth at downstream of sluice gates and grade-control structures. The neuro-fuzzy GMDH has higher flexibility and lower complexity compared to the GMDH network. The other advantages of the NF-GMDH models were presented in literatures [24].

In cases of practical contributions, the NF-GMDH model in the field of geotechnical engineering has not been applied yet. In this study, a computer program is coded for the NF-GMDH network with MATLAB. Also, the GSA model is applied in topology design of the NF-GMDH model for prediction of the lateral ground spread. Results of the proposed NF-GMDH-GSA are compared with those obtained using the best formulation of the gene-expression programming (GEP) model. The performance of the proposed NF-GMDH-GSA is also evaluated with empirical equations based on regression models. The results showed that the GMDH models are able to learn, with a very high accuracy, the complex relationship between lateral ground displacement and its contributing factors; and generalize the learning to provide, predictions for new cases that are not used in the construction of the model, however the new contribution dictates that a precise correlation can be an easier method to be used in the routine geotechnical projects in comparison with the field measurement techniques.

2. INFLUENTIAL PARAMETERS AND DATABASE DEVELOPMENT

A thorough understanding of the factors affecting lateral ground displacement is needed in order to obtain accurate lateral displacement estimations. Based on previous researches [e.g., 25 and 26], the most important factors that affect the lateral ground displacement due to liquefaction can be categorized as moment magnitude of the earthquake (M), the nearest distance to the seismic energy source (R), the cumulative thickness of saturated granular layers with corrected blow counts of SPT less than 15 (T_{15}), the average fines content for granular materials included within T_{15} (F_{15}), the average mean size of granular materials within T_{15} ($D50_{15}$), the ground slope (S) and the free-face ratio (W). In this research the same factors are presented for the NF-GMDH-GSA and GEP models as input variables. Lateral ground displacement (D_h) is the single output variable.

A wide-range database was compiled from previously different earthquakes (1906 San Francisco, 1964 Anchorage, 1964 Niigata, 1971 San Fernando, 1979 Imperial Valley, 1983 Nihonkai-Chubu, 1983 Borah Peak, 1987 Superstition Hills, 1989 Loma Prieta, and 1995 Hyogo-Ken Nanbu).

It is common practice to divide the available data into two subsets; a training set to construct the soft computing model, and an independent validation set to estimate the performance of the trained model. In this study 80% of the data set was used for training and 20% for validation of the model. The data division process was performed so that the main

statistical parameters of the training and testing subsets (i.e., maximum, minimum, mean, and standard deviation) become close to each other. For this purpose, a trial selection procedure was carried out and the most possible consistent division was determined [27].

The case histories involving the lateral displacement towards a free face and those corresponding to gently sloping ground, have been analyzed separately. The lateral ground displacement database includes 426 case histories gathered from the literature. In the collected database, 219 cases are related to sloping ground condition and the 207 cases involve free face ground. Descriptive statistics of these two groups variables used in the model development for both sloping ground and free face conditions are presented in Table 1 and Table 2, respectively.

3. DESCRIPTIONS OF THE NF-GMDH MODEL

The GMDH network is a learning machine based on the principle of heuristic self-organizing. Also, it is a series of operations of seeding, rearing, crossbreeding, selection and rejection of seeds correspond to the determination of the input variables, and structure and parameters of model, and selection of model by principle of termination. The other descriptions of the GMDH network were presented in literatures [28].

In this research a neuro-fuzzy GMDH model based on PSO algorithm has been proposed for the lateral ground spread prediction. The structure of neuro-fuzzy GMDH is constructed automatically using heuristic self-organized algorithm. The neuro-fuzzy GMDH network is a very flexible algorithm, and it can be hybridized easily by other iterative and evolutionary algorithms. Furthermore, a simplified fuzzy reasoning rule is utilized to improve the GMDH network as follows [29]:

Table 1: Descriptive statistical analysis of parameters used for development of free face model

Statistical parameters	subset	Inputs						Output
		M	R (km)	W (%)	T_{15} (m)	F_{15} (%)	$D50_{15}$ (mm)	D_h (m)
Max.	Training	9.2	100	56.8	16	70	1.98	10.16
	Validation	9.2	95	41.38	16.7	47	1.98	8.39
Min.	Training	6.4	0.5	1.64	0.2	1	0.04	0.01
	Validation	6.4	0.5	2.27	0.5	3	0.07	0.01
Mean	Training	7.17	17.38	11.31	8.2	18.13	0.37	2.61
	Validation	7.27	21.13	10.83	9.02	16.78	0.33	2.1
S.D.	Training	0.55	15.5	9.38	4.87	13.92	0.44	2.36
	Validation	0.53	19.1	8.45	4.86	11.94	0.33	1.95

Table 2: Descriptive statistical analysis of parameters used for development of sloping ground model

Statistical parameters	subset	Inputs						Output
		M	*R* (*km*)	*W* (%)	T_{15} (*m*)	F_{15} (%)	$D50_{15}$ (*mm*)	D_h (*m*)
Max.	Training	9.2	100	11	19.7	59	10	3.36
	Validation	7.7	65	11	11.6	68	12	3.55
Min.	Training	6.4	0.2	0.05	0.01	0	0.06	0.01
	Validation	6.4	0.2	0.21	0.7	0	0.06	0.01
Mean	Training	7.55	24.91	0.92	6.63	9	0.38	1.92
	Validation	7.43	21.69	1.54	5.95	11.66	0.7	1.77
S.D.	Training	0.33	13.1	1.2	3.94	10.41	0.74	1.03
	Validation	0.34	10.35	2.68	3.95	15.16	1.84	1.01

If x_1 is F_{k1} and x_2 is F_{k2}, then, output of y is w_k .Gaussian membership function is used in term of F_{kj} which is related to the *k*th fuzzy rules in the domain of the *j*th input values x_j:

$$F_{kj}(x_j) = \exp(-(x_j - a_{kj})^2 / b_{kj}) \tag{1}$$

which a_{kj} and b_{kj} are constant values for each rules. Also, the y parameter is defined as output that is expressed as follows:

$$y = \sum_{k=1}^{K} u_k w_k \tag{2}$$

$$u_k = \prod_j F_{kj}(x_j) \tag{3}$$

which, w_k is a real value for *k*th fuzzy rules [31].

The NF-GMDH model is one of the adaptive learning networks that have hierarchical structure. In this model, each neuron has two input variables and one output. General configuration of the neuro-fuzzy GMDH with two fuzzy rules for each partial description (*PD*) is presented in Fig. 1.

Through Fig. 1, output of each neuron in a layer is considered as the input variable in the next layer. The final output is calculated using the average of the outputs from the last layer. Fig. 1 indicates that the inputs from the *m*th model and *p*th layer are the output variables of the (*m-1*)th and *m*th model in the (*p-1*)th layer. The mathematical function for calculating y^{pm} is:

$$y^{pm} = f(y^{p-1,m-1}, y^{p-1,m}) = \sum_{k=1}^{K} \mu_k^{pm} . w_k^{pm} \tag{4}$$

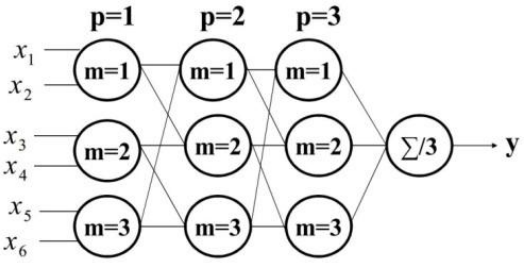


Figure 1. General structure of the NF-GMDH

which, μ_k^{pm} and w_k^{pm} are the *k*th Gaussian function and its corresponding weight parameter, respectively, and are related to the *m*th model in the *p*th layer. In addition, a_k^{pm} and b_k^{pm} are the Gaussian parameters that are utilized for the *i*th input variable from the *m*th model and *p*th layer. Also, the final output is expressed using the following function:

$$\mu_k^{pm} = \exp\left\{-\frac{(y^{p-1,m-1} - a_{k,1}^{pm})^2}{b_{k,1}^{pm}} - \frac{(y^{p-1,m} - a_{k,2}^{pm})^2}{b_{k,2}^{pm}}\right\} \tag{5}$$

$$y = \frac{1}{M}\sum_{m=1}^{M} y^{pm} \tag{6}$$

Neuro-fuzzy GMDH network is known as an iterative method to solve the complicated systems. In iteration, the error parameter for the network can be obtained as follows:

$$E = \frac{1}{2}(y^* - y)^2 \tag{7}$$

which, y^* is the predicted value.

4. DEVELOPMENT OF THE NEURO-FUZZY GMDH MODEL USING GSA

In this section, the NF-GMDH model is developed using GSA approach. The basic structure of NF-GMDH is consisted of partial descriptions (neurons). As mentioned in previous section, the grouped parameters in the form of Gaussian variables and weights related to the fuzzy rule are unknown in each partial description (*PD*). GSA algorithm has been applied to optimized grouped-unknown parameters in PDs.

Recently, a novel heuristic search algorithm which is called gravitational search algorithm (GSA) has been proposed motivated by gravitational law and laws of motion. Application of the GSA model indicated efficient performance for different optimization problems such as filter modeling, identification of hydraulic turbine governing system, scour process at group piles under waves, [30, and 31].

In GSA, a set of agents called masses are introduced to find the optimum solution by

simulation of Newtonian laws of gravity and motion [25]. To describe the GSA consider a system with S masses in which the position of the i^{th} mass is defined as follows:

$$X_i = (x_i^1, ..., x_i^d, ..., x_i^n) \quad , i = 1,2,..., s \tag{8}$$

where, x_i^d is the position of i^{th} mass in the d^{th} dimension and n is dimension of the search space. The mass of each agent is calculated after computing current population's fitness as follows:

$$q_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{9}$$

$$M_i(t) = \frac{q_i(t)}{\sum_{j=1}^{s} q_j(t)} \tag{10}$$

where, $M_i(t)$ and $fit_i(t)$ represent the mass and the fitness value of the agent i at t, and, $worst(t)$ and $best(t)$ are defined as follows (for a minimization problem):

$$worst(t) = \max_{j \in \{1,..,s\}} fit_j(t) \tag{11}$$

$$best(t) = \min_{j \in \{1,..,s\}} fit_j(t) \tag{12}$$

To compute the acceleration of an agent, total forces from a set of heavier masses should be considered based on gravity law (Eq. 13). Also, it is followed by calculation of agent acceleration using motion law (Eq. 14).

$$F_i^d(t) = \sum_{j \in kbest,\, j \neq i} rand_j G(t) \frac{M_j(t) M_i(t)}{R_{ij}(t) + \varepsilon} (x_j^d(t) - x_i^d(t)) \tag{13}$$

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} = \sum_{j \in kbest,\, j \neq i} rand_j G(t) \frac{M_j(t)}{R_{ij}(t) + \varepsilon} (x_j^d(t) - x_i^d(t)) \tag{14}$$

Afterward, the next velocity of an agent is calculated as a fraction of its current velocity added to its acceleration (Eq. 15). Then, its position could be calculated using Eq. (16).

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \tag{15}$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{16}$$

where, $rand_i$ and $rand_j$ are two uniform random in the interval [0, 1], ε is a small value, and $R_{ij}(t)$ is the Euclidian distance between two agents i and j that were defined as

$R_{ij}(t) = \left\| X_i(t), X_j(t) \right\|_2$. *kbest* is the set of first K agents with the best fitness value and biggest mass. *kbest* is a function of time, initialized to K_0 at the beginning and decreasing with time. Here, K_0 is set to s (total number of agents) and is decreased linearly to 1. In GSA, the gravitational constant, G will take an initial value of G_0, and it will be reduced by time:

$$G(t) = G(G_0, t) \tag{17}$$

In present study, Eq. (28) used for the gravitational constant:

$$G(t) = G_0 e^{-\alpha \frac{t}{T}} \tag{18}$$

Performing the NF-GMDH and GSA models is a parallel action in each PD. Also, two fuzzy rules were used to model the neuro-fuzzy in each *PD*.

The value of the control parameters of the GSA algorithm such as α, G_0, number of agents, and maximum number are presented in Table 3. In case of free space status, the NF-GMDH-GSA network is modeled with three layers, 45 PDs, and 90 fuzzy rules generated through an optimization process. In addition, for ground status, the proposed NF-GMDH-GSA model has three layers with 30 PDs and 60 fuzzy rules.

Through modeling the NF-GMDH-GSA model for the lateral spread in free face, 45 partial descriptions (PDs) are produced in the first layer. Then, the second layer is generated using 45 PDs from the first layer. This process could be continued until minimum error of training network is obtained. In conclusion, the NF-GMDH-GSA network is modeled with three layers, 45 PDs, and 90 fuzzy rules generated through an optimization process. In addition, for sloping ground, the proposed NF-GMDH-GSA model has three layers with 30 PDs and 60 fuzzy rules. Table 3 demonstrates the values of the GSA properties for predicting the lateral displacements in both conditions.

Table 3: Values of controlling parameters for GSA model

Parameter	Range
Alpha	20
G_0	100
Number of Variables	6
Maximum Iteration	100
Error	0.00001
Number of Agents	50
Weighting Coefficients	0.1-1.5

Furthermore, three of partial descriptions (PDs) generated in the first layer of the proposed NF-GMDH-GSA networks are expressed as follows:

For free face conditions:

$$(D_h)_1^1 = 0.282\exp[-\frac{(M-0.617)^2}{0.918} - \frac{(R-0.617)^2}{0.918}] + 0.32\exp[-\frac{(M-1.42)^2}{0.383} - \frac{(R-1.42)^2}{0.383}] \quad (19)$$

$$(D_h)_2^1 = 0.8188\exp[-\frac{(R-0.3186)^2}{0.2255} - \frac{(W-0.3186)^2}{0.2255}] + 0.3536\exp[-\frac{(R-1.074)^2}{0.27} - \frac{(W-1.074)^2}{0.27}] \quad (20)$$

$$(D_h)_3^1 = 0.3098\exp[-\frac{(W-1.0474)^2}{1.0364} - \frac{(T_{15}-1.0474)^2}{1.0364}] + 0.4731\exp[-\frac{(W-1.157)^2}{1.082} - \frac{(T_{15}-1.157)^2}{1.082}] \quad (21)$$

And, for sloping ground:

$$(D_h)_1^1 = 0.149\exp[-\frac{(R-0.265)^2}{0.968} - \frac{(S-0.265)^2}{0.968}] + 0.7711\exp[-\frac{(R-0.577)^2}{0.816} - \frac{(S-0.577)^2}{0.816}] \quad (22)$$

$$(D_h)_2^1 = 0.5417\exp[-\frac{(S-0.6078)^2}{0.3966} - \frac{(T_{15}-0.6078)^2}{0.3966}] + 0.2644\exp[-\frac{(S-0.596)^2}{0.618} - \frac{(T_{15}-0.596)^2}{0.618}] \quad (23)$$

$$(D_h)_3^1 = 0.562\exp[-\frac{(T_{15}-0.548)^2}{0.548} - \frac{(F_{15}-0.548)^2}{0.548}] + 0.896\exp[-\frac{(T_{15}-0.8573)^2}{0.3399} - \frac{(F_{15}-0.8573)^2}{0.3399}] \quad (24)$$

The superscript and subscript of each parameter present the number of pertaining layer and partial description, respectively.

5. DEVELOPMENT OF THE GENE-EXPRESSION PROGRAMMING

Most recently a new technique called gene-expression programming (GEP) was developed which is an extension of the GP approach. The GEP is a search model that evolves computer programming in the forms of mathematical expressions, decision trees, and logical expressions [32, 33, 34, and 35]. In addition, the GEP model has attracted the attention of investigators in prediction of characterizations in hydraulic problems. This research represents GEP models for evaluation of scour hole geometry at downstream of stilling basins. The GEP approach is coded in the forms of linear chromosomes, which are expressed into Expression Trees (ETs).

In fact, the ETs are sophisticated computer programming which are usually evolved to

solve a practical problem, and are selected accordingly to their fitness at solving that problem. The corresponding empirical expressions can be obtained from these trees structures. A population of the ETs will discover traits, and therefore will adapt to the particular problem they are employed to solve [32, 33, 34, and 35].

Development of the GEP approach includes five steps. The first step is the selection of the fitness function, f_i, of an individual program (i). This item is evaluated as follows:

$$f_i = \sum_{j=1}^{C_t} (M - \left| C_{(i,j)} - T_j \right|) \tag{25}$$

which, M , $C_{(i,j)}$, and T_j are the selection range, value returned by the individual chromosome i for fitness case j, and the largest value for fitness case j.

In second stage, the set of terminals T and the set of function F are selected to generate the chromosomes. In this research, for free space conditions, the terminal includes six independent parameters in the form of $T(D_h) = S, M, R, D50_{15}, F_{15}$, and T_{15} and $T(D_h) = W, M, R, D50_{15}, F_{15}$, and T_{15} also is considered for sloping ground status. To find the appropriate function set, it is necessary to peer review previous investigations made on liquefaction-induced lateral ground displacements in this area. In this way, four basic operators (+, -, *, /) and basic mathematical functions (√, power, sin, cos, exp) are applied to lateral displacements modeling. The third step is to configure the chromosomal architecture. The fourth step is selection of liking function. Finally, for the fifth stage, the sets of genetic operators and their rates are selected. The other details related to the architecture of the GEP modeling are expressed in the literature [32]. In this study, lateral ground displacements are predicted using the GEP model for different conditions of the free space and the sloping ground. The best explicit form of the GEP approach for lateral ground displacements in the free face and the ground sloping conditions are given as follows, respectively:

$$\begin{aligned} D_h = {} & 15.31 - 2.183\frac{D_{5015} + \cos(R)}{\sin(S) + R + F_{15}} + 0.003615 D_{5015} S^3 - 0.9099 D_{5015} \\ & - 0.9099\cos(F_{15}) - 1.82(S + F_{15}) + 0.02087\frac{(S + F_{15} - T_{15})M}{(S + F_{15})} \end{aligned} \tag{26}$$

$$\begin{aligned} D_h = {} & 0.6815 + 0.0001367\frac{(M.T_{15})^2}{D_{5015}^2 F_{15}} + 11.32\frac{(W - R)T_{15}}{(W + M)M^2} \\ & + 0.000778\left(M.T_{15} + \frac{W}{R}\right)\left(T_{15} - M + M^2\right) + 0.0484\left(T_{15}^2 + \frac{F_{15}}{W}\right)(\sin(M) - 1) \end{aligned} \tag{27}$$

Furthermore, the functional set and the operational parameters applied in the proposed GEP models are presented in Table 4.

Table 4: Functional set and the operational parameters used for proposed GEP models

Value	Parameters
Population size	100
Number of generations	500
Tournament size	5
Max. gene	4
Max. tree depth	4
Operators and Functions	(√, +, -, *, /, power, sin, cos, exp)

6. RESULTS

The results of the NF-GMDH-GSA networks and GEP models are presented in this section. In this way, correlation coefficient (*R*), mean absolute percentage of error (MAPE), root mean square error (RMSE), BIAS, and scatter index (*SI*) are defined to evaluate error indicators in the training and validation stages according to the following equations:

$$R = \frac{\sum_{i=1}^{M}\left[Y_{i(Actual)} - \bar{Y}_{(Actual)}\right]\left[Y_{i(model)} - \bar{Y}_{(model)}\right]}{\sqrt{\sum_{i=1}^{M}[Y_{i(Actual)} - \bar{Y}_{(Actual)}]^2 \cdot \sum_{i=1}^{M}[Y_{i(model)} - \bar{Y}_{(model)}]^2}} \tag{28}$$

$$MAPE = \frac{1}{M}\left[\frac{\sum_{i=1}^{M}\left|Y_{i(model)} - Y_{i(Actual)}\right|}{\sum_{i=1}^{M} Y_{i(Actual)}} \times 100\right] \tag{29}$$

$$RMSE = \left\{\frac{\sum_{i=1}^{M}[Y_{i(model)} - Y_{i(Actual)}]^2}{M}\right\}^{1/2} \tag{30}$$

$$Bias = \sum_{i=1}^{M}\left[\frac{Y_{i(model)} - Y_{i(Actual)}}{M}\right] \tag{31}$$

$$SI = \frac{RMSE}{(1/M)\sum_{1}^{M} Y_{i(Actual)}} \tag{32}$$

where $Y_{i(Model)}$ is the predicted values (network output), $Y_{i(Actual)}$ is the observed values (target), and *M* is the total of events.

Numerous runs were performed with various initial settings and the performance of the developed model was analyzed for each run. Consequently, the best model selected according to statistical criteria such as *R*, *MAPE*, *RMSE*, *Bias*, and *SI*. In addition, a comprehensive parametric study was performed to monitor the behavior of each model versus variations of input variables. The Proposed NF-GMDH-GSA model that selected as the most appropriate model, was constituted by six input parameters (*M*, *R*, *W* or *S*, T_{15}, F_{15}, $D50_{15}$) and one output (D_h).

Precision of the proposed model examined by comparison of the actual and NF-GMDH-GSA predicted values of lateral displacements (D_h) for the free face (Figs. 2 and 3) and the sloping ground (Figs. 4 and 5) conditions. For the free face cases, the values of *R*, *MAPE*, *RMSE*, *Bias*, and *SI* are equal to 0.951, 0.483, 0.672, 0.302, and 0.242, respectively, for training set (Fig. 2) and 0.939, 0.587, 0.814, 0.331, and 0.257, respectively, for validation

set (Fig. 3). Also for the sloping ground cases, the values of *R*, *MAPE*, *RMSE*, *Bias*, and *SI* are equal to 0.940, 0.357, 0.409, 0.319, and 0.361, respectively, for training set (Fig. 4) and 0.927, 0.396, 0.469, 0.344, and 0.377, respectively, for validation set (Fig. 5).

The results shown in Figs. 2–5 indicate reasonable good performance of NF-GMDH-GSA model for assessment of the lateral displacement for both free face and sloping ground cases.

Figs. 6 and 7 illustrate the precision of the proposed GEP model of free face (*R*= 0.653, *MAPE*=1.014, *RMSE*=1.483, *Bias*=0.856, *SI*=1.198) and sloping ground (*R*=0.614, *MAPE*=1.110, *RMSE*=1.582, *Bias*=1.268, *SI*=1.657) conditions, respectively. These figures are plotted for all data.

The results indicate that NF-GMDH-GSA model obtains the better performance than the GEP model in prediction of liquefaction induced lateral displacements (D_h). This can be inferred from the performance parameters presented in Table 5, showing the values of statistical parameters of NF-GMDH-GSA and GEP models for training, validation, and all datasets. In fact, the evolved NF-GMDH-GSA model has obtained enough accuracy for both testing and validation sets.

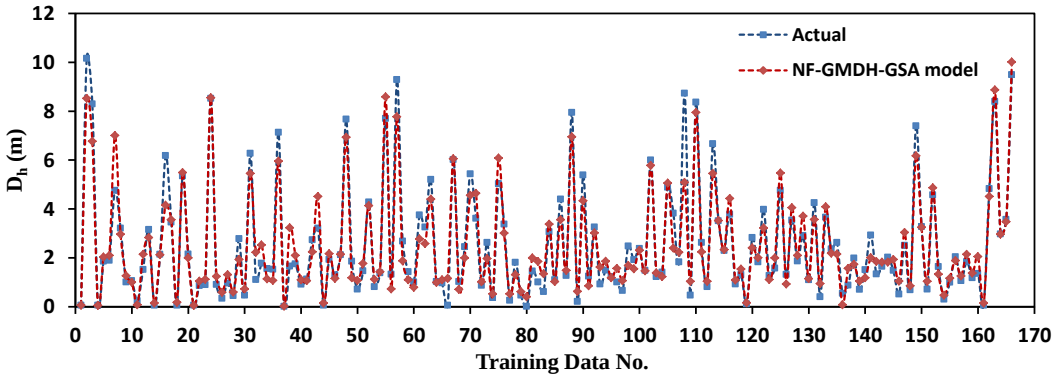


Figure 2. Measured versus NF-GMDH-GSA predicted values of D_h for free face cases - training dataset

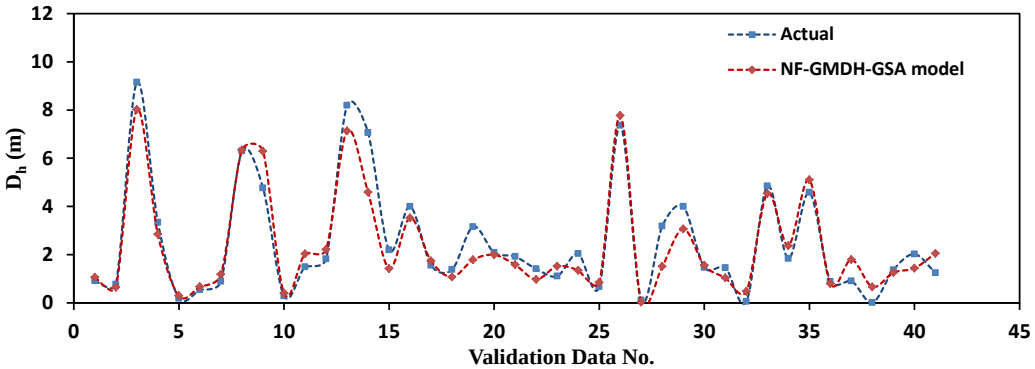


Figure 3. Measured versus NF-GMDH-GSA predicted values of D_h for free face cases - validation dataset

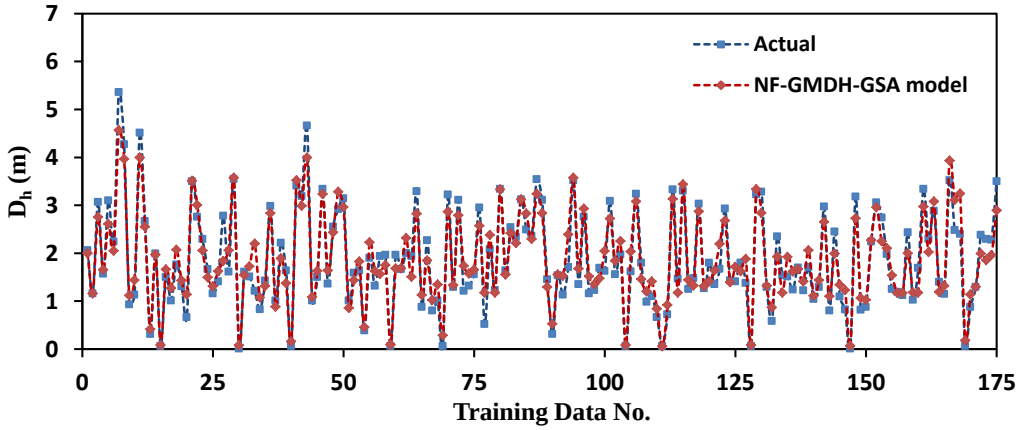


Figure 4. Measured versus NF-GMDH-GSA predicted values of D_h for sloping ground cases - training dataset

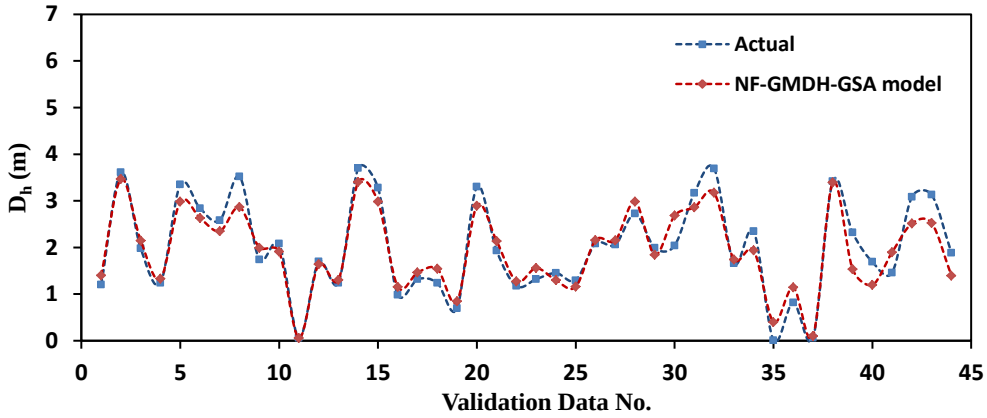


Figure 5. Measured versus NF-GMDH-GSA predicted values of D_h for sloping ground cases - validation dataset

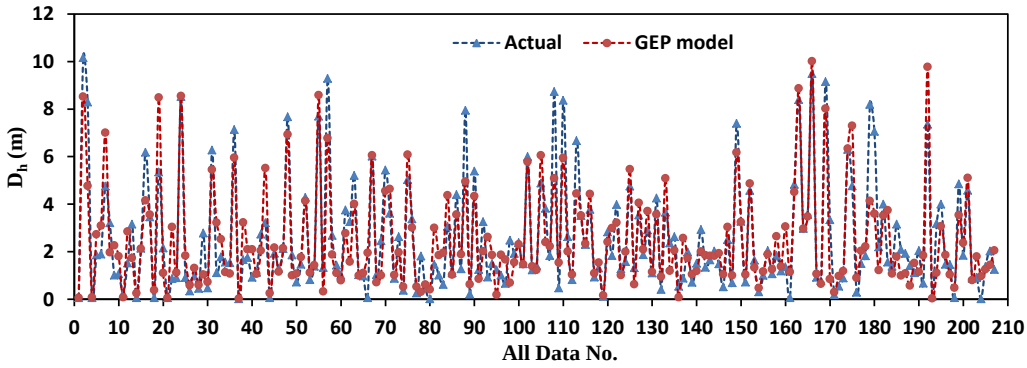


Figure 6. Measured versus GEP predicted values of D_h for free face cases – all data

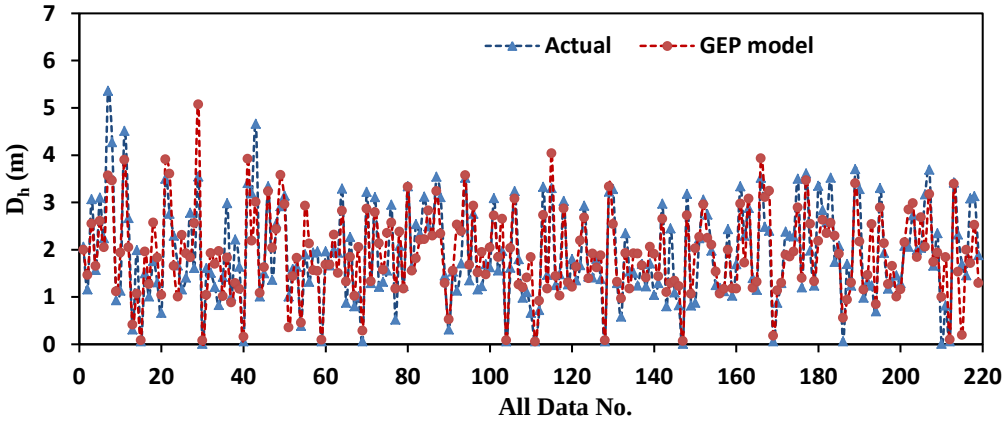


Figure 7. Measured versus GEP predicted values of D_h for sloping ground cases - all data

7. MODEL ACCURACY

Presentation of the model accuracy is relied on the NF-GMDH-GSA model because of high precision of this model in comparison with the GEP model. For all element tests data, values of relative errors for free face and sloping ground conditions are shown in Figs. 8 and 9, respectively. Relative error is the difference between the actual and predicted D_h. It is observed that the developed NF-GMDH-GSA models can predict the lateral displacements (D_h) with reasonable accuracy because the relative error is satisfactorily less than 1m for majority of data in both the free face and the sloping ground conditions.

Table 5: Target error parameters of proposed NF-GMDH-GSA and GEP models

Error parameters	Performance											
	NF-GMDH-GSA based model						GEP based model					
	Free face			Sloping ground			Free face			Sloping ground		
	Training	Validation	All data	Training	Validation	All data	Training	Validation	All data	Training	Validation	All data
R	0.951	0.939	0.944	0.940	0.927	0.931	0.667	0.6379	0.653	0.623	0.601	0.614
MAPE	0.483	0.587	0.533	0.357	0.396	0.374	0.856	1.157	1.014	0.981	1.307	1.11
RMSE	0.672	0.814	0.772	0.409	0.469	0.428	1.325	1.512	1.483	1.424	1.816	1.582
Bias	0.302	0.331	0.312	0.319	0.344	0.330	0.821	0.902	0.856	1.209	1.413	1.268
SI	0.242	0.257	0.249	0.361	0.377	0.368	1.154	1.312	1.198	1.558	1.813	1.657

8. COMPARISON BETWEEN NEW RESULTS AND PREVIOUS STUDIES

Predictions of some previously published empirical relationships for evaluation of liquefaction induced lateral ground displacements are compared with what is found in this study. Table 6 presents the models. Due to the complexities exist in the liquefaction induced lateral ground displacement phenomenon, the aforementioned constitutive models as well as

simplified analytical methods have failed to capture the full effect, and thus empirical models based on case histories have remained as a popular method in the past decades. [8, 11, and 15].

In present research, the empirical correlations and multi-linear regression (MLR) models for estimation of lateral ground displacements (D_h) are introduced. Table 7 presents the values of *R*, *MAPE*, *RMSE*, *Bias*, and *SI* for proposed neuro-fuzzy group method of data handling (NF-GMDH) gravitational search algorithm (GSA) models, and the values estimated by empirical relationships (Table 6) for lateral displacements (D_h) for free face and the sloping ground conditions. The results indicate higher precision of proposed NF-GMDH-GSA model with respect to the previously published empirical equations. However, the developed NF-GMDH model proposed can be an applicable and more reliable tool for predicting liquefaction induced lateral ground displacement.

Table 6: Empirical relationships for estimation of the lateral ground displacement

Condition	Equation	Proposer
Free face	$Log\ D_h = -20.71 + 25.32LogM - 1.39Log(10^{(0.89M-5.64)} + R) - 0.009R + 1.15LogW + 0.19T_{15}0.5 - 0.02F_{15} - 0.84Log(D50_{15} + 0.1mm)$	Kanibir [15]
	$Log\ D_h = -16.713 + 1.532M - 1.406Log\left(10^{(0.89M-5.64)} + R\right) - 0.012R + 0.592LogW + 0.540LogT_{15} + 3.413Log(100 - F_{15}) - 0.795Log(D50_{15} + 0.1mm)$	Youd et al. [8]
	$Log\ (D_h + 0.01) = -17.372 + 1.248M - 0.923LogR - 0.014R + 0.685LogW + 0.3LogT_{15} + 4.826Log(100 - F_{15}) - 1.091D50_{15}$	Bardet et al. [11]
Sloping ground	$Log\ D_h = -7.52 + 8.44LogM + 0.001(10^{(0.89M-5.64)} + R) - 0.23R + 0.11S + 0.6LogT_{15} - 0.22F_{15} - 0.89LogD50_{15}$	Kanibir [15]
	$Log\ D_h = -16.213 + 1.532M - 1.406Log\left(10^{(0.89M-5.64)} + R\right) - 0.012R + 0.338LogS + 0.540LogT_{15} + 3.413Log(100 - F_{15}) - 0.795Log(D50_{15} + 0.1mm)$	Youd et al. [8]
	$Log\ (D_h + 0.01) = -14.152 + 0.988M - 1.049LogR - 0.011R + 0.318LogS + 0.619LogT_{15} + 4.287Log(100 - F_{15}) - 0.705D50_{15}$	Bardet et al. [11]

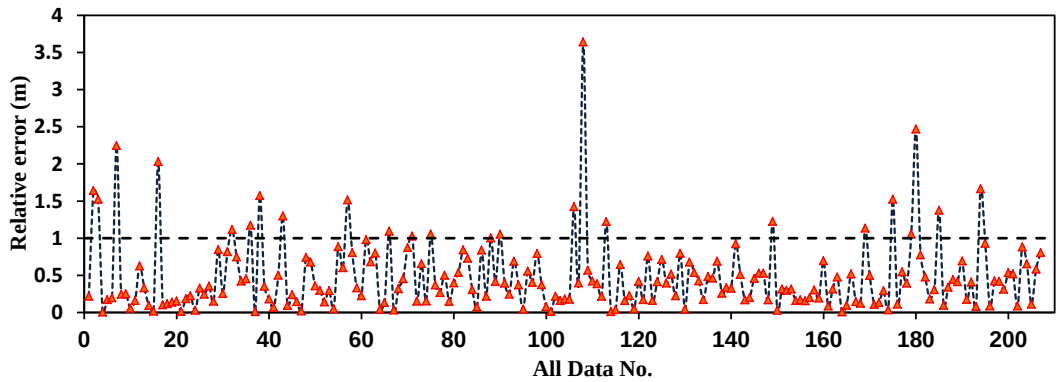


Figure 8. Variation of absolute relative error between the actual and NF-GMDH-GSA predicted

values of D_h – free face cases

Table 7: Comparison between statistical parameters for NF-GMDH model and previous models

Model	Performance									
	Free face					Sloping ground				
	R	*MAPE*	*RMSE*	*Bias*	*SI*	*R*	*MAPE*	*RMSE*	*Bias*	*SI*
NF-GMDH-GSA based model (present study)	0.944	0.533	0.772	0.312	0.249	0.931	0.374	0.428	0.330	0.368
Kanibir [15]	0.825	0.600	0.654	0.517	0.331	0.792	0.514	0.612	0.682	0.511
Youd et al. [8]	0.889	2.335	3.061	0.602	0.296	0.719	1.728	1.915	0.745	0.630
Bardet et al. [11]	0.835	0.832	1.257	0.532	0.327	0.758	0.509	0.667	0.739	0.548

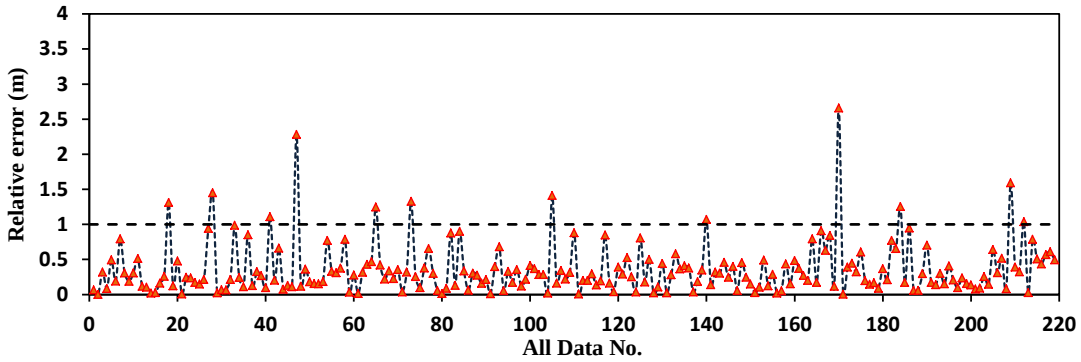


Figure 9. Variation of absolute relative error between actual and NF-GMDH-GSA predicted values of D_h – sloping ground cases

9. CONCLUSIONS

Determination of liquefaction induced lateral ground spreading is a complex geotechnical engineering problem. A robust neuro-fuzzy group method of data handling (NF-GMDH) based on gravitational search algorithm (GSA) and gene-expression programming (GEP) models are developed for prediction of liquefaction induced lateral ground displacement using a large data. A wide-range database of case histories consisting of 426 data of liquefaction-induced lateral displacement for the free face and the sloping ground conditions obtained from ten earthquakes, and were compiled and analyzed. Based on data analysis and the previous researches, the following results are made:

1. the most important factors that affect the lateral ground displacement during liquefaction are categorized as seismological (*M*, *R*), topographical (*S*, *W*), and geotechnical (T_{15}, F_{15}, $D50_{15}$) parameters.

2. It was shown that for both free face and sloping ground cases, the NF-GMDH-GSA models are able to learn with a very high accuracy,

A. For free face conditions, *R*= 0.944, *MAPE*=0.533, *RMSE*=0.772, *Bias*=0.312, *SI*=0.249, and displacements were from 0.01 to 10.16 m.

B. For sloping ground conditions, R= 0.931, MAPE=0.374, RMSE=0.428, Bias=0.330, SI=0.368, and displacements were from 0.01 to 3.36 m.

3. It is clearly observed that the NF-GMDH-GSA model yields a much better performance than GEP models and the previous equations.

4. The results of present study confirm higher precision of proposed model in comparison with previous models. This accuracy shows the superiority of the proposed NF-GMDH-GSA models over available relationships and models in literature.

5. It is clear that a precise correlation is easier to be used in routine geotechnical projects compared with the field measurement techniques.

REFERENCES

1. Javadi AA, Rezania M, Mousavi Nezhad M. Evaluation of liquefaction induced lateral displacements using genetic programming, *Comput Geotech* 2006; **33**(4-5): 222-33.
2. Newmark NM. Effects of earthquakes on dams and embankments, *J Geotech* 1965; **5**(2): 137-60.
3. Shi LP, Towhata I, Wieland M. Prediction of seismically induce deformation of Liyutan Dam, Taiwan, by means of cyclic triaxial testing and finite element analysis, *J Comput Geotech* 1989; **7**(3): 205-20.
4. Baziar MH, Dobry R. Residual strength and large-deformation potential of loose silty sand, *J Geotech Geoenviron Eng, ASCE* 1995; **121**(2): 896-906.
5. Finn WDL. Assessment of liquefaction potential and post-liquefaction behavior of earth structures: developments 1981–1991, *Proceedings of Second International Conference on Recent Advances in Geotechnical Engineering and Soil Dynamics* 1991; **II**: pp. 1833-1850.
6. Bartlet SF, Youd TL. Empirical analysis of horizontal ground displacement generated by liquefaction-induced lateral spreads. Technical Report No. NCEER-91- 021, National Center for Earthquake Engineering Research, State University of New York at Buffalo, NY, 1992.
7. Juang CH, Chen CJ. CPT-based liquefaction evaluation using artificial neural networks, *Computer-Aided Civil Infrastruct Eng* 1999; **14**: 221-9.
8. Youd TL, Hansen CM, Bartlett SF. Revised multi-linear regression equations for prediction of lateral spread displacement, *J Geotech Geo-environ Eng ASCE* 2002, **128**(12): 1007-17.
9. Hamada M, Yasuda R, Isoyama S. Study on liquefaction induced permanent ground displacement. Report for the Association for the Development of Earthquake Prediction, Japan, 1986.
10. Youd TL, Perkins DM. Mapping of liquefaction induced ground failure potential, *J Geotech Eng Div, ASCE* 1987, **104**: 433-46.
11. Bardet JP, Mace N, Tobita T. Liquefaction-induced ground deformation and failure, a report to PEER/PG&E. Task 4A - Phase 1, Civil Engineering Department, University of Southern California, Los Angeles, 1999.
12. Mohebi B, Ghodrati Amiri Gh, Taheri M. Selection of suitable records for nonlinear analysis using genetic algorithm (GA) and particle swarm optimization (PSO), *Int J Optim Civil Eng* 2014; **4**(4): 509-24.

13. Fattahi H. Prediction of slope stability state for circular failure: a hybrid support vector machine with harmony search algorithm, *Int J Optim Civil Eng* 2015; **5**(1): 103-15.
14. Kaveh A, Zakian P. Performance based optimal seismic design of rc shear walls incorporating soil–structure interaction using CSS algorithm, *Int J Optim Civil Eng* 2012; **2**(3): 383-405.
15. Kanibir A. Investigation of the lateral spreading at sapanca and suggestion of empirical relationships for predicting lateral spreading, M.Sc. Thesis, Department of Geological Engineering, Hacettepe University, Ankara, Turkey [in Turkish], 2003.
16. Zhang G, Robertson KP, Brachman IWR. Estimating liquefaction-induced lateral displacements using the standard penetration test or cone penetration test, *J Geotech Geoenviron Eng, ASCE* 2004; **130**(8): 861-71.
17. Idriss IM, Boulanger RW. Soil liquefaction during earthquakes, In EERI Monograph, Earthquake Engineering Research Institute, California, 2008, pp. 262.
18. Kalantary F, MolaAbasi H, Salahi M, Veiskarami M. Prediction of liquefaction induced lateral displacements using robust optimization model, *Scientia Iranica* 2013; **20**(2): 242-50.
19. Penumadu D, Zhao R. Triaxial compression behavior of sand and gravel using artificial neural networks (ANN), *J Comput Geotechn* 1999; **24**: 207-30.
20. McCombie P, Wilkinson P. The use of the simple genetic algorithm in finding the critical factor of safety in slope stability analysis, *Comput Geotech* 2002; **29**: 699-714.
21. Shahin MA, Maier HR, Jaksa MB. Predicting settlement of shallow foundations using neural networks, *J Geotech Geoenviron Eng, ASCE* 2002; **128**(9): 785-93.
22. Baziar MH, Jafarian Y, Shahnazari H, Movahed V, Tutunchian MA. Prediction of strain energy-based liquefaction resistance of sand–silt mixtures: An evolutionary approach, *Comput Geosci* 2011; **37**: 1883-93.
23. Kalantary F, Ardalan H, Nariman-Zadeh N. An investigation on the S_u-N_{SPT} correlation using GMDH type neural networks and genetic algorithms, *Eng Geo* 2009; **104**: 144-55.
24. Nagasaka K, Ichihashi H, Leonard R. Neuro-fuzzy GMDH and Its application to modeling grinding characteristics, *Int J Prod Res* 1995; **33**(5): 1229-40.
25. Baziar MH, Ghorbani A. Evaluation of lateral spreading using artificial neural networks, *Soil Dyn Earthquake Eng* 2005; **25**: 1-9.
26. Rezania M, Faramarzi A, Javadi AA. An evolutionary based approach for assessment of earthquake-induced soil liquefaction and lateral displacement, *Eng Appl Artif Intel* 2011; **24**(1): 142-53.
27. Masters T. *Practical Neural Network Recipes In C++*, San Diego, Academic press, 1993.
28. Ivakhnenko AG. The Group Method of Data Handling In Prediction Problems, *Soviet Automat Control Avtomat* 1976; **9**(6): 21-30.
29. Takashi O, Hidetomo I, Tetsuya M, Kazunori N. Orthogonal and successive projection methods for the learning of neurofuzzy GMDH, *Inform Sci* 1998; **110**: 5-24.
30. Rashedi E, Nezamabadi-pour H, Saryazdi S. Filter modeling using gravitational search algorithm, *Eng Appl Artif Intel* 2011; **24**(1): 117-22.
31. Li C, Zhou J. Parameters identification of hydraulic turbine governing system using improved gravitational search algorithm, *Energy Conversion and Management* 2011; **52**: 374-81.

32. Azamathulla HM, Haque AAM. Prediction of scour depth at culvert outlets using gene-expression programming, *Int J Innov Comput, Inform Control* 2012; **8**(7B): 5045-54.
33. Ferreria C. *Gene-Expression Programming Mathematical Modeling by an Artificial Intelligence*, Spriner, Berling, Heidelberg, New York, 2001a, 21.
34. Ferreria C. Gene-expression programming in problem solving The *6th Online Word Conference on Soft Computing in Industrial Applications* (Invited tutorial), 2001b.
35. Ferreria C. *Gene-expression Programming a New Adaptive Algorithm for Solving Problems* (Complex Systems), 2001c; **13**(2): 87-129.