EVALUATION OF LIQUEFACTION SUSCEPTIBILITY OF SOIL USING GENETIC PROGRAMMING AND MULTIVARIATE ADAPTIVE REGRESSION SPLINE

A Thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF TECHNOLOGY

IN CIVIL ENGINEERING

(GEOTECHNICAL ENGINEERING)

RUPASHREE SAHOO

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA,

2012-2014

EVALUATION OF LIQUEFACTION SUSCEPTIBILITY OF SOIL USING GENETIC PROGRAMMING AND MULTIVARIATE ADAPTIVE REGRESSION SPLINE

A Thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF TECHNOLOGY

IN CIVIL ENGINEERING

(GEOTECHNICAL ENGINEERING)

Under the guidance and supervision of

Dr. Sarat Kumar Das

Submitted by:

RUPASHREE SAHOO

(ROLL NO-212CE1028)

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA,

2012-2014

Dedicated to my Parents and my family members, who had been a constant inspiration for me.

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

CERTIFICATE

This is to certify that the thesis entitled "Evaluation of Liquefaction susceptibility of soil By using Genetic Programming and Multivariate Adaptive regression Spline" being submitted by **Rupashree Sahoo** in partial fulfillment of the requirements for the award of **Master of Technology** in **Civil Engineering** with specialization in **Geotechnical Engineering** at National Institute of Technology Rourkela, is an authentic work carried out by her under my guidance and supervision.

To the best of my knowledge, the contents of this thesis in part or full, have not been submitted to any other university or Institution for the award of any degree or diploma.

PLACE: NIT ROURKELA

DATE: 26/05/2014

DR. SARAT KUMAR DAS

ASSOCIATE PROFESSOR

DEPARTMENT OF CIVIL ENGINEERING

NIT ROURKELA

ACKNOWLEDGEMENT

It would not have been possible to complete the thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

First and foremost, I would like to express my gratitude and sincere thanks to my esteemed supervisor **Prof. Sarat Kumar Das** for his consistent guidance, valuable suggestions and encouragement throughout the work and in preparing this thesis.

His inspiring words always motivated me to do hard labour which helped me to complete my work in time.

I am grateful to the Dept. of Civil Engineering, NIT Rourkela, for giving me the opportunity to execute this project, which is an important part of the curriculum in M.Tech programme.

I also thank my friends who have directly or indirectly helped in my project work. Many special thanks to my senior **Swagatika Senapati** and my friend **Partha Sarathi Parhi** for their help & co-operation with me in my work.

Last but not the least I would like to thank my family for providing me this platform for study and their support as and when required.

RUPASHREE SAHOO

212CE1028

ABSTRACT

Now a day the human life and the environment have frequently been endangered by the natural hazards like earthquake, tsunami, flood, cyclone and landslides. As a consequence of which the human society and the nation's economy get hampered immediately after the occurrence of a natural disaster. In developing countries like India, where the population is very large and is increasing day by day, the social and economic factors force the people to live in vulnerable areas, due to which the effects of these natural disasters are catastrophic. Among all these threats, liquefaction of soil can be pointed out as one of the most disastrous seismic hazards. Hence evaluation of liquefaction susceptibility is an important aspect of geotechnical engineering. The widely used procedures for evaluation of liquefaction potential of soil are the simplified procedure. This procedure was developed from empirical evaluation of field observations and field and laboratory test data. For evaluation of liquefaction potential of soil generally two variables are required, such as: (i) the seismic demand on a soil layer expressed in terms of CSR, (ii) the capacity of the soil to resist liquefaction expressed in terms of CRR. The method for evaluation of CRR is to test undisturbed soil specimens in the laboratory. To avoid the difficulties associated with sampling and laboratory testing, field tests have become the state-of–exercise for routine liquefaction inquiries. The various field tests used for the liquefaction resistance of the soil are (i) Standard Penetration Test(SPT), (ii) Cone Penetration Test (CPT) , (iii) Shear Wave velocity Measurements and (iv) Becker Penetration test(BPT).

Artificial intelligent techniques such as artificial neural network (ANN), support vector machine (SVM), relevance vector machine (RVM) are used to develop liquefaction prediction models based on in-situ database , which are found to be more efficient as 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 a 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 a comprehensive relationship between the inputs and output, and are also called as ‘black box’ system.

In the present study an attempt has been made to predict the liquefaction potential of soil based post liquefaction cone penetration test (CPT) , standard penetration test (SPT) and shear wave velocity (V_s) data using multivariate adaptive regression splines (MARS) and genetic programming (GP). A comparative analysis is made among the existing methods and the proposed MARS and GP model for prediction of liquefied and non-liquefied cases in terms of percentage success rate with respect to the field manifestations. It is observed that the prediction as per MARS and GP model is more accurate towards field manifestation in comparison to other existing methods.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Now a days the human life and the environment have frequently been threatened by the natural calamities like earthquake, tsunami, flood, cyclone and landslides. As a consequence of which the human society and the nation's economy get hampered immediately after the occurrence of a natural disaster. In developing countries like India, where the population is very large and is increasing day by day, the effects of these natural disasters are catastrophic and the social and economic factors force the people to live in vulnerable areas. Among all these threats, liquefaction of soil can be pointed out as one of the most disastrous seismic hazards. During the last century, it has been estimated that these seismic hazard accounts around 30% of total casualties and 60% of the total property loss (Das and Muduli 2011). Though, soil liquefaction phenomena have been recognized since long, it was more comprehensively brought to the attention of engineers, seismologists and scientific community of the world by several devastating earthquakes around the world; Niigata and Alaska (1964), Loma Prieta (1989), Kobe (1995),Kocaeli (1999) and Chi-Chi (1999) earthquakes (Baziar and Jafarian 2007). Since then, a numerous investigations on field and laboratory revealed that soil liquefaction may be better described as a disastrous failure phenomenon in which saturated soil losses strength due to increase in pore water pressure and reduction in effective stress under rapid loading and the failed soil acquires a degree of mobility sufficient to permit movement from meters to kilometres. Soil liquefaction can cause ground failure in the way of sand boils, major landslides, surface settlement, lateral spreading, lateral movement of bridge supports, settling and tilting of buildings, failure of water front structure and severe damage to the lifeline systems etc.

The liquefaction hazard evaluation involves liquefaction susceptibility analysis, liquefaction potential evaluation, assessment of effect of liquefaction (i.e., the extent of ground failure caused by liquefaction) and study of response of various foundations in liquefied soil. These are the major concern of geotechnical engineers. In the present study, the focus is on liquefaction potential evaluation, which determines the likelihood of liquefaction triggering in a particular soil in a given earthquake. Evaluation of the liquefaction potential of a soil subjected to a given seismic loading is an important first step towards mitigating liquefaction-induced damage. Though, different approaches like cyclic strain-based, energy based and cyclic stress-based approaches are in use, the stress based approach is the most widely used method for evaluation of liquefaction potential of soil (Krammer, 1996). Thus, the focus in present study is on the evaluation of liquefaction potential on the basis of the cyclic stress-based approach. There are two types of cyclic stress based-approach available for assessing liquefaction potential. One is by means of laboratory testing (e.g., cyclic tri-axial test and cyclic simple shear test) of undisturbed samples, and the other involves use of empirical relationships that relate observed field behaviour with in-situ tests such as standard penetration test (SPT), cone penetration test (CPT), shear wave velocity measurement (Vs) and the Becker penetration test (BPT).

The widely used procedures for evaluation of liquefaction resistance are simplified procedure. This simplified procedure was originally developed by Seed and Idriss (1971). Various in-situ tests are also there for evaluation of liquefaction potential of soil such as Standard Penetration Test(SPT) which were developed by Seed and Idriss (1971), Tokimatsu andYoshimi (1983), Seed et al. (1985), Berrill and Davis (1985),and Law et al. (1990), Cone Penetration test(CPT) developed by Robertson and Campanella (1985), Seed and De Alba(1986), Shibata and Teparaksa (1988), and Stark and Olson(1995).Other in-situ test methods for evaluation of liquefaction potential of soil are Dilatometer test (Marchetti 1982) and Shear wave velocity

Test (Andrus and Stokoe,1997). However, there are several limitations in using these methods to determine the liquefaction resistance of saturated sandy soils. Because of the difficulty and the cost constraint of obtaining high-quality undisturbed samples of saturated sandy soils, there is a need for simple, economic procedures for evaluation of earthquake induced liquefaction resistance of these soils.

With the rapid increases in processing speed and memory of low-cost computers, it is not surprising that various advanced computational learning tools such as neural networks have been increasingly used for analyzing or modeling highly nonlinear multivariate engineering problems. These algorithms are useful for analyzing many geotechnical problems, particularly those that lack a precise analytical theory or understanding of the phenomena involved. To take care of such problems Soft computing have been developed. There are various methods of soft computing which are shown in the Figure 1.1

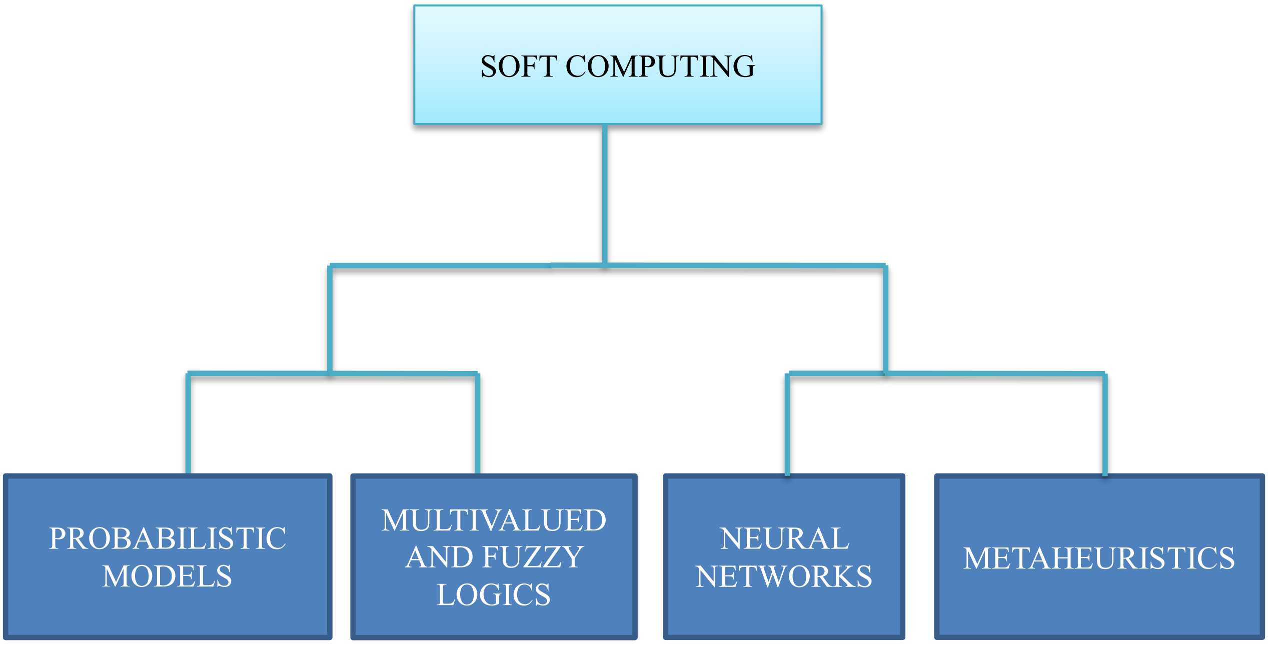

Figure 1.1 Methods of Soft Computing

In situations where measured or numerical data are available, neural networks have been shown to offer great promise for mapping the nonlinear interactions (dependency) between the system's inputs and outputs. Unlike most computational tools, in neural networks no predefined mathematical relationship between the dependent and independent variables is required. Although this is successful in many regards, a major disadvantage, compared to other statistical models is that they provide no information about the relative importance of the various parameters involved, as also implied by some previous studies. It has also been noted that as the knowledge acquired during training is stored in an implicit manner in the ANN, it is very difficult to come up with a reasonable interpretation of the overall structure of the network. These inherent limitations wherein the information or the intervening steps are not available have earned ANN, the reputation of being a "black box" approach. In addition, ANN has several inherent drawbacks such as over fitting, slow convergence speed, poor generalizing performance, and arriving at local minimum. Recently support vector machine (SVM), based on statistical learning theory and structural risk minimization is being used as an alternate prediction model. The SVM developed by Vapnik (1998) , uses structural constrained minimization penalizing the error margin during training. The error function being a convex function better generalization used to observe in SVM as compared to ANN. However, neural networks have been criticized for its long training process since the optimal configuration is not self-evident. Another technique, called the Genetic Programming (GP), developed by Koza, in 1992, mimics biological evolution of living organisms and makes use of principle of genetic algorithm (GA). It is also called as 'grey box' model. Various attempts have been made in the recent past to use GP to some Geotechnical engineering problems. GP helps in achieving greatly simplified model formula compared to ANN model, but a trade-off is made between the complexity of the formula and accuracy of the model.

Another class of model may be termed as ‘white box’ model is the multivariate adaptive regression spline (MARS) developed based on statistical model developed by Friedman (1991). MARS is a fairly simple nonparametric regression algorithm known as multivariate adaptive regression splines, which has the ability to approximate the relationship between the inputs and outputs, and express the relationship mathematically. The main advantages of MARS are its capacity to produce simple, easy-to-interpret models, its ability to estimate the contributions of the input variables, and its computational efficiency. MARS can adjust any functional form, hence suitable for exploratory data analysis. Samui et al. (2011) observed that the MARS model for uplift capacity of suction caisson has better statistical performance comparable to ANN and FEM model. It may be mentioned here that, though above AI techniques are based on sound mathematical/numerical background, its application to different problems is an art.

In the present study liquefaction index is predicted by using Multivariate Adaptive Regression splines (MARS) and Genetic Programming (GP) based on post Liquefaction CPT, SPT and Shear wave velocity data. Then a comparison is made between the developed MARS and GP model and the existing methods.

1.2 OBJECTIVE AND SCOPE

The objective of the project work is to develop liquefaction susceptibility analysis of soils using AI techniques; GP and MARS.

SCOPE:

- Evaluation of liquefaction potential of soil from CPT data by using MARS and GP.
- Evaluation of liquefaction potential of soil from SPT data by using MARS and GP.
- Evaluation of liquefaction potential of soil from Shear Wave velocity by using MARS and GP.

1.3 THESIS OUTLINE

This thesis consists of seven chapters and the chapters has been organised in the following order.

After a brief introduction in Chapter 1 the Literature review and the methodology are described in the Chapter 3 and 4 respectively.

Chapter 4,5, 6, describes the application of MARS and GP for evaluation of Liquefaction potential of soil from post liquefaction CPT, SPT and shear wave velocity data. A comparison is made between the existing methods and the developed MARS and GP model.

In Chapter 7 conclusions drawn from above chapters and scope for the future studies are presented. The general layout of the thesis work based on each chapter is presented in a flow diagram as shown below.

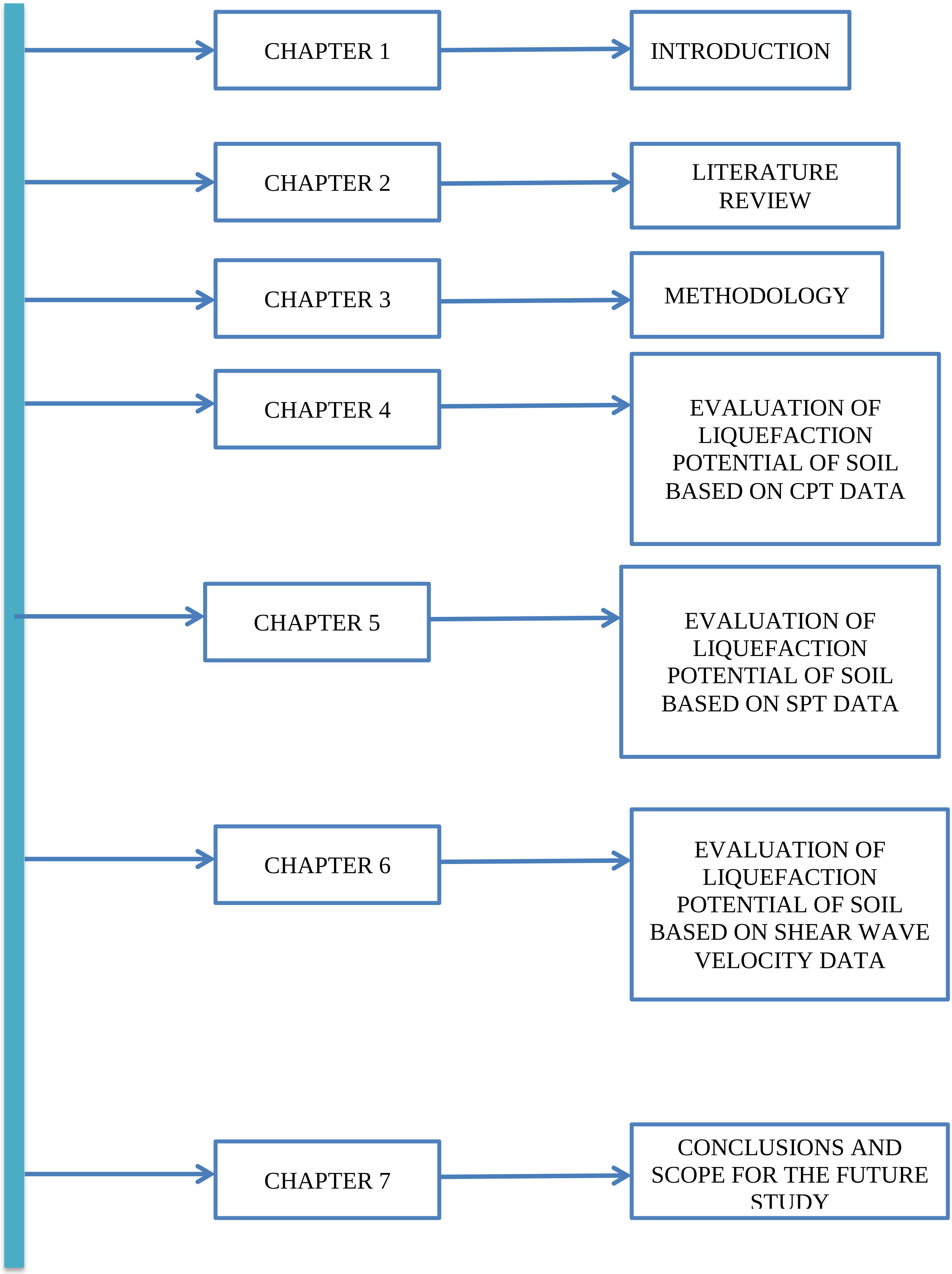

Figure 1.2 Basic Outline of the thesis

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

Liquefaction hazard evaluation involves liquefaction susceptibility analysis, liquefaction potential evaluation, assessment of effect of liquefaction and study of response of various foundations in liquefied soil. These are the major concerns of Geotechnical engineers. In the present study the focus is on liquefaction potential evaluation for determination of the likelihood of liquefaction triggering in a particular soil in a given earthquake. A review of the various liquefaction potential evaluation methods are presented in this chapter. In the present section, application of various methods for liquefaction susceptibility analysis is only discussed in International and natural scenario.

2.2 LITERATURE REVIEW

Once a particular soil is found to be susceptible to liquefaction on the basis of various susceptibility criteria as mentioned in Kramer (1996) the next step in the liquefaction hazard evaluation process is the evaluation of liquefaction potential, which is the main topic of the present study. The major factors controlling the liquefaction potential of a saturated cohesion less soil in level ground are the intensity and duration of earthquake shaking and the density and effective confining pressure of the soil. Several approaches are used for evaluating liquefaction potential, including (i) the cyclic stress-based approach, (ii) the cyclic strain-based approach, and (iii) the energy-based approach.

2.2.1 INTERNATIONAL STATUS

Ishihara (1993) suggested that in case of liquefaction resistance evaluated by using CPT value for silty sands (> 5%fines), the effects of fines could be estimated by adding some tip resistance increments to the measured tip resistance to obtain an equivalent clean sand tip resistance.

Goh (1994), studied the feasibility of using neural networks to model the complex relationship between the seismic and soil parameters, and the liquefaction potential. He used a simple back propagation neural-network algorithm. He concluded that, the performance of the neural-network models improved as more input variables are provided. The model consisting of eight input variables was the most successful. Out of which the standard penetration test (SPT) value and the fines content were the most important variables. Comparisons indicate that the neural-network model is more reliable than the conventional dynamic stress method by Seed et al.

Goh (1996), examined the feasibility of neural-network to assess liquefaction potential from the actual CPT field test data. In comparison to other conventional deterministic methods, the proposed neural-network method precludes a probabilistic assessment of the risk of liquefaction.

Juang et al. (2002), investigated the data's collected from the largest earthquake of the century in Taiwan, They examined the Three CPT- based simplified methods, the Olsen method, the Robertson method and the Juang method, using the case histories derived from the Chi-Chi earthquake. As per them, the comparison shows that the Juang method is more accurate than the other two methods in predicting the liquefaction potential of soils based on the cases derived from the Chi-Chi earthquake.

Juang et al. (2003) also developed an ANN-based simplified method using soil type index (*Ic*) for evaluation of *CRR* of soil using post liquefaction CPT database and also used Bayesian mapping function approach to relate *Fs* with *PL*.

Moss (2003) and Moss et al. (2005) presented a CPT-based probabilistic model for evaluation of liquefaction potential using reliability approach and a Bayesian updating technique.

Juang et al. (2006) used first order reliability method (FORM) for probabilistic assessment of soil liquefaction.

Baziar and Jafarian (2007), developed an artificial neural network (ANN) model to establish a correlation between soils initial parameters and the strain energy required to trigger liquefaction in sands and silty sands. A relatively large database, presenting laboratory cyclic data of clean and silty sands, and also data of several centrifuge liquefaction tests were utilized to develop an ANN model to predict the amount of strain energy required up to liquefaction triggering. A subsequent parametric study was carried out and the trends of the results have been confirmed via some previous laboratory studies. In addition, the data recorded during some real earthquakes at Wildlife, Lotung and Port Island Kobe sites plus some available centrifuge tests data have been utilized in order to validate the proposed ANN-based liquefaction energy model. The results clearly demonstrate the capability of the proposed model and the strain energy concept to assess liquefaction resistance (capacity energy) of soils.

2.2.2 NATIONAL STATUS

Das and Samui (2007), proposed and investigated with the use of the Relevance Vector Machine (RVM) to determine the liquefaction potential of soil by using actual cone penetration test (CPT) data. RVM is based on a Bayesian formulation of a linear model with an appropriate prior that results in a sparse representation. He compared the results with widely used artificial neural network (ANN) model. Overall, the RVM shows good performance and is proven to be more accurate than the ANN model. It also provides probabilistic output. The model provides a viable tool for earthquake engineers to assess seismic conditions for sites that are susceptible to liquefaction.

Pal (2006) examined the potential of Support Vector Machines(SVMs) for accessing the liquefaction potential from the actual SPT and CPT field data. The data were taken from

Goh (1994) and Goh (1996).SPT field data consists of a total of 85 records out of which there were 42 liquefied sites and the remaining 43 are non-liquefied sites after earthquakes. CPT field data sets consists of a total of 109 data sets out of which 74 are liquefied sites and 35 are non-liquefied sites. The advantage of SVMs approach is the use of SPT(N_1)value which was changed to SPT$(N_1)_{60}$ value in Goh (1994).SVMs method showed better result than Neural network method and SVMs requires small training time. The number of data used to provide the results is smaller than the neural network approach.

Samui(2011), proposed Least Square Support Vector Machine(LSSVM) and Relevance Vector Machine(RVM) for evaluation of seismic liquefaction potential of soil from actual Standard Penetration Test(SPT) data. He collected datas from Goh (1994). He proposed that LSSVM and RVM method suggest the standardized SPT $\{(N_1)_{60}\}$ value is not required for the determination of liquefaction potential of soil. A comparison were made between the two developed model and ANN model developed by Goh(1994).The comparison indicated that the RVM model is more reliable than the ANN and LSSVM model.

Das and Muduli, (2011), analysed the liquefaction potential of soil based on cone penetration test (CPT) data obtained after 1999 Chi-Chi, Taiwan, earthquake using genetic programming (GP), and made a comparative study among the three CPT based statistical methods, i.e. ANN, SVM and GP, for prediction of liquefied and non-liquefied cases in terms of liquefaction index, and found that the developed GP model is more efficient as compared to the other two statistical methods.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

In the present study, two artificial intelligence techniques, GP and MARS have been used to to separate liquefaction and non-liquefaction cases in terms of liquefaction index(LI). As Genetic programming and its variant MGGP and MARS have been used in very limited geotechnical engineering problems, and are not very common to geotechnical engineering professionals, hence are discussed in brief as follows.

3.2 MULTIVARIATE ADAPTIVE REGRESSION SPLINES (MARS)

MARS is basically a nonparametric regression procedure that does not assume any functional relationship between independent and dependent variables. Instead, MARS uses the regression data to construct this relation and forms, some sets of coefficients and basis functions. In other word, it can be said that this method is based on “divide and conquer” strategy, which divides the input parameters into groups or say regions, each having its own regression equation. So this makes MARS particularly suitable for problems with higher input dimensions (i.e., with more number of variables), whereas other techniques face problem of dimensionality with large number of input variables. MARS is an adaptive procedure because the selection of basis functions is data-based and specific to the problem at hand. It is very useful for high dimensional problems and shows a great promise for fitting nonlinear multivariate functions. A special advantage of MARS lies in its ability to estimate the contributions of the basis functions so that both the additive and the interactive effects of the predictors are allowed to determine the response variable. For this model an algorithm was proposed by Friedman (1991) as a flexible approach to high dimensional nonparametric regression, based on a modified recursive partitioning methodology. The general form of a MARS predictor is as follows:

$$F(x) = \beta + \sum_{i=1}^{n} B_i a_i \qquad (3.1)$$

Where

β=Intercept, B_i= Basis functions,a_i=Coefficient of Basis function I, n= no of basis functions.

The MARS algorithm for estimating the model function F(x) consists of two algorithms Friedman (1991)

(i) The forward stepwise algorithm: Here, forward stepwise search for the basis function takes place with the constant basis function, the only one present initially. At each step, the split that minimized some “lack of fit” criterion from all the possible splits on each basis function is chosen. The process stops when a user-specified value max M is reached. At the end of this process, there will be a large expression. This model typically over fits the data and so a backward deletion procedure is applied.

(ii) The backward stepwise algorithm: The purpose of this algorithm is to prevent from over-fitting by decreasing the complexity of the model without degrading the fit to the data.

Therefore, the backward stepwise algorithm involves removing from the model basis functions that contribute to the smallest increase in the residual squared error at each stage, producing an optimally estimated model f^{α} with respect to each number of terms, called α. We note that α expresses some complexity of our estimation. To estimate the optimal value of α, generalized cross-validation can be used which shows the lack of fit when using MARS.

Here in the present study ‘EARTH’ package of R to predict the model of some geotechnical problems. R is a system for statistical computation and graphics. Nowadays it is used in various statistical problems related to engineering, medical, economics etc. Moreover it can also be used for regression problems such as linear, nonlinear, and single or multivariate. The advantage of using R is that, it is very easy to work on R. We don’t have to write long syntax, each and every function of R consists of small syntax. Also data from excel can be directly entered into R from clipboard.

3.3 GENETIC PROGRAMMING

Genetic programming (GP) is a pattern recognition technique where the model is developed on the basis of adaptive learning over a number of cases of provided data, developed by[8]. It mimics biological evolution of living organisms and makes use of the 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 solution in the form of tree structure or in the form of compact equation using the given dataset. A brief description about GP is presented for the completeness, but the details can be found in [9].

Genetic programming model is composed of nodes, which resembles a tree structure and, thus, it is also known as a 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 (.), tanh (.) or ln(.)), Boolean operators (AND, OR, NOT, *etc.*), logical expressions (IF, or THEN) or any other suitable functions defined by the user. The terminal set includes variables (like x_1, x_2, x_3, *etc.*) or constants (like 3, 5, 6, 9, *etc.*) or both. The functions and terminals are randomly chosen to form a GP tree with a root node and the branches extending from each function nodes to end in terminal nodes.

Initially, a set of GP trees, as per user defined population size, are randomly generated using various functions and terminals assigned by the user. The fitness criterion is calculated by the objective function and it determines the quality of the each individual in the population competing with rest. At each generation a new population is created by selecting individuals as per the merit of their fitness from the initial population and then implementing various evolutionary mechanisms like reproduction, crossover and mutation to the functions and terminals of the selected GP trees. 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 best GP model, based on its fitness value that appeared in any generation, is selected as the result of genetic programming. A brief description of various evolutionary mechanisms in GP is presented below.

Initial population:- In the first step of genetic programming, a number of GP trees are generated by randomly selecting user defined functions and terminals. These GP trees form initial population.

Reproduction:- . In the second stage of the GP, a proportion of the initial population is selected and copied to the next generation and this procedure is called reproduction. Roulette wheel selection, tournament selection, ranking selection, *etc*., are the methods generally followed.

Crossover:- In crossover operation, two trees are selected randomly from the population in the mating pool. One node from each tree is selected randomly, the sub-trees under the selected nodes are swapped and two offspring are generated.

Mutation:- A genetic programming tree is first selected randomly from the population in the mating pool and any node of the tree is replaced by any other node from the same function or terminal set. A function node can replace only a function node and the same principle is applicable for the terminal nodes.

3.3.1 MULTI-GENE GENETIC PROGRAMMING

Multi-gene genetic programming (MGGP) is a variant of GP and is designed to develop input output relationship of a system in terms of empirical mathematical model which is weighted linear combination of outputs from a number of GP trees. It is also referred as symbolic regression. Each tree represents lower order nonlinear transformations of input variables termed as "gene" and the linear combination of these genes are termed as "multi gene".

Figure 3.1 shows an example of MGGP model where the output is represented as linear combination of two genes (Gene-1 and Gene-2) that are developed using four input variables (x_1, x_2, x_3, x_4). Each gene is a nonlinear model as it contains nonlinear terms (sin (.) / log(.)). In the MGGP model development it is important to make a tradeoff between accuracy and complexity in terms of maximum allowable number of genes (G_{max}) and maximum depth of GP tree (d_{max}). The user specifies the values of G_{max} and d_{max} to have a control over the complexity of MGGP based models. Thus, there are optimum values of G_{max} and d_{max} , which produces a relatively compact model (Searson et al, 2010). The linear coefficients termed as weights of Gene-1 and Gene-2 (c_1 and c_2) and the bias (c_0) of the model are obtained from the training data using statistical regression analysis (ordinary least square method).

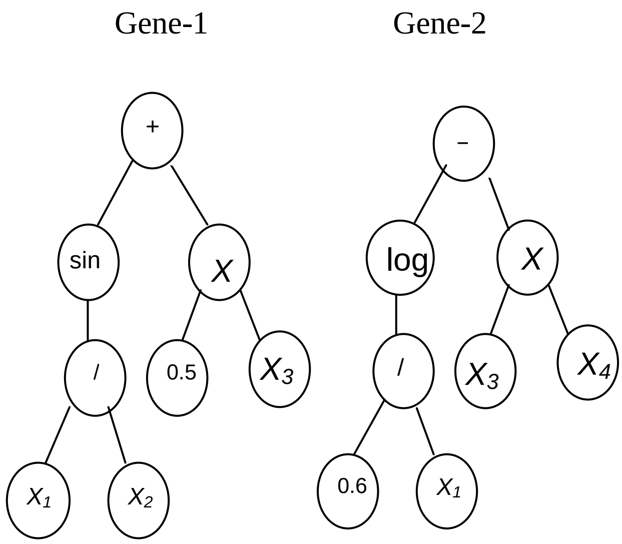

Figure 3.1 An Example of typical Multi-Gene Genetic Programming (MGGP) model.

In MGGP procedure the user defined initial population is generated by creating individuals that contain randomly evolved genes varying from 1 to $G_{\max}$. In addition to the standard GP evolution mechanisms as discussed earlier, there are some special MGGP crossover mechanisms (Searson et al ,2010) which allow the exchange of genes between individuals.

Similarly, MGGP also provides six methods of mutation of genes (Gandomi and Alavi, 2012b). The probabilities of the various evolutionary mechanisms can be set by the user for achieving the best MGGP model. These mechanisms are grouped into categories referred as events. Therefore, the probability of crossover event, mutation event and direct reproduction event are to be specified by the user in such a way that the sum of these probabilities is 1.

The general form of MGGP based model of the present study can be presented as:

$$LI_p = \sum_{i=1}^{n} F[X, f(X), b_i] + b_0 \qquad (3.2)$$

where LI is the predicted value of liquefaction index (LI), F is the function created by the MGGP process referred herein as liquefaction index function, X is the vector of input variables = {$(N_1)_{60}$, $CSR_{7.5}$}, where $(N_1)_{60}$ is the corrected blow count and $CSR_{7.5}$ is the cyclic stress ratio adjusted to the benchmark earthquake (moment magnitude, M_w, of 7.5) as presented by Juang et al. (2000) ,b_i is constant, f is the MGGP function defined by the user, and n is the number of terms of model equation. The MGGP as per Searson et al (2010) is used and the present model is developed and implemented using MatLab (2010).

CHAPTER 4

EVALUATION OF LIQUEFACTION POTENTIAL OF SOIL FROM CPT DATA USING MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND GENETIC PROGRAMMING

4.1 INTRODUCTION

Simplified methods based on standard penetration test (SPT), cone penetration test (CPT), shear wave velocity test, Becker penetration test (BPT) 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 (1971) , 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 used worldwide, CPT is becoming more acceptable due to consistent, repeatable and identification of continuous soil profile (Mayne ,2007) . Hence, recently CPT is being widely used for liquefaction susceptibility analysis of soil using various statistical and regression analysis techniques (Robertson and Campanella (1985) ,Seed and Alba (1986) , Stark and Olson(1995)) . Artificial intelligence (AI) techniques such as; artificial neural network (ANN) (Goh (1994), Juang et al.(2002), Samui and Sitharam (2011)) , support vector machine (SVM)(Pal (2006) , Goh and Goh (2007), Samui and Sitharam (2011)) , and relevance vector machine (RVM) (Samui , 2007) have been used to develop liquefaction prediction models based on an in-situ test database, which 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 a 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 a comprehensive relationship between the inputs and output, and are also called as 'black box' 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 (Ku et al., 2004) using MARS. A comparative evaluation is made among the existing CPT based statistical method.

4.2 DATABASE AND PROCESSING

In the present study database of CPT-based liquefaction case histories in 1999, Chi-Chi, Taiwan, earthquake is used (Juang et al.,2002). The database consists of total 125 total cases, 41 out of them are liquefied cases and other 84 are non-liquefied cases. out of 125 cases 91 data are selected for training and 34 data are selected for testing. Out of 91 training cases 31 sites are liquefied and 60 sites are non-liquefied and out of 34 testing cases 10 sites are liquefied and 24 sites are non-liquefied.

Table 4.1 Data for prediction of liquefaction Index (Training data)

CPT ID	Depth (m)	q_c (MPa)	fs (kPa)	Pw (kPa)	Rf (%)	σv (kPa)	σ' (kPa)	a_{max} (g)	Liq?
C-K1-NT	12.5	7.52	30.9	38.1	0.42	231.3	121.3	0.21	No
C-K1-NT	13.5	7.02	24.3	-2	0.36	249.8	129.8	0.19	No
C-22-YL	14.5	16.89	44	-35.8	0.27	268.3	138.3	0.19	No
C-24-YL	3.5	1.5	24.4	-85.5	2.16	66.6	43	0.12	Yes
C-15-NT	7.5	7.04	30	45.6	0.43	138.6	75	0.12	No
C-LW-A3	5	6.61	41.5	11	0.62	93.6	55	0.12	No
C-36-YL	3.5	2.45	17.1	-18.7	0.72	64.8	44.8	0.19	Yes
C-44-YL	14.5	17.08	69.1	-22.5	0.37	268.3	138.3	0.19	No
C-LW-D2	7.4	5.46	45.9	28	0.84	136.8	74.2	0.12	No
C-36-YL	5	2.96	21.1	-0.2	0.71	92.5	57.5	0.19	Yes
C-LW-A2	3.5	2.09	8.2	-33.2	0.39	64.8	39.8	0.19	Yes
C-LW-D2	3.2	2.66	19.2	-10.4	0.73	59.2	42.2	0.19	Yes
C-LW-A7	8	5.77	25	-45.1	0.45	148	83	0.43	Yes
C-LW-A2	16.5	13.65	17.6	-21.5	0.13	305.3	150.3	0.19	No
C-LW-D1	7.5	7.57	41.4	45.5	0.55	142.5	78.8	0.12	No

C-K2-YL	13.5	14.67	9.8	-30.6	0.07	249.8	124.8	0.19	No
C-43-YL	3.1	1.41	4.9	-40.4	0.39	57.4	46.4	0.43	Yes
C-LW-A3	10.1	7.72	15.5	19.9	0.2	186.9	100.9	0.19	No
C-LW-A5	10.5	6.08	31.7	71.6	0.52	192.6	99	0.12	No
C-LW-A7	6.5	7.03	36.1	10.8	0.51	120.6	67	0.12	No
C-44-YL	14.5	8.01	20.9	23.6	0.26	268.3	138.3	0.19	No
C-LW-A5	18.5	10.05	46.1	2.1	0.45	346	172.3	0.19	No
C-LW-C1	12.5	9.19	33	-63.6	0.4	231.3	121.3	0.19	No
C-LW-A9	12.5	8.3	12.7	-0.1	0.15	231.3	121.3	0.19	No
C-LW-A5	6.5	7.12	50.7	23.1	0.71	120.6	67	0.12	No
C-25-YL	2.5	3.26	9.5	-10.2	0.29	48.6	35	0.12	Yes
C-LW-D2	2.5	2.54	23	-31.2	0.97	46.3	36.3	0.19	Yes
C-K2-YL	6.5	2.69	28.8	11.8	1.09	120.3	65.3	0.19	Yes
C-K2-YL	2.5	3	7.4	-13.8	0.25	46.3	31.3	0.19	Yes
C-24-YL	8.5	7.47	34.8	45.2	0.47	156.6	83	0.12	No
C-LW-D2	4.05	2.61	23.5	-26.5	0.95	74.9	49.4	0.19	Yes
C-LW-A1	12.5	5.47	63.3	16.8	1.17	228.6	115	0.12	No
C-19-YL	3.1	2.54	11.9	-40.9	0.57	57.4	41.4	0.19	Yes
C-LW-A10	12.5	7.38	42.9	52.6	0.57	228.6	115	0.12	No
C-42-ST	14	13.65	21.8	-13.8	0.16	259	134	0.19	No
C-5-YL	2.5	0.23	0.9	76.2	0.42	50	36.3	0.12	Yes
C-22-YL	6.5	7.94	45.1	31.2	0.57	124	70.3	0.12	No
C-LW-A10	17	7.68	60.8	-19.9	0.81	314.5	159.5	0.19	No
C-LW-A2	3.5	2.49	10	3.9	0.41	68.5	44.8	0.12	Yes
C-LW-C2	11.8	8.15	37	-35.9	0.46	218.3	115.3	0.19	No
C-24-YL	18.5	9.48	86.1	48.2	0.79	336.6	163	0.12	No
C-24-YL	2.5	0.92	18.9	-90.2	2.54	48.6	35	0.12	Yes
C-LW-A1	9	6.67	14.2	4.8	0.21	166.5	91.5	0.19	No
C-32-YL	10.35	11.32	114	-79.9	0.73	191.5	108	0.43	No
C-K2-YL	9.5	6.76	64.9	27.2	0.96	174.6	91	0.12	No
C-25-YL	15.5	8.74	41	-48.1	0.46	286.8	146.8	0.19	No
C-15-NT	11.6	7.72	62.6	22.1	0.81	218.3	113.6	0.12	No
C-K2-YL	8.5	5.38	26.1	48.6	0.48	156.6	83	0.12	No
C-LW-D1	8.5	6.73	49.2	37.7	0.73	156.6	83	0.12	No
C-K2-YL	10.5	7.46	35.8	-40	0.48	189	99	0.19	No
C-7-NT	10	11.96	162.2	-55	1.35	185	105	0.43	No
C-9-YL	4.5	6.01	27.2	-30.6	0.46	83.3	58.3	0.43	Yes
C-LW-C2	10.5	8.25	70.6	-39.7	0.86	194.3	104.3	0.19	No
C-LW-A9	3.5	2.65	9.3	13.5	0.36	66.6	43	0.12	Yes
C-36-YL	3.5	11.56	170	-20.8	1.51	68.5	49.8	0.43	No
C-36-YL	12.5	8.27	0.2	0.2	0.24	231.3	121.3	0.19	No
C-24-YL	4.5	1.73	25.8	-15.7	1.59	83.3	53.3	0.21	Yes

C-31-YL	5	2.22	23.4	31.1	1.06	92.5	57.5	0.19	Yes
C-LW-C2	5.5	1.89	6.7	38	0.37	105.5	61.8	0.12	Yes
C-LW-A2	4.5	0.64	9.9	-14.7	1.91	84.6	51	0.12	Yes
C-22-YL	3.7	2.7	32.4	-43.8	1.24	68.5	46.5	0.19	Yes
C-K2-YL	11.5	7.62	27.9	-27.1	0.36	207	107	0.19	No
C-K2-YL	11.5	6.83	24.5	36.6	0.35	212.8	112.8	0.21	No
C-LW-A2	3.5	3.86	24.3	-35.5	0.78	64.8	49.8	0.43	Yes
C-43-YL	3.5	2.62	11	-2.2	0.41	64.8	44.8	0.19	Yes
C-LW-A9	14	12.77	22.8	-0.6	0.18	259	134	0.19	No
C-2-DC	9.5	7.43	57.7	11.7	0.77	179.5	95.8	0.12	No
C-LW-A7	14.5	10.61	19.2	-26.3	0.18	268.3	133.3	0.19	No
C-24-YL	2.6	1.18	11.4	-48.6	0.79	48.1	37.1	0.19	Yes
C-LW-D2	7.5	6.23	1.7	-0.9	0.27	138.8	78.8	0.19	No
C-LW-D1	6.5	7.4	30.3	32.3	0.4	120.6	67	0.12	No
C-10-YL	3.5	0.2	3.7	100.8	1.96	68.5	44.8	0.12	Yes
C-LW-C1	10.5	6.49	55.2	17.3	0.86	192.6	99	0.12	No
C-22-YL	9.5	6.62	37	54.8	0.57	174.6	91	0.12	No
C-LW-C2	9	12.89	138.8	-93	1.08	170.2	96.5	0.43	No
C-LW-A10	5	2.54	13.8	-30.8	0.54	92.5	57.5	0.19	Yes
C-7-YL	12.5	6.8	37.2	-23.6	0.55	231.3	121.3	0.19	No
C-K1-NT	13.5	6.85	59.1	55.3	0.87	246.6	123	0.12	No
C-LW-A10	13.5	16.3	130.1	11.8	0.8	249.8	134.8	0.43	No
C-36-YL	7.9	6.05	43.3	26	0.71	145.8	78.2	0.12	No
C-LW-C1	7.5	8.03	2.6	-5.3	0.32	138.8	78.8	0.19	No
C-LW-A7	11.5	7.41	55.5	23.4	0.76	212.8	112.8	0.19	No
C-LW-A9	6.5	1.54	5.8	53.4	0.41	124	70.3	0.12	Yes
C-3-DC	12.5	7.76	53.9	46.5	0.7	228.6	115	0.12	No
C-7-NT	13.5	8.3	43.3	-8.3	0.53	249.8	129.8	0.19	No
C-2-NT	4.1	0.9	9	8.9	0.59	75.9	54.9	0.43	Yes
C-25-YL	14	12.43	28.2	-28.5	0.23	259	134	0.19	No
C-LW-A2	3.5	1.28	8.8	3.6	1	63	43	0.12	Yes
C-K5-NT	6.5	6.68	41.2	25.4	0.62	124	70.3	0.12	No
C-1-DC	7.5	5.91	28	44.2	0.47	138.6	75	0.12	No
C-35-ST	6	6.64	36.9	40.8	0.55	111.6	63	0.12	No

Table 4.2 Data table for prediction of Liquefaction (Testing Data).

CPT ID	**Depth (m)**	**q_c (MPa)**	**Fs (kPa)**	**Pw (kPa)**	**Rf (%)**	**σv (kPa)**	**σ' (kPa)**	**a_{max} (g)**	**Liq?**
C-19-YL	2.5	0.94	22.4	-27.5	2.54	46.3	41.3	0.43	Yes
C-LW-C1	3.5	1.47	24.6	-40.9	1.94	64.8	49.8	0.43	Yes
C-LW-A1	12.5	10.08	22	-27.1	0.23	231.3	121.3	0.19	No

C-4-DC	2.5	1.62	15.5	-42.9	1	46.3	36.3	0.19	Yes
C-3-DC	4	1.87	23.6	-19	1.3	74	49	0.43	Yes
C-42-ST	12.5	7.58	44.6	77	0.6	228.6	115	0.12	No
C-2-NT	13.5	8	26.8	-65.7	0.36	249.8	129.8	0.19	No
C-22-YL	11.5	8.32	27.1	11.1	0.34	216.5	112.8	0.19	No
C-31-YL	3.5	0.18	0.6	110.5	0.37	68.5	44.8	0.12	Yes
C-LW-C1	19.5	11.26	35.5	-31	0.32	364.5	180.8	0.19	No
C-K2-YL	12.5	7.68	58.7	41.5	0.77	228.6	115	0.12	No
C-32-YL	6.1	7.24	41.4	18.1	0.57	116.6	66.9	0.12	No
C-19-YL	11.5	7.99	43.3	28	0.54	210.6	107	0.12	No
C-LW-A2	13.5	6.54	49.8	26.7	0.76	246.6	123	0.12	No
C-LW-D1	5	5.93	54.4	14.9	0.92	96.2	57.5	0.12	No
C-LW-A9	4.5	2.78	20.7	-15.4	0.74	96.2	48.3	0.19	Yes
C-LW-A7	8	6.61	26	3.6	0.4	148	83	0.19	No
C-K5-NT	7.5	5.59	21.8	52.3	0.4	138.6	75	0.12	No
C-31-YL	8.5	6.12	30.6	47.9	0.51	161	87.3	0.12	No
C-5-YL	13.5	7.41	58.9	37.9	0.79	246.6	123	0.12	No
C-LW-A3	13.9	11.58	29.5	-9.4	0.28	257.2	133.2	0.19	No
C-2-NT	9.5	7.18	45.5	40.7	0.64	179.5	95.8	0.12	No
C-LW-A10	4.5	2.01	5.1	23.7	0.25	87	53.3	0.12	Yes
C-5-YL	13.5	6.32	61.5	-7.9	0.98	246.6	123	0.12	No
C-25-YL	7.5	5.21	28.8	44	0.55	142.5	78.8	0.12	No
C-LW-C1	4.5	1.82	22.8	-31	1.25	83.3	53.3	0.19	Yes
C-LW-C2	8.5	6.21	24.8	47.3	0.4	161	87.3	0.12	No
C-43-YL	15.5	14.74	26.2	-28.3	0.2	286.8	141.8	0.19	No
C-7-NT	7.5	3.05	32.5	4.6	1.07	138.8	73.8	0.19	Yes
C-22-YL	11.1	6.7	46.9	-54.2	0.72	205.4	109.4	0.19	No
C-LW-A1	12.5	8.83	57.7	37.1	0.66	235	121.3	0.12	No
C-LW-A5	13	5.16	62	19.8	1.21	237.6	119	0.12	No
C-LW-A7	14	12.15	0.3	-0.7	0.25	259	134	0.19	No
C-LW-A3	4.5	0.64	27.5	-78.9	4.2	84.6	51	0.12	Yes

4.3 RESULTS AND DISCUSSION.

4.3.1 MARS MODELLING FOR LIQUEFACTION INDEX

The MARS model is developed taking P_L =1 for liquefaction and P_L =0 for the non - liquefaction case. The probability of liquefaction/non-liquefaction of the total 125 cases as obtained using the statistical CPT based method (Juang et al.,2002) , and the liquefaction index (*LI*) determined for the same 125 cases by the proposed MARS model are evaluated.

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 liquified cases and similarly for a 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 as per Juang et al. (2002).

Table 4.3 Results for prediction of occurrence of Liquefaction by MARS.

<table>
<tr><td rowspan="4">CRITERIA FOR P_L</td><td colspan="4">PRESENT STUDY BY MARS</td></tr>
<tr><td colspan="2">TRAINING</td><td colspan="2">TESTING</td></tr>
<tr><td>NO OF SUCCESSFUL PREDICTION</td><td>RATE (%)</td><td>NO OF SUCCESSFUL PREDICTION</td><td>RATE (%)</td></tr>
<tr><td colspan="2">BASED ON 31 LIQUEFIED CASES</td><td colspan="2">BASED ON 10 LIQUEFIED CASES</td></tr>
<tr><td>A (P_L > 0.85)</td><td>28</td><td>90</td><td>8</td><td>80</td></tr>
<tr><td>B (P_L > 0.65)</td><td>28</td><td>90</td><td>10</td><td>100</td></tr>
<tr><td>C (P_L > 0.5)</td><td>30</td><td>97</td><td>10</td><td>100</td></tr>
<tr><td></td><td colspan="2">BASED ON 60 NON-LIQUEFIED CASES</td><td colspan="2">BASED ON 24 NON-LIQUEFIED CASES</td></tr>
<tr><td>A (P_L < 0.15)</td><td>60</td><td>100</td><td>22</td><td>92</td></tr>
<tr><td>B (P_L< 0.35)</td><td>60</td><td>100</td><td>24</td><td>100</td></tr>
<tr><td>C (P_L< 0.5)</td><td>60</td><td>100</td><td>24</td><td>100</td></tr>
<tr><td colspan="3">BASED ON ALL 91 CASES</td><td colspan="2">BASED ON ALL 34 CASES</td></tr>
<tr><td>A</td><td>88</td><td>97</td><td>30</td><td>88</td></tr>
<tr><td>B</td><td>88</td><td>97</td><td>34</td><td>100</td></tr>
<tr><td>C</td><td>90</td><td>99</td><td>34</td><td>100</td></tr>
</table>

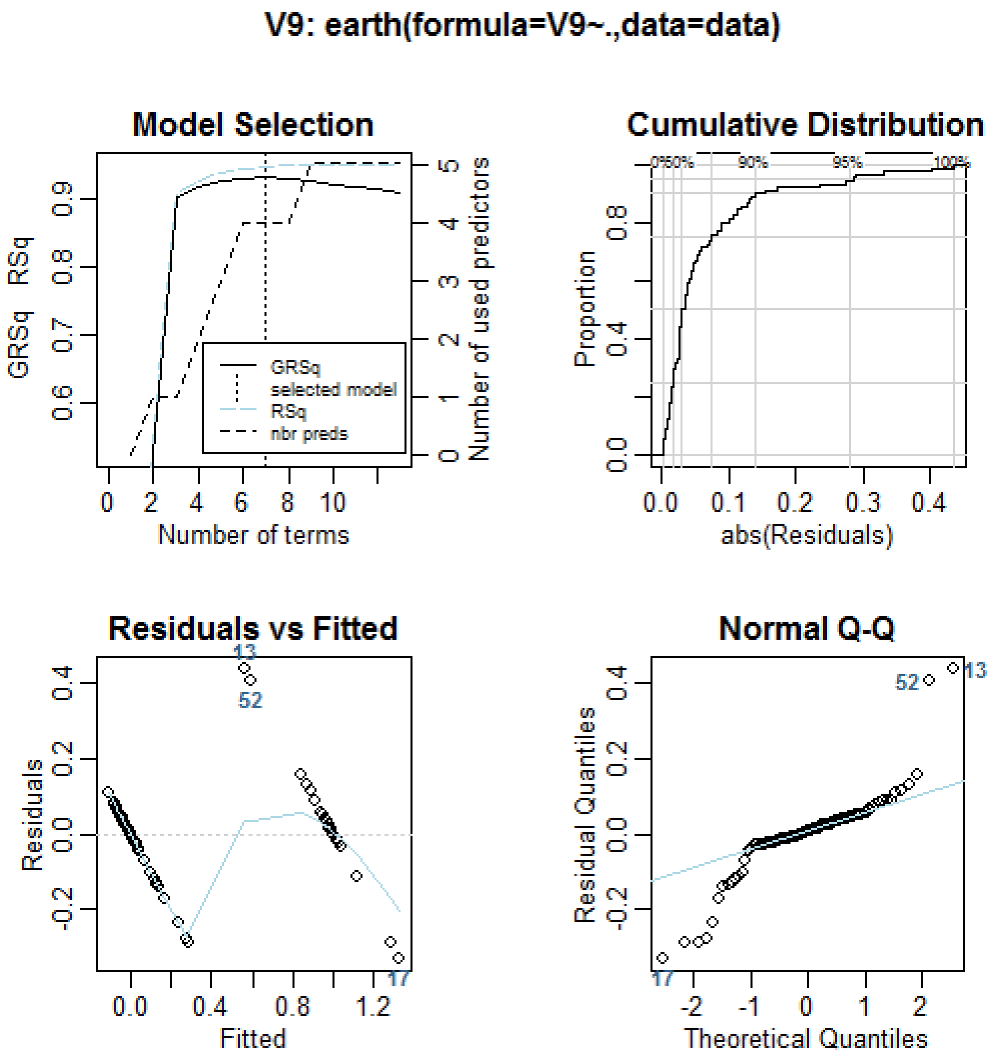

Figure 4.1 Model selection graphs

The model selection graph shows that the RS_q and GRSq lines do not run together at all, this is due to more number of terms. The graph indicates that the best model has 7 terms and uses all 5 predictors. The cumulative distribution graph (Figure 4.1) shows the cumulative distribution of the absolute values of residuals. In this graph, the median absolute residual is about 0.03 and it is observed that the absolute value of residuals are less than about 0.27, i.e. in the training data's 95% of the time the predicted value is within 0.27 units of the observed values. The important variables are also identified and is shown in Table 4.4. It can be seen that **σ'** is the most important input followed by fs and a_{max}.

Table 4.4 Variables and their importance in MARS model.

VARIABLES	N SUBSETS	GCV	RSS
q_c	6	100	100
a_{max}	4	17.8	20
fs	3	12.4	14.7
σ'	2	6.7	9.7

The residual vs fitted graph is for showing the residual for each value of the predicted response. The case 52 has the largest residual while case 13 and 17 have the smallest residual, and they appear suspiciously in separate clusters. The Q-Q graph compares the distribution of the residuals to a normal distribution. The normally distributed residuals will lie on the line. In this case, the divergence is observed from the normality in the left as well as right tail.

MARS model includes 6 Basis functions and the basis functions are listed in Table 4.5.

Table 4.5 Basis functions considered in MARS model and their corresponding Equations.

Basis functions(B_i)	Equations	Coefficients(a_i)
B_1	Max($0, q_c - 2.96$)	-0.2205
B_2	Max($0, q_c - 7.03$)	0.3069
B_3	Max($0, q_c - 8.25$)	-0.0908
B_4	Max($0, f_s - 55.2$)	-0.0043
B_5	Max($0, 75 - \sigma'$)	0.0051
B_6	Max($0, a_{max} - 0.19$)	1.5110

The final equation for prediction of liquefaction index based on MARS model is given below

$$LI=0.8218+\sum_{i=1}^{6} B_i \times a_i \qquad (4.1)$$

4.3.2 GP MODELING FOR PREDICTION OF LIQUEFACTION INDEX.

The GP model is developed taking P_L =1 for liquefaction and P_L =0 for the non - liquefaction case. The probability of liquefaction/non-liquefaction of the total 125 cases as obtained using the statistical CPT based method (Juang et al.,2002), and the liquefaction index (*LI*) determined for the same 125 cases by the proposed GP model are evaluated. 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 a 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 as per Juang et al. (2002)

Table 4.6 Results for prediction of occurrence of Liquefaction by GP.

CRITERIA FOR LI	PRESENT STUDY BY GP			
	TRAINING		TESTING	
	NO OF SUCCESSFUL PREDICTION	RATE (%)	NO OF SUCCESSFUL PREDICTION	RATE (%)
	BASED ON 31 LIQUEFIED CASES		BASED ON 10 LIQUEFIED CASES	
A (P_L > 0.85)	28	90	10	100
B (P_L > 0.65)	28	90	10	100

C ($P_L > 0.5$)	30	97	10	100
	BASED ON 60 NON-LIQUEFIED CASES		BASED ON 24 NON-LIQUEFIED CASES	
A ($P_L < 0.15$)	60	100	23	96
B ($P_L < 0.35$)	60	100	23	96
C ($P_L < 0.5$)	60	100	23	96
BASED ON ALL 91 CASES			BASED ON ALL 34 CASES	
A	88	97	33	97
B	88	97	33	97
C	90	99	33	97

The performance of a GP model depends upon the population size, number of generation, reproduction, crossover and mutation probability, tree depth(d_{max}),number of Genes(G_{max}).In the present study the Liquefaction Index is obtained with population size 1000 indivisuals at 100 generations with reproduction probability of 0.05,crossover probability 0.85, mutation probability of 0.1 and with tournament size of 7.The best result was obtained with G_{max} as 4 and d_{max} as 4.

The develop model is presented below as follows

$$LI = \frac{47.16}{\exp\left(\frac{9.905}{q_c}\right)} - \frac{1385}{exp\left(\frac{7.208}{q_c - a_{max}}\right)} + \frac{1341}{exp\left(\frac{7.158}{q_c - a_{max}}\right)} - \frac{0.0854 \times exp(a_{max}) \times (q_c - a_{max})}{exp\left(\frac{6.781}{\sigma'}\right)} + 1.031 \quad (4.2)$$

Criteria for LI	PRESENT STUDY BY MARS			Criteria for LI	PRESENT STUDY BY GP				Criteria for LI	RESULT BY ANN			
	TRAINING	TESTING			TRAINING		TESTING			TRAINING		TESTING	
	No of successful prediction	No of successful prediction	Rate (%)		No of successful prediction	Rate (%)	No of successful prediction	Rate (%)		No of successful prediction	Rate (%)	No of successful prediction	Rate (%)
	Based on 31 liquefied cases	Based on 10 liquefied cases			Based on 31 liquefied cases		Based on 10 liquefied cases			Based on 31 liquefied cases		Based on 10 liquefied cases	
A (P_L > 0.85)	28	8	80	A (P_L > 0.85)	28	90	10	100	A (P_L > 0.85)	22	71	9	90
B (P_L > 0.65)	28	10	100	B (P_L > 0.65)	28	90	10	100	B (P_L > 0.65)	28	90	9	90
C (P_L > 0.5)	30	10	100	C (P_L > 0.5)	30	97	10	100	C (P_L > 0.5)	30	97	9	90
	Based on 60 non-liquefied cases	Based on 24 non-liquefied cases			Based on 60 non-liquefied cases		Based on 24 non-liquefied cases			Based on 60 non-liquefied cases		Based on 24 non-liquefied cases	
A (P_L < 0.15)	60	22	92	A (P_L < 0.15)	60	100	23	96	A (P_L > 0.85)	4	7	0	0
B (P_L< 0.35)	60	24	100	B (P_L< 0.35)	60	100	23	96	B (P_L > 0.65)	24	40	11	46
C (P_L< 0.5)	60	24	100	C (P_L< 0.5)	60	100	23	96	C (P_L > 0.5)	33	55	17	71
	Based on all 91 cases	Based on all 34 cases			Based on all 91 cases		Based on all 34 cases			Based on all 91 cases		Based on all 34 cases	
A	88	30	88	A	88	97	33	97	A	26	29	9	26
B	88	34	100	B	88	97	33	97	B	52	57	19	79
C	90	34	100	C	90	99	33	97	C	63	69	26	76

4.3.3 COMPARISON OF THE DEVELOPED MARS AND GP MODEL WITH THE EXISTING METHOD.

The present study discussed about the evaluation of liquefaction susceptibility of soil using a non-parametric regression technique, based on statistical methods MARS and GP , and post liquefaction CPT database. In the case of MARS model the probability of liquefaction based on 91 liquefied cases are 88%, 88% and 90% for A,B and C respectively where as for GP it is 97%,97% and 99% for cases A, B and C respectively.The probability of liquefaction for MARS model based on 34 non-liquefied cases are 88%,100% and 100% for A,B and C respectively and for GP model it is 97%,97% and 97% fo the cases A,B and C respectively.It was observed that the prediction as per MARS model is more accurate towards field manifestation in comparison to the developed GP model and the exsisting ANN method. Based on sensitivity analysis σ' is the most important input followed by q_c and R_f.

CHAPTER 5

EVALUATION OF LIQUEFACTION POTENTIAL OF SOIL BASED ON STANDARD PENETRATION TEST USING MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND MULTI-GENE GENETIC PROGRAMMING

5.1 INTRODUCTION

When a saturated or partially saturated soil substantially loses its strength and stiffness in response to an applied stress, such as ground shaking due to earthquake or sudden change in stress condition, causes the soil to behave like a liquid. This phenomenon is known as liquefaction of soil. This phenomenon is most often observed in saturated, loose (low density or un-compacted), sandy soils. The pressures generated during large earthquakes with many cycles of shaking can cause the liquefied sand and excess water to force its way to the ground surface from several meters below the ground. This is often observed as sand boils at the ground surface. Liquefaction is classified in two groups as flow liquefaction and cyclic liquefaction. Flow liquefaction can occur when the shear stress required for static equilibrium of a soil is greater than the shear strength of soil in its liquefied state. Cyclic liquefaction occurs when static shear stress is less than the shear strength of liquefied soil. Liquefaction of saturated sandy soils during earthquakes causes building settlement or tipping, sand blows, lateral spreading, ground cracks, landslides, dam instability, high embankment failures and other hazards. Prediction of liquefaction of saturated sandy soils due to an earthquake is an important task in earthquake geotechnical engineering. Since it is very difficult to get high-quality undisturbed samples of sandy soils, in situ tests have been used to determine the liquefaction resistance of saturated sandy soils. The method of liquefaction resistance based on standard penetration test (SPT) data has been developed by Seed and Alba (1986) . However, there are several limitations in using their methodology to determine the liquefaction resistance of saturated sandy soils. Because of its reliability, speed, economy and continuity of profiling, the CPT test is considered a superior technique for determination of liquefaction resistance. Liquefaction analysis based on probabilistic and statistical methods have been done by many researchers. But all of the above methods have been developed based on some empirical formulae, which are associated with some inherent uncertainties. More recently artificial neural network (ANN) (Goh (1994), Juang et al (2000) , Hanna et al (2007), Samui and Sitharam (2011)) model, Support Vector Machine (SVM) Pal (2006) , Gog and Gog (2007), Samui and Sitharam (2011)) model has been used for prediction of liquefaction potential as a classification problem. It has also been noted that as the knowledge acquired during training is stored in an implicit manner in the ANN, it is very difficult to come up with a reasonable interpretation of the overall structure of the network. These inherent limitations wherein the information or the intervening steps are not available have earned ANN, the reputation of being a "black box" approach. In addition, ANN has several inherent drawbacks such as over fitting, slow convergence speed, poor generalizing performance, and arriving at local minimum. Recently support vector machine (SVM), based on statistical learning theory and structural risk minimization is being used as an alternate prediction model. Another technique, called the Genetic Programming (GP), developed by (Koza,1992) , mimics biological evolution of living organisms and makes use of principle of genetic algorithm (GA). It is also called as 'grey box' model. Various attempts have been made in the recent past to use GP to some Geotechnical engineering problems. GP helps in achieving greatly simplified model formula compared to ANN model, but a trade-off is made between the complexity of the formula and accuracy of the model. Another class of model may be termed as 'white box' model is the multivariate adaptive regression spline (MARS) developed based on statistical model developed by Friedman (1991) MARS can adjust any functional form, hence suitable for exploratory data analysis. Samui et al.(2011)

observed that the MARS model for uplift capacity of suction caisson has better statistical performance comparable to ANN and FEM model.

In the present study an attempt has been made to present MARS and GP models based on post liquefaction SPT database. These models are used to evaluate liquefaction potential of soil. A comparative study among the developed MARS model, MGGP model and the available ANN model are made in terms of rate of successful liquefaction and non-liquefaction cases.

5.2 DATABASE AND PROCESSING

In the present study, database comprising of 288 numbers of Standard Penetration Test (SPT), of liquefaction case history of Chi-Chi, Tawain, 1999 earthquake is used (Hwang and Yang,2001) Out of these 288 cases, 164 cases are identified as liquefied cases and 124 are non-liquefied cases, based on the field test values. Here 202 cases are selected randomly for training and the rest 88 are selected for testing. Samui and Sitharam (2011) and Muduli and Das (2013) also have used the above stated database with the same number of training and testing data while developing ANN, SVM and MGGP based liquefaction models. In case of MARS and MGGP approach, normalization or scaling of data sets are not required as in the case of ANN and SVM approach.

Table 5.1 Data table for prediction of liquefaction Index from post liquefaction SPT data.

Sl no	Depth	N_m	FC (%)	CC (%)	$D_{50(mm)}$	a_{max}	CSR	$N_{1,60}$	Liquefied?
1	9	14	17	9	0.13	0.124	0.140	14.290	0
2	9	21	14	3	0.230	0.124	0.127	20.700	0
3	5	16	46	5	0.090	0.124	0.127	20.610	0
4	7.5	12	55	8	0.080	0.428	0.384	12.140	1
5	8.2	1	42	6	0.111	0.084	0.069	0.970	0
6	7.8	7	16	4	0.300	0.420	0.363	6.990	1
7	1.3	1.5	65	23	0.055	0.789	0.741	3.600	1
8	4.3	9	26	4	0.140	0.211	0.165	10.650	1
9	3.6	6	11	3	2.000	0.420	0.289	7.530	1
10	4.5	7	26	4	0.135	0.211	0.222	9.850	1

11	9	19	10	1	0.260	0.124	0.113	17.720	0
12	6.3	11	30	6	0.110	0.420	0.363	12.130	1
13	8.3	12	13	3	0.560	0.428	0.386	11.620	1
14	16.2	28	31	9	0.300	0.420	0.374	20.820	0
15	12.8	5	26	10	0.110	0.211	0.178	7.950	1
16	7	16	8	1	0.220	0.124	0.131	17.970	0
17	10.3	14	15	5	0.380	0.211	0.228	13.700	1
18	13.2	12	61	6.9	0.068	0.055	0.042	8.950	0
19	6	2	33	7	0.160	0.124	0.130	2.390	1
20	9	21	12	2	0.200	0.124	0.133	21.090	0
21	7.3	13	40	11	0.095	0.789	0.644	12.030	1
22	3.8	6	17	2	0.170	0.211	0.208	8.650	1
23	2.2	6	23	5	0.150	0.420	0.304	7.510	1
24	4	5	21	3	0.140	0.124	0.126	7.270	1
25	13.5	13	14	3	0.160	0.211	0.223	11.530	1
26	9	16	29	6	0.200	0.124	0.135	16.110	0
27	10	20	18	4	0.190	0.124	0.125	19.010	0
28	10	22	15	3	0.180	0.124	0.118	19.910	0
29	19.5	9	46	18	0.093	0.211	0.196	6.600	1
30	5	18	14	3	0.200	0.124	0.137	23.860	0
31	3	2	36	5	0.100	0.124	0.118	3.260	1
32	5.8	11	22	4	0.130	0.789	0.780	12.790	1
33	10	25	14	3	0.220	0.124	0.123	23.490	0
34	5.8	6	10	3	0.280	0.211	0.231	7.460	1
35	7.3	5	16	2	0.210	0.211	0.244	5.760	1
36	4.2	7	27	5	0.190	0.428	0.634	8.520	1
37	10	18	7	1	0.290	0.124	0.126	17.270	0
38	13.5	7	47	5	0.091	0.211	0.226	6.290	1
39	15.7	46	29	5	0.100	0.420	0.384	35.340	0
40	10.9	26	31	8	0.120	0.420	0.355	21.870	1
41	2.8	3	38	11	0.097	0.789	0.760	4.940	1
42	9	22	16	3	0.150	0.124	0.139	22.390	0
43	3.7	9	11	3	0.190	0.165	0.128	11.570	1
44	8.8	13	40	14	0.100	0.789	0.611	11.010	1
45	4	8	15	3	0.180	0.124	0.124	11.410	0
46	13.3	16	11	4	0.340	0.211	0.200	13.290	1
47	12	8	41	6	0.104	0.211	0.234	7.690	1
48	5.5	15	17	3	0.700	0.420	0.420	18.660	1
49	10	20	15	3	0.170	0.124	0.134	19.330	0
50	8.8	5	31	±	0.125	0.165	0.193	5.310	1
51	17.3	13	23	3	0.148	0.181	0.095	7.110	0

52	18.8	9	45	10	0.110	0.181	0.123	5.340	0
53	3.8	3	24	±	0.138	0.165	0.217	5.040	1
54	3	7	5	0	0.200	0.124	0.121	11.800	1
55	8.8	5	24	2	0.400	0.211	0.219	5.130	1
56	8	20	13	2	0.220	0.124	0.129	20.820	0
57	7.4	13	25	4	0.180	0.420	0.416	14.340	1
58	5.3	10	21	4	0.230	0.420	0.294	10.860	1
59	6.8	8	59	9	0.070	1.000	0.633	7.390	1
60	4.2	3	62	17	0.100	0.211	0.417	4.050	1
61	5.8	7	25	7	0.150	0.428	0.283	6.550	1
62	9.5	27	17	4	1.000	0.420	0.296	23.580	0
63	6.3	16	15	5	0.270	0.428	0.410	16.590	1
64	8.8	7	37	11	0.110	0.211	0.228	7.220	1
65	8	20	13	2	0.170	0.124	0.113	19.550	0
66	5	15	17	4	0.170	0.124	0.130	19.800	0
67	3	2	9	1	0.200	0.124	0.118	3.280	1
68	9.2	12	49	3	0.078	0.211	0.203	11.490	1
69	16.8	40	39	8	0.100	0.420	0.275	27.730	0
70	9	7	42	9.4	0.100	0.789	0.714	6.380	1
71	11.8	10	24	5	0.200	0.211	0.165	7.880	1
72	9.1	44	32	5	0.150	1.000	0.613	37.490	0
73	6.5	26	17	4	0.280	0.420	0.353	27.800	0
74	14.1	36	35	12	0.300	0.420	0.364	27.920	0
75	14.8	14	17	1	0.170	0.211	0.187	10.810	1
76	10	18	16	3	0.180	0.124	0.124	17.040	0
77	14.3	14	14	3	0.500	0.211	0.143	9.570	1
78	10.3	10	45	14	0.100	0.211	0.176	8.590	1
79	4.3	4	18	6	0.190	0.211	0.177	5.020	1
80	6	11	8	0	0.200	0.165	0.142	12.070	0
81	9	15	18	4	0.180	0.124	0.141	14.680	0
82	5.8	6	47	7	0.080	0.330	0.325	7.030	1
83	2.4	6	41	9	0.095	0.789	0.514	8.710	1
84	6	17	9	2	0.280	0.124	0.108	18.440	0
85	6.2	6	23	5	0.130	0.420	0.411	7.110	1
86	10	19	10	2	0.250	0.124	0.124	17.990	0
87	10.9	30	21	5	0.013	0.420	0.354	25.900	0
88	5.7	8	16	2	0.170	0.165	0.139	8.820	1
89	9.8	15	23	3	0.149	0.181	0.128	11.840	0
90	12.8	12	44	3	0.111	0.181	0.126	8.400	0
91	6.2	1	42	4.6	0.108	0.084	0.064	1.040	0
92	7.2	6	13	5	0.140	0.165	0.145	6.050	1

93	8	20	18	3	0.180	0.124	0.143	21.570	0
94	11.8	18	12	1	0.610	0.211	0.213	16.150	1
95	10.3	6	31	5	0.110	0.211	0.207	5.600	1
96	2.8	2	55	13	0.060	0.330	0.271	3.020	1
97	10	17	12	2	0.300	0.124	0.134	16.430	0
98	3.8	6	13	4	0.500	0.420	0.391	8.750	1
99	8	20	24	6	0.170	0.124	0.129	20.720	0
100	5	8	14	2	0.130	0.124	0.104	9.320	0
101	10	18	28	5	0.130	0.124	0.134	17.430	0
102	13.5	7	42	3	0.102	0.211	0.228	6.340	1
103	3	12	33	8	0.160	0.124	0.156	18.830	0
104	10	16	13	3	0.290	0.124	0.135	14.920	0
105	5	20	18	4	0.180	0.124	0.119	24.840	0
106	6	7	11	0	0.167	0.211	0.232	8.820	1
107	15.4	20	33	10	0.180	1.000	0.663	14.000	1
108	2.8	4	33	9	0.188	0.428	0.506	5.370	1
109	5.8	3	47	5	0.078	0.211	0.247	3.850	1
110	9	17	16	4	0.180	0.124	0.112	15.740	0
111	4	6	43	3	0.090	0.165	0.124	7.460	0
112	3	3	38	6	0.110	0.211	0.203	4.940	1
113	6.8	10	15	7	0.037	0.789	0.790	10.780	1
114	10	21	21	9	0.150	0.124	0.123	19.620	0
115	8.8	17	39	9	0.100	1.000	0.606	14.480	1
116	11.7	10	13	2	1.200	0.165	0.146	8.440	1
117	7.2	5	30	13	0.024	0.165	0.144	4.990	1
118	9	17	31	5	0.100	0.124	0.139	17.260	0
119	10	20	12	3	0.200	0.124	0.124	18.950	0
120	12	9	39	5	0.108	0.211	0.222	8.330	1
121	10.8	44	32	5	0.160	1.000	0.670	35.220	0
122	4	11	11	0	0.120	0.428	0.356	14.340	1
123	13.2	26	31	8	0.100	0.420	0.352	20.150	1
124	8.1	17	18	6	0.200	0.420	0.347	16.640	1
125	9	12	4	3	0.220	0.124	0.127	11.860	1
126	6	18	19	4	0.190	0.124	0.138	21.970	0
127	18.8	13	15	1	0.164	0.181	0.142	8.290	0
128	10.6	40	14	1	0.300	0.420	0.353	33.430	0
129	2.8	6	22	5	0.180	0.428	0.458	10.620	1
130	4.8	9	29	±	0.129	0.165	0.214	13.400	1
131	3	5	24	6	0.130	0.428	0.300	6.710	1
132	16	11	20	0	0.300	0.165	0.152	8.790	1
133	15.4	28	18	6	0.100	0.420	0.337	19.590	0

134	5	9	20	6	0.150	0.428	0.296	9.400	1
135	17.3	22	39	3	0.120	0.181	0.100	12.260	0
136	7.3	12	18	5	0.210	0.211	0.201	12.540	1
137	8.8	4	30	4	0.100	0.211	0.208	3.940	1
138	6	7	61	13.5	0.075	0.428	0.451	8.470	1
139	7.5	7	47	9	0.091	0.211	0.247	8.140	1
140	3.7	7	28	1	0.100	0.165	0.148	9.930	1
141	10	20	14	4	0.180	0.124	0.125	19.040	0
142	5.3	17	21	3	0.300	0.420	0.339	19.510	1
143	2.8	4	18	2	0.180	0.211	0.197	6.530	1
144	12	7	48	10	0.089	0.211	0.232	6.670	1
145	5.7	5	40	10	0.080	0.165	0.140	5.580	1
146	9	22	10	1	0.220	0.124	0.128	21.800	0
147	5.8	12	49	12	0.075	0.428	0.650	12.450	1
148	8	22	15	4	0.190	0.124	0.138	23.370	0
149	7.7	7	18	3	0.170	0.165	0.142	6.720	1
150	2.1	2	18	4	0.130	0.420	0.330	2.690	1
151	7	17	20	3	0.115	0.124	0.140	19.310	0
152	5.7	13	34	7	0.150	0.420	0.322	14.400	1
153	2.3	8	29	6	0.100	0.420	0.336	10.630	1
154	4.2	5	43	7	0.143	0.428	0.596	5.870	1
155	5.8	10	30	7	0.970	0.428	0.433	12.040	1
156	7.2	6	22	15	0.180	0.165	0.148	6.160	1
157	7.7	11	16	3	0.220	0.165	0.132	10.190	1
158	12	12	13	3	0.162	0.211	0.231	11.400	1
159	7.5	26	17	4	0.300	0.420	0.354	26.280	0
160	9	19	14	3	0.200	0.124	0.127	18.710	0
161	10.3	11	28	5	0.110	0.211	0.191	9.870	1
162	18	13	22	0	0.104	0.165	0.142	9.340	0
163	13.3	10	30	3	0.110	0.211	0.214	8.580	1
164	8.1	11	19	2	0.170	0.165	0.148	10.710	1
165	10	21	17	4	0.180	0.124	0.124	19.890	0
166	9	18	14	3	0.200	0.124	0.139	18.350	0
167	10	23	15	4	0.220	0.124	0.128	21.980	0
168	3	6	30	3	0.127	0.211	0.195	9.580	1
169	5	14	14	3	0.200	0.124	0.138	18.680	0
170	11.3	15	47	4	0.106	0.181	0.124	10.980	0
171	16.2	15	43	16	0.113	0.128	0.124	11.810	0
172	5.8	10	48	12	0.080	0.789	0.822	11.780	1
173	8.7	6	44	12	0.080	0.165	0.148	5.680	1
174	12	10	28	3	0.131	0.211	0.238	9.760	1

175	10.2	6	13	4	0.140	0.165	0.145	5.320	1
176	7.3	5	23	1	0.310	0.211	0.208	5.310	1
177	13	31	20	5	0.400	0.420	0.331	23.490	0
178	8	20	13	2	0.200	0.124	0.144	21.720	0
179	10	21	12	2	0.300	0.124	0.135	20.430	0
180	7.1	25	15	1	0.800	0.420	0.341	24.340	0
181	9	21	14	3	0.200	0.124	0.122	20.230	0
182	14.3	43	39	8	0.080	0.420	0.272	29.280	0
183	4.5	4	32	6	0.123	0.211	0.221	5.610	1
184	7.2	10	36	17	0.126	0.128	0.141	11.430	0
185	4	4	30	4	0.100	0.124	0.124	5.730	1
186	9	20	9	2	0.200	0.124	0.128	19.860	0
187	5.7	6	15	4	0.180	0.165	0.143	6.620	1
188	12.8	15	25	±	0.138	0.165	0.174	13.110	1
189	6.9	46	44	12	0.100	1.000	0.715	49.290	0
190	7.8	7	31	18	0.135	0.128	0.139	7.650	0
191	5.8	5	19	4	0.500	0.211	0.211	5.940	1
192	14.3	13	46	8	0.108	0.181	0.120	8.490	0
193	10	8	45	4	0.080	0.165	0.145	7.250	1
194	12.6	33	29	6	0.300	0.420	0.368	26.670	0
195	3	5	12	2	0.190	0.124	0.119	8.220	1
196	5	14	25	3	0.800	0.420	0.324	15.450	1
197	2.7	4	22	15	0.180	0.165	0.105	5.520	1
198	5	18	14	3	0.200	0.124	0.137	23.860	0
199	2.8	11	33	18	0.155	0.211	0.161	15.950	0
200	10	18	23	5	0.130	0.124	0.111	16.090	0
201	14.7	13	25	6.91	0.160	0.055	0.041	9.200	0
202	9.2	1	39	4.8	0.128	0.084	0.070	0.930	0
203	3.7	9	17	1	0.160	0.165	0.153	12.630	1
204	7.3	11	9	0	0.490	0.211	0.199	11.320	1
205	4.2	7	40	10	0.130	0.428	0.672	8.720	1
206	4.7	12	29	6	0.250	0.420	0.344	14.740	1
207	7.8	10	46	±	0.094	0.165	0.209	11.880	1
208	6	7	14	1	0.160	0.211	0.193	7.790	1
209	5.8	9	40	7	0.160	0.428	0.718	9.880	1
210	10	22	16	3	0.170	0.124	0.122	20.420	0
211	4.2	8	34	7	0.200	0.428	0.609	9.500	1
212	7.7	6	13	0	0.180	0.165	0.134	5.700	1
213	8.8	11	38	12	0.400	0.420	0.308	9.860	1
214	5.8	4	35	7	0.125	0.428	0.360	4.350	1
215	10	19	11	2	0.180	0.124	0.124	18.030	0

216	4.3	4	10	3	0.250	0.211	0.167	4.870	1
217	4.3	4	9	2	0.310	0.211	0.193	5.240	1
218	4.2	3	6	0	0.331	0.084	0.054	3.640	0
219	15.8	17	39	5	0.120	0.181	0.141	11.550	0
220	9	22	9	2	0.200	0.124	0.128	21.870	0
221	10	16	14	2	0.180	0.124	0.123	15.030	0
222	8.2	8	10	2	0.450	0.211	0.202	7.980	1
223	8	17	16	3	0.140	0.124	0.139	18.120	0
224	18	15	32	3	0.100	0.165	0.142	10.770	0
225	9.8	8	36	14	0.126	0.128	0.136	7.950	0
226	11.8	11	31	9	0.130	0.211	0.194	9.410	1
227	8.7	7	42	1	0.080	0.165	0.152	6.860	1
228	17.1	50	20	5	0.100	0.420	0.339	35.090	0
229	16.9	28	49	7	0.080	1.000	0.670	19.250	1
230	7.2	9	29	5	0.185	0.428	0.687	8.930	1
231	11.8	12	17	3	0.220	0.211	0.194	10.270	1
232	9	20	12	2	0.220	0.124	0.128	19.920	0
233	12	12	8	0	0.201	0.165	0.149	10.200	0
234	14.8	11	12	3	0.210	0.211	0.208	8.940	1
235	18.1	48	18	5	0.850	0.420	0.367	34.180	0
236	5.8	3	34	6	0.100	0.211	0.182	3.310	1
237	9	17	15	4	0.190	0.124	0.128	16.940	0
238	6.4	11	16	5	0.400	0.420	0.406	12.630	1
239	14.7	4	39	5.7	0.126	0.084	0.071	3.130	0
240	8.8	8	25	8	0.120	0.211	0.201	7.760	1
241	5	3	22	3	0.065	0.428	0.379	3.670	1
242	10	18	13	3	0.180	0.124	0.110	15.990	0
243	10.4	33	30	12	0.040	0.420	0.406	30.440	0
244	8.8	4	46	19	0.110	0.211	0.199	3.860	1
245	7.7	9	48	5	0.080	0.165	0.150	8.890	1
246	4.9	9	29	6	0.200	0.420	0.323	10.660	1
247	7.5	7	13	1	0.162	0.211	0.248	8.140	1
248	5.8	5	10	4	0.360	0.211	0.217	6.030	1
249	11.8	12	13	1	0.300	0.211	0.215	10.810	1
250	6.5	4	17	4	0.300	0.420	0.412	4.580	1
251	4	2	36	5	0.100	0.124	0.127	2.930	1
252	9	22	6	1	0.220	0.124	0.128	21.930	0
253	8.1	9	16	3	0.190	0.165	0.136	8.480	1
254	3	3	6	3	0.080	0.124	0.120	5.000	1
255	17.6	19	26	5	0.200	1.000	0.699	13.430	1
256	7.7	7	19	1	0.170	0.165	0.151	6.960	1

257	14.5	28	18	6	0.100	0.420	0.322	25.300	0
258	3	4	26	2	0.135	0.211	0.235	6.940	1
259	9	22	7	1	0.280	0.124	0.128	21.930	0
260	2.3	2	22	6	0.150	0.789	0.641	3.330	1
261	8.7	3	23	3.5	0.227	0.084	0.070	2.870	0
262	16	16	18	0	0.140	0.165	0.145	12.080	0
263	11.8	9	36	14	0.126	0.128	0.131	8.190	0
264	16	12	18	4	0.140	0.165	0.145	9.060	0
265	18.1	42	39	8	0.080	0.420	0.274	27.040	0
266	7.7	4	24	4.7	0.220	0.084	0.068	3.910	0
267	14.3	16	33	2	0.132	0.181	0.147	11.530	0
268	15.6	40	39	8	0.100	0.420	0.274	28.470	0
269	10	12	40	7	0.080	0.124	0.130	11.540	0
270	18.8	9	45	4	0.110	0.181	0.130	5.500	0
271	7.3	11	21	7	0.150	0.428	0.306	9.750	1
272	10	22	11	2	0.190	0.124	0.124	20.780	0
273	4	12	26	0	0.110	0.428	0.362	15.780	1
274	7.2	17	18	6	0.200	0.420	0.349	17.390	1
275	3.3	6	34	8	0.110	1.000	0.673	8.610	1
276	3	13	45	14	0.090	0.789	0.697	19.240	1
277	5	16	31	7	0.130	0.124	0.127	20.520	0
278	5.7	4	30	10	0.024	0.165	0.136	4.310	1
279	5.8	6	27	5	0.195	0.428	0.659	6.440	1
280	6.2	5	18	7.9	0.254	0.084	0.065	5.450	0
281	4.2	3	24	5	0.200	0.789	0.378	3.580	1
282	11.7	6	48	14	0.075	0.165	0.144	4.960	1
283	10	22	20	4	0.190	0.124	0.134	21.240	0
284	8.2	7	17	2	0.180	0.165	0.163	7.040	1
285	9	14	13	3	0.180	0.124	0.135	14.120	0
286	10.8	11	20	5	0.130	0.420	0.376	9.770	1
287	20.3	17	30	2	0.137	0.181	0.098	8.980	0
288	10	23	15	4	0.220	0.124	0.128	21.980	0

5.3 RESULTS AND DISCUSSION

For MARS modelling of liquefaction index, following assumptions or predictions were made:

i. For all liquefied cases, the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction)

ii. Similarly for all non-liquefied cases, the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).

iii. If the LI predicted by the technique is greater than 0.5 then the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction) and,

iv. If the LI predicted by the technique is less than 0.5, then the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).

The Training and Testing performance (%) are calculated by using the following formula:

Training Performance (%) or Testing Performance (%)

$$=\left(\frac{No\ of\ data\ predicted\ accurately\ by\ MARS}{Total\ data}\right)100 \qquad (5.1)$$

The various statistical parameters which are required for determining the best model are as defined below:-

The coefficient of efficiency (E) (Das and Basudhar , 2008) is given by the equation below. A model which gives better prediction has always higher value of E

$$E = \frac{E_1 + E_2}{E_1} \qquad (5.2)$$

Where, $E_1 = \sum_1^n \left(LI_m - \overline{LI_m}\right)^2$ (5)

$$E_2 = \sum_1^n (LI_p - LI_m)^2 \qquad (5.3)$$

The fitness of each model is determined by minimizing the root mean square error (RMSE) between the predicted and actual value of the output variable (LI) as the objective function. The RMSE is defined as follows

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(LI-LI_P)^2}{n}} \quad (5.4)$$

Where n is the number of cases in the fitness group. If the errors calculated by using equation (5.4) 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 average absolute error (AAE) for the models are calculated by using the following formulae,

$$\text{AAE}=\text{abs}(LI_m\text{-}LI_p) \qquad (5.5)$$

5.3.1 MARS Modeling For Liquefaction Index (LI).

For predicting the liquefaction index using MARS modeling, two models were considered, which were chosen on the basis of number of input parameters considered for modeling. The successful prediction rates for liquefied cases are shown in Table 5.2:-

Table 5.2

MARS modeling showing input parameters, overall, training and testing performances.

Model No	Input parameters	Overall Performance (%)	Training Performance (%)	Testing Performance (%)
1	Depth, N_m, FC, CC, a_{max}, CSR, $N_{1(60)}$, LI_m	96.53	92.1	90.37
2	CSR, $N_{1(60)}$	92.36	91.58	95.35

It can be seen from the above table that Model 1 gives better prediction accuracy for the occurrence of liquefaction than the Model 2. The various statistical performances of the two models are depicted in Table 5.3

Table 5.3

Statistical performances obtained by mars modeling.

Statistical Performances	Model 1		Model 2	
	TRAINING	TESTING	TRAINING	TESTING
Overall Performance (%)	96.53		92.36	
R	0.93	0.928	.849	.808
E	0.859	0.855	.699	0.615
RMSE	0.186	0.187	.273	.305
AAE	0.035	0.035	0.074	.093

Table 5.4

Variable importance, n subsets, gcv and rss, obtained for model no 1

Variables	n subsets	Gcv	rss
N_m	6	95	95.4
CSR	6	95	95.4
CC	5	27.5	35.6
a_{max}	3	100	100
Depth	2	15.7	20.9
$N_{1(60)}$	2	15.7	20.9

A variable can be designated as important when, it has got maximum number of n subsets value. From the above table it is concluded that, based on sensitivity analysis, N_m and CSR are the most important variables among all. Model 1 includes 7 basis functions, which are listed in Table 45.5 together with their corresponding equations.

Table 5.5 Basis function considered in model 1 and their corresponding equations

Basis functions(B_i)	Equations	Coefficients(a_i)
B_1	Max(0,D-8.1)	-0.0334
B_2	Max(0,N_m-6)	-0.030
B_3	Max(0, N_m-16)	0.030
B_4	18*CC	-0.663
B_5	Max(0,0.221-a_{max})	-3.490
B_6	Max(0,0.208-CSR)	-3.445
B_7	Max(0,$(N_1)_{60}$-13.11)	-0.040

The final equation for prediction of Liquefaction Index based on MARS model no 1 is given below:

$$LI = 1.143 + \sum_{i=1}^{7} Bi \times a_i \qquad (5.5)$$

Table 5.6 Variable importance, n subsets, gcv and rss, obtained for model no 2

Variables	n subsets	gcv	rss
CSR	4	100	100
$N_{1(60)}$	3	70.8	70.7

From the above table it is concluded that, based on sensitivity analysis, CSR is the most important variables among all. Model 1 includes 4 basis functions, which are listed in Table 6 together with their corresponding equations.

Table 5.6 Basis function considered in model 1 and their corresponding equations

Basis functions(B_i)	Equations	Coefficients(a_i)
B_1	Max(0,.194-CSR)	-8.591
B_2	Max(0,$(N_1)_{60}$-6.34)	-0.0396
B_3	Max(0, $(N_1)_{60}$-21.93)	-0.092
B_4	Max(0, $(N_1)_{60}$-25.30)	0.127

The final equation for prediction of Liquefaction Index based on MARS model no 2 is given below:

$$LI = 1.125 + \sum_{i=1}^{4} Bi \times a_i \tag{5.6}$$

5.3.2 GP Modeling for Liquefaction Index (LI).

For predicting the liquefaction index using GP modeling, only one model was formulated by considering all the input variables. The successful prediction rates for liquefied cases are given in Table 7:-

Table 5.7 GP modeling showing input parameters, overall, training and testing performances.

Model No	Input parameters	Overall Performance	Training Performance	Testing Performance
1	Depth, N_m, FC, CC, a_{max}, CSR, $N_{1(60)}$	96.88	97.03	96.51

The equation for prediction of Liquefaction Index based on MGGP model 1 is given below:

$$LI=0.03\times D_{50}-0.03\times D-\frac{2.96}{e^{10.35}}+0.02\times(D_{50}+\left(N_1\right)_{60})\times(CSR-4.732)$$

$$+0.03\times a_{max}(N_1)_{60}-\frac{8.769\times CSR}{\left(CSR+(N_1)_{60}\right)}$$

$$+\frac{0.0002\times(N_1)_{60}\times(N_m-CC)}{a_{max}}+2.425 \qquad (5.7)$$

Table 5.8 Comparison of results of developed mars and GP based LI model with ANN model

Model No	Input Variables	Performance in terms of successful prediction (%)					
		MARS	MGGP	ANN	MARS	MGGP	ANN
		Training data			Testing data		
1	Depth, N_m, FC, CC, a_{max}, CSR, $N_{1(60)}$, LI_m	92.1	97.03	-----	90.37	96.51	-----
2	CSR, $N_{1(60)}$	91.58	94.55	94.55	95.35	94.19	88.37

Generally the efficiency of different models are compared firstly on testing data then on training data (Das and Basudhar,2008) . It is evident from the above table that, for the same database as presented by Samui and Sitharam (2011) the accuracy of prediction of liquefaction susceptibility of soil based on ANN model were 94.55 and 83.77 % for training and testing respectively. In case of MARS the accuracy of prediction by model no 1 were found out to be 92.1 and 90.37 for training and testing respectively, whereas in the case of model no 2 the accuracy of prediction obtained were 91.58 and 95.35 % respectively. Similarly in case of MGGP the accuracy of prediction by model no 1 were found out to be 97.03 and 96.51 respectively, whereas in case of model no 2, the accuracy of prediction were 94.55 and 94.19. Thus, it can be stated that MGGP model in both the cases outperforms the MARS and ANN Models.

The present paper successfully adopted MARS and GP for prediction of Liquefaction Index of a soil. The developed models using MARS and MGGP have shown good predictive abilities than ANN, but in comparison between the four above stated methods, the developed MGGP models outperforms the MARS and ANN models. The equation developed (Eqn 5.7)

by MGGP model can be helpful to the geotechnical engineers for predicting liquefaction susceptibility of soils.

CHAPTER 6

EVALUATION OF LIQUEFACTION POTENTIAL OF SOIL FROM SHEAR WAVE VELOCITY DATA BY USING MULTIVARIATE ADPTIVE REGRESSION SPLINES AND GENETIC PROGRAMMING

6.1 INTRODUCTION

The commonly used technique for evaluation of liquefaction potential of soil is simplified technique which was first developed by Seed and Idriss(1971).Liquefaction evaluation involves evaluation of liquefaction susceptibility, evaluation of liquefaction potential and study of the response of various foundations in liquefied soils.Various in-situ tests are also there for evaluation of liquefaction potential of soil such as Standard Penetration Test(SPT) which were developed by Seed and Idriss (1971), Tokimatsu andYoshimi (1983), Seed et al. (1985), Berrill and Davis (1985),and Law et al. (1990), Cone Penetration test(CPT) developed by Robertson and Campanella (1985), Seed and De Alba(1986), Shibata and Teparaksa (1988), and Stark and Olson(1995).Other in-situ test methods for evaluation of liquefaction potential of soil are Dilatometer test (Marchetti 1982) and Shear wave velocity Test (Andrus and Stokoe,1997).Among all the above discussed methods , Andrus and Stokoe(1997) suggested that Shear Wave Velocity method has the following advantages over all methods.(1) the measurements are possible in soils that are hard to sample, such as gravelly soils where penetration tests may be unreliable (2) measurements can be performed on small laboratory specimens(3) Shear wave velocity is a basic mechanical property of soil materials,directly related to small-strain shear modulus (G_{max}).(4)Shaer wave velocity is a required property in earthquake site response and soil-structure interaction analysis. Liquefaction analysis based on probabilistic and statistical methods have been done by many researchers. But all of the above methods have been developed based on some empirical formulae, which are associated with some inherent uncertainties. More recently artificial neural network (ANN) (Goh(1994),Juang et al(2000),Hanna et al(2007),Samui and Sitharam(2011)) model, Support Vector Machine (SVM) (Pal (2006),Goh and Goh (2007) ,Samui and Sitharam(2011)) model has been used for prediction of liquefaction potential as a classification problem. It has also been noted that as the knowledge acquired during training is stored in an implicit manner in the ANN, it is very difficult to come up with a reasonable interpretation of the overall structure of the network. These inherent limitations wherein the information or the intervening steps are not available have earned ANN, the reputation of being a "black box" approach. In addition, ANN has several inherent drawbacks such as over fitting, slow convergence speed, poor generalizing performance, and arriving at local minimum. Recently support vector machine (SVM), based on statistical learning theory and structural risk minimization is being used as an alternate prediction model. Another technique, called the Genetic Programming (GP), developed by (Koza, 1992), mimics biological evolution of living organisms and makes use of principle of genetic algorithm (GA). It is also called as 'grey box' model. Various attempts have been made in the recent past to use GP to some Geotechnical engineering problems. GP helps in achieving greatly simplified model formula compared to ANN model, but a trade-off is made between the complexity of the formula and accuracy of the model. Another class of model may be termed as 'white box' model is the multivariate adaptive regression spline (MARS) developed based on statistical model developed by (Friedman, 1991). MARS can adjust any functional form, hence suitable for exploratory data analysis. Samui et al. (2011) observed that the MARS model for uplift capacity of suction caisson has better statistical performance comparable to ANN and FEM model.

In the present study an attempt has been made to present MARS and GP techniques to separate Liquefied and non-liquefied cases based on post liquefaction Shear wave velocity data. These models are used to evaluate the liquefaction potential of soil. A comparative study among the developed MARS and GP model and the available Neural Network model are also made in terms rate of successful liquefaction and non- liquefaction cases.

6.2 DATA BASE AND PROCESSING

In the present study, database comprising of 186 numbers of shear wave velocity tests , of liquefaction case history Taken from Andrus and Stokoe (1997). Out of these 186 cases, 89 cases are identified as liquefied cases and 97 are non-liquefied cases, based on the field test values. Here 130 cases are selected randomly for training and the rest 56 are selected for testing.

Table 6.1 data table for prediction of liquefaction index from shear wave velocity data

SL NO	σ_{v0} (kPa)	σ'_{v0} (kPa)	Soil type	V_s (m/s)	a_{max} (g)	M	LIQ?
1	104.4	82	2	145	0.16	7.1	0
2	87.1	75.2	2.5	193	0.27	7.1	0
3	193.6	111	3	179	0.12	6.9	0
4	70.4	54.8	2.5	101	0.13	5.9	0
5	178.2	140.8	1.5	195	0.15	7.1	0
6	122.9	70	4	144	0.19	8.3	1
7	85.4	35.4	1.5	127	0.16	7.6	0
8	142.2	123.5	2.5	168	0.32	7.7	1
9	63	48	2	131	0.03	5.9	0
10	85.4	35.4	1.5	127	0.18	6.6	0
11	85.4	35.4	1.5	133	0.18	6.6	0
12	117	82.2	2.5	105	0.15	7.1	1

13	79.2	55.8	2	90	0.2	6.5	0
14	142.2	123.5	2.5	149	0.15	7.1	0
15	147.2	83	2	157	0.14	7.1	1
16	85.4	35.4	1.5	146	0.22	6.6	0
17	160.8	98.8	4	197	0.5	6.9	1
18	60.1	46.6	2.5	173	0.5	6.5	0
19	178.2	140.8	1.5	177	0.32	7.7	0
20	74.7	57.8	1	105	0.24	5.9	0
21	73.8	57.9	3	135	0.2	6.9	1
22	75.4	58.2	4	161	0.19	7.1	0
23	70.4	54.8	2.5	101	0.3	5.9	1
24	74.7	57.8	1	105	0.21	6.5	0
25	83.8	53.9	1.5	127	0.12	5.9	0
26	60.1	46.6	2.5	173	0.02	5.9	0
27	69.9	60.9	2.5	97	0.27	7.1	1
28	85.4	35.4	1.5	130	0.22	6.6	0
29	85.4	35.4	1.5	146	0.05	6.2	0
30	60.1	46.6	2.5	173	0.18	6.5	0
31	70.4	54.8	2.5	101	0.12	6.5	0
32	85.4	35.4	1.5	127	0.04	6.2	0
33	98.6	79.4	2	138	0.16	7.1	1
34	142.2	123.5	2.5	131	0.32	7.7	1
35	191.8	101.2	1	143	0.12	7.1	0

36	124.9	73.7	1	103	0.12	7.1	1
37	47.3	36	4	122	0.3	6.9	1
38	62.1	49.4	4	109	0.36	6.9	1
39	41	30.5	3	126	0.42	7.1	1
40	63	48	2	131	0.5	6.5	1
41	90	54.7	3	115	0.16	7.5	1
42	109.2	84.3	2	133	0.16	7.1	0
43	60.5	45.6	4	122	0.36	6.9	1
44	102.6	71	3	163	0.15	7.1	1
45	115.8	83.1	3	157	0.24	7.1	1
46	57.5	46.3	4	134	0.36	6.9	1
47	90	501	1.5	98	0.12	7.1	1
48	48.5	38.1	4	154	0.36	6.9	1
49	113.1	81.7	3	176	0.24	7.1	1
50	85.4	35.4	1.5	133	0.05	6.2	0
51	251.6	120.2	2	195	0.08	5.9	0
52	136.6	92.4	2.5	148	0.24	7.1	1
53	60.1	48.1	2.5	145	0.42	7.1	1
54	83.6	62.1	4	136	0.36	7.7	1
55	75.4	58.2	4	154	0.36	7.7	1
56	75.2	53.5	4	274	0.46	6.9	0
57	108.3	78.8	2.5	146	0.24	7.1	1
58	117	82.2	2.5	120	0.15	7.1	1

59	85.4	35.4	1.5	146	0.18	6.2	0
60	67	59.6	2.5	125	0.27	7.1	1
61	98.6	79.4	2	121	0.16	7.1	1
62	110.1	84.7	1.5	143	0.16	7.1	1
63	67.7	57.8	2	135	0.42	7.1	0
64	101.3	69.8	3	162	0.25	7.1	1
65	91.8	57.8	1.5	115	0.27	5.9	1
66	63	48	2	133	0.5	6.5	1
67	60.1	46.6	2.5	164	0.02	5.9	0
68	63.6	46.9	3	116	0.25	7.1	1
69	52.7	40.5	4	94	0.36	6.9	1
70	98.9	79.4	1.5	152	0.16	7.1	1
71	121.6	85.8	3	179	0.24	7.1	1
72	69.9	59.6	3	120	0.15	7.1	0
73	83.8	53.9	1.5	127	0.27	5.9	1
74	251.6	120.2	2	195	0.06	6	0
75	44.5	36	4	102	0.36	6.9	1
76	178.2	140.8	1.5	199	0.15	7.1	0
77	85.6	59.5	3	171	0.25	7.1	1
78	139.9	78.6	2	148	0.14	7.1	1
79	45.5	33.4	4	79	0.19	8.3	1
80	85.4	35.4	1.5	130	0.18	6.6	0
81	83.8	53.9	1.5	124	0.13	6.5	0

82	60.1	46.6	2.5	164	0.18	6.5	0
83	54.2	42.4	3	130	0.25	7.1	1
84	85.4	35.4	1.5	130	0.16	7.6	0
85	142.2	123.5	2.5	158	0.32	7.7	1
86	63	48	2	133	0.03	5.9	0
87	178.2	140.8	1.5	200	0.32	7.7	0
88	154.4	86.4	2	152	0.14	7.1	1
89	105.4	59.2	2	155	0.06	6	0
90	118.5	69.2	2	136	0.14	7.1	1
91	6	48	2	133	0.18	6.5	0
92	67.8	44.5	3	118	0.16	7.5	1
93	91.9	77	2.5	204	0.27	7.1	0
94	85.4	35.4	1.5	146	0.18	6.6	0
95	69.9	60.9	2.5	116	0.27	7.1	1
96	54.3	38.1	2	126	0.19	6.5	0
97	39.4	33.8	4	131	0.36	6.9	1
98	91.8	57.8	1.5	115	0.2	6.5	0
99	74.7	57.8	1	105	0.36	5.9	1
100	79.2	55.8	2	90	0.2	5.9	1
101	142.2	123.5	2.5	149	0.32	7.7	1
102	79.2	55.8	2	90	0.11	5.9	0
103	74.8	53.1	2.5	150	0.25	7.1	1
104	85.4	35.4	1.5	146	0.16	7.6	0

105	105.4	59.2	2	155	0.08	5.9	0
106	51	36	4	206	0.46	6.9	0
107	85.4	35.4	1.5	133	0.04	6.2	0
108	63	48	2	131	0.02	5.9	0
109	185.9	110.9	4	174	0.5	6.9	1
110	85.4	35.4	1.5	146	0.04	6.2	0
111	163	90.6	2	137	0.14	7.1	1
112	115.9	82.4	2.5	134	0.24	7.1	1
113	85.4	35.4	1.5	130	0.05	6.2	0
114	75.4	58.2	4	161	0.36	7.7	1
115	74.7	57.8	1	105	0.12	6.5	0
116	159.4	90.1	1.5	147	0.12	7.1	1
117	85.4	35.4	1.5	130	0.04	6.2	0
118	178.2	140.8	1.5	195	0.32	7.7	0
119	61.9	54.4	3	153	0.15	7.1	1
120	40.6	28.7	4	106	0.29	6.9	1
121	178.2	140.8	1.5	199	0.32	7.7	0
122	918	57.8	1.5	115	0.13	6.5	0
123	39	27.8	4	105	0.29	6.9	1
124	54.3	38.1	2	126	0.51	6.5	1
125	110.9	97.7	3	163	0.16	7.5	0
126	69.2	49.8	2	158	0.42	7.1	0
127	63	48	2	131	0.18	6.5	0

128	81.7	43.6	1	101	0.12	7.1	1
129	168.2	140.2	2.5	150	0.1	6.5	0
130	83.8	53.9	1.5	124	0.27	5.9	1
131	148.5	87.7	3	209	0.25	7.1	1
132	85.4	35.4	1.5	127	0.05	6.2	0
133	38.4	32.4	4	107	0.36	6.9	1
134	57.2	46.2	4	107	0.36	6.9	1
135	148.7	83.7	2	157	0.14	7.1	1
136	142.2	123.5	2.5	131	0.15	7.1	0
137	60.1	46.6	2.5	173	0.03	5.9	0
138	83.8	53.9	1.5	127	0.13	6.5	0
139	106.6	63.5	2	130	0.14	7.1	1
140	60.1	46.6	2.5	164	0.5	6.5	0
141	122.5	86.2	3	145	0.24	7.1	1
142	104.4	82	2	148	0.16	7.1	0
143	85.4	35.4	1.5	133	0.22	6.6	0
144	85.4	35.4	1.5	130	0.18	6.2	0
145	122.5	86.2	3	142	0.24	7.1	1
146	83.8	53.9	1.5	124	0.12	5.9	0
147	82.1	63.9	2.5	143	0.15	7.1	1
148	83.6	62.1	4	136	0.19	7.1	0
149	178.2	140.8	1.5	177	0.15	7.1	0
150	85.4	35.4	1.5	127	0.22	6.6	0

151	51	41	2	126	0.42	7.1	0
152	101.3	60.9	2	131	0.14	7.1	1
153	75.4	58.2	4	173	0.36	7.7	1
154	87.1	75.2	2.5	212	0.27	7.1	0
155	54.3	38.1	2	126	0.06	5.9	0
156	48.1	28.7	2	116	0.42	7.1	1
157	85.4	35.4	1.5	127	0.18	6.2	0
158	85.4	35.4	1.5	133	0.18	6.2	0
159	66.6	57.4	4	271	0.23	6.9	0
160	133.1	77.5	2.5	178	0.14	7.1	0
161	142.2	123.5	2.5	158	0.15	7.1	0
162	142.2	123.5	2.5	168	0.15	7.1	0
163	140.6	105.7	3	220	0.15	7.1	0
164	60.1	46.6	2.5	164	0.03	5.9	0
165	85.4	35.4	1.5	133	0.16	7.6	0
166	83.8	53.9	1.5	124	0.2	6.5	0
167	63	48	2	133	0.02	5.9	0
168	121.6	85.8	3	145	0.24	7.1	1
169	110.1	84.7	1.5	135	0.16	7.1	1
170	70.4	54.8	2.5	101	0.2	6.5	0
171	38.8	32.9	4	128	0.36	6.9	1
172	60.3	39.6	3	143	0.25	7.1	1
173	75.4	58.2	4	173	0.19	7.1	0

174	75.4	58.2	4	154	0.19	7.1	0
175	146.3	82.5	2	146	0.14	7.1	1
176	178.2	140.8	1.5	200	0.15	7.1	0
177	43.3	38.3	4	122	0.36	6.9	1
178	158.5	139.1	3	149	0.48	6.9	0
179	83.8	53.9	1.5	115	0.12	5.9	0
180	54.3	38.1	2	126	0.06	5.9	0
181	79.2	55.8	2	90	0.21	6.5	1
182	83.8	53.9	1.5	127	0.2	6.5	0
183	41.1	32.7	4	105	0.3	6.9	1
184	97	78.8	1.5	117	0.16	7.1	1
185	75.4	48.8	3	154	0.14	7.1	1
186	54.7	35.3	1	122	0.12	7.1	1

6.3 RESULTS AND DISCUSSION

In the present study, database comprising of 186 numbers of shear wave velocity test, of liquefaction case history Taken from Andrus and Stokoe (1997). Out of these 186 cases, 89 cases are identified as liquefied cases and 97 are non-liquefied cases, based on the field test values. Here 130 cases are selected randomly for training and the rest 56 are selected for testing.

6.3.1 MARS MODELLING FOR LIQUEFACTION INDEX

For MARS modeling of liquefaction index, following assumptions or predictions were made:

i. For all liquefied cases, the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction)

ii. Similarly for all non-liquefied cases, the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).

iii. If the LI predicted by the technique is greater than 0.5 then the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction) and,

iv. If the LI predicted by the technique is less than 0.5, then the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).

The Training and Testing performance (%) are calculated by using the following formula:

Training Performance (%) or Testing Performance (%)

$$=\left(\frac{No\ of\ data\ predicted\ accurately\ by\ MARS}{Total\ data}\right)100 \qquad (6.1)$$

Table 6.2 MARS modelling showing overall training and testing performances

Criteria for LI	Present study by MARS			
	Training		Testing	
	No of successful prediction	Rate(%)	No of successful prediction	Rate(%)
	Based on 67 liquefied cases		Based on 22 liquefied cases	
	61	91	20	91
	Based on 63 non- liquefied cases		Based on 34 non- liquefied cases	
	54	86	29	85
	Based on all 186 cases		Based on all 186 cases	
	164		89	

A sensitivity analysis was made to identify the important input parameters and is shown in Table 6.3.

VARIABLES	NSUBSETS	GCV	RSS
ST	9	100	100
v_s	8	80.2	82.6
a_{max}	7	69.3	72.2
σ_v	6	61.3	64.1
σ_v'	4	41.1	45.5

A variable can be designated as important when, it has got maximum number of n subsets value. From the above table it is concluded that, based on sensitivity analysis, ST and $v_s$are the most important variables among all. MARS model includes 9 basis functions, which are listed in Table 3 together with their corresponding equations.

Table 6.4 Basis functions considered in mars model and their corresponding equations

Basis functions(B_i)	Equations	Coefficients(a_i)
B_1	Max(0,$\sigma_v - 142.2$)	-0.0010
B_2	Max(0, $1442.2 - \sigma_v$)	-0.0107
B_3	Max(0, $\sigma_v' - 79.4$)	-0.0143
B_4	Max(0, $\sigma_v' - 123.4$)	0.0175
B_5	Max(0,$3 - ST$)	-0.3479
B_6	Max(0,$V_s - 155$)	-0.0105
B_7	Max(0,$155-, V_s$)	0.0060
B_8	Max(0,$a_{max} - 0.5$)	7.5353
B_9	Max(0,$a_{max} - 0.27$)	-7.2389

The final equation for prediction of Liquefaction Index based on MARS model is given below.

$$Liq = 1.15903 + \sum_{i=1}^{9} B_i \times a_i \qquad (6.2)$$

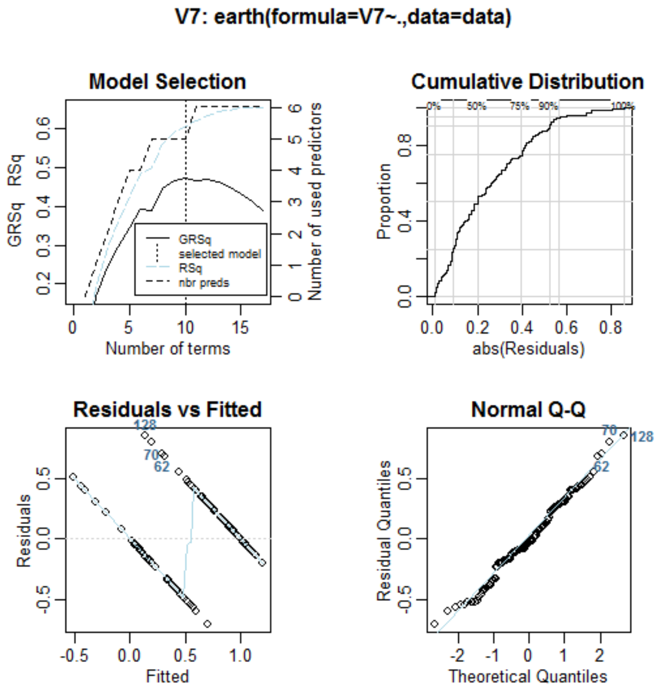

Figure 6.1 Model selection graphs

The graph indicates that the best model has 10 terms and uses all 6 predictors.

The cumulative distribution graph (Figure 6.2) shows the cumulative distribution of the absolute values of residuals. In this graph, the median absolute residual is about 0.50 . The residual vs fitted graph is for showing the residual for each value of the predicted response. The case 128 has the largest residual while case 70 and 62 have the smallest residual, and they appear suspiciously in separate clusters.

The Q-Q graph compares the distribution of the residuals to a normal distribution. The normally distributed residuals will lie on the line.

6.3.2 GP MODELLING FOR LIQUEFACTION INDEX

For GP modeling of liquefaction index, following assumptions or predictions were made:

i. For all liquefied cases, the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction)
ii. Similarly for all non-liquefied cases, the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).
iii. If the LI predicted by the technique is greater than 0.5 then the Liquefaction Index (LI) is assumed as 1, i.e. LI= 1 (for liquefaction) and,
iv. If the LI predicted by the technique is less than 0.5, then the Liquefaction Index is assumed to be 0, i.e. LI = 0 (for non-liquefaction).

The Training and Testing performance (%) are calculated by using equation 6.1

The overall training and testing performances are shown in the table 6.5.

Table 6.5 GP modeling showing overall training and testing performances

Present study by GP			
Training		Testing	
No of successful prediction	Rate(%)	No of successful prediction	Rate(%)
Based on 67 liquefied cases		Based on 22 liquefied cases	
67	100	19	87
Based on 63 non liquefied cases		Based on 34 non liquefied cases	
54	86	27	80
Based on all 186 cases		Based on all 186 cases	
167		90	

The equation for prediction of liquefaction index by GP model is given below

$$Liq = \frac{243.36 \times (a_{max} - 1)}{M} - \frac{0.000123 \times ST \times V_s}{M - 6.395} - \frac{0.00984(a_{max} + M)}{ST \times (2a_{max} - M + 6.395)}$$
$$- \frac{0.00017 \times {V_s}^2 \times a_{max}}{ST}$$
$$- \frac{0.1837 \times V_s \times a_{max} \times M \times (M - 6.395)}{ST(\sigma_v{}' - ST)} \qquad (6.3)$$

Generally the efficiency of different models are compared firstly on testing data then on training data (Das and Basudhar, 2008). It is evident from the above table that, for the same database as presented by Andrus and Stokoe (1997) the accuracy of prediction of liquefaction susceptibility of soil based on Neural Network model was 68% for training and testing respectively. In case of MARS the accuracy of prediction were found out to be 91% for both training and testing r, whereas in the case of GP the accuracy of prediction obtained were 100 and 87% respectively. Thus, it can be stated that MGGP model in both the cases outperforms the MARS and Neural Network Models.

Table 6.6 Comparison of results of developed MARS and GP based LI model with neural network model (Andrus and Stokoe,1997).

<table>
<tr><td rowspan="10">Criteria for LI</td><td colspan="4">Present study by MARS</td><td colspan="4">Present study by GP</td><td rowspan="2">Comparison with existing methods(Andrus and stokoe,1997)</td></tr>
<tr><td colspan="2">Training</td><td colspan="2">Testing</td><td colspan="2">Training</td><td colspan="2">Testing</td></tr>
<tr><td>No of successful prediction</td><td>Rate(%)</td><td>No of successful prediction</td><td>Rate(%)</td><td>No of successful prediction</td><td>Rate(%)</td><td>No of successful prediction</td><td>Rate(%)</td><td>Rate(%)</td></tr>
<tr><td colspan="2">Based on 67 liquefied cases</td><td colspan="2">Based on 22 liquefied cases</td><td colspan="2">Based on 67 liquefied cases</td><td colspan="2">Based on 22 liquefied cases</td><td rowspan="7">68</td></tr>
<tr><td>61</td><td>91</td><td>20</td><td>91</td><td>67</td><td>100</td><td>19</td><td>87</td></tr>
<tr><td colspan="2">Based on 63 non liquefied cases</td><td colspan="2">Based on 34 non liquefied cases</td><td colspan="2">Based on 63 non liquefied cases</td><td colspan="2">Based on 34 non liquefied cases</td></tr>
<tr><td>54</td><td>86</td><td>29</td><td>85</td><td>54</td><td>86</td><td>27</td><td>80</td></tr>
<tr><td colspan="2">Based on all 186 cases</td><td colspan="2">Based on all 186 cases</td><td colspan="2">Based on all 186 cases</td><td colspan="2">Based on all 186 cases</td></tr>
<tr><td colspan="2">164</td><td colspan="2">89</td><td colspan="2">167</td><td colspan="2">90</td></tr>
</table>

The present study successfully adopted MARS and GP for prediction of Liquefaction Index of a soil. The developed models using MARS and GP have shown good predictive abilities than Neural Network, but in comparison between the three above stated methods, the developed GP model outperforms the MARS and Neural Network model. The equation developed (Eqn 6.3) by MGGP model can be helpful to the geotechnical engineers for predicting liquefaction susceptibility of soils.

CHAPTER 7

CONCLUSIONS AND SCOPE FOR THE FUTURE STUDY

7.1 CONCLUSIONS

Applications of MARS and GP in geotechnical engineering are very limited. The focus of this research work was to explore the applications of MARS and GP in evaluation of liquefaction potential of soil.

Based on above study the following conclusions can be made.

- Chapter 4, discussed about the evaluation of liquefaction potential of soil from the post liquefaction CPT data. It was observed that the In the case of MARS model the probability of liquefaction based on 91 liquefied cases are 88%, 88% and 90% for A,B and C respectively where as for GP it is 97%,97% and 99% for cases A, B and C respectively.The probability of liquefaction for MARS model based on 34 non-liquefied cases are 88%,100% and 100% for A,B and C respectively and for GP model it is 97%,97% and 97% fo the cases A,B and C respectively.It was observed that the prediction as per MARS model is more accurate towards field manifestation in comparison to the developed GP model and the exsisting ANN method. Based on sensitivity analysis σ' is the most important input followed by q_c and R_f.
- Chapter 5, discussed about the evaluation of liquefaction potential of soil from the post Liquefaction SPT data. The accuracy of prediction of liquefaction susceptibility of soil based on ANN model was 94.55 and 83.77 % for training and testing respectively. In case of MARS the accuracy of prediction by model no 1 were found out to be 92.1 and 90.37 for training and testing respectively, whereas in the case of model no 2 the accuracy of prediction obtained were 91.58 and 95.35 % respectively. Similarly in case of MGGP the accuracy of prediction by model no 1 were found out to be 97.03 and 96.51 respectively,

whereas in case of model no 2, the accuracy of prediction were 94.55 and 94.19. Thus, it can be stated that MGGP model in both the cases outperforms the MARS and ANN Models.

- Chapter 6 discussed about the evaluation of liquefaction potential of soil from the post liquefaction Shear wave velocity data. The overall performances for MARS and GP and ANN model are 89% , 90% and 68%, respectively The developed models using MARS and GP have shown good predictive abilities than Neural Network, but in comparison between the three above stated methods, the developed GP model outperforms the MARS and Neural Network model.

7.2 SCOPE FOR THE FUTURE STUDIES

The following are the recommendation for further research.

- There is a scope to improve the developed GP and MARS models using new high quality post liquefaction SPT and CPT data.
- Effort should be made to include pore pressure into the limit state function to study its effect on liquefaction triggering.
- After solving liquefaction potential evaluation part of the liquefaction hazard analysis efforts can be focused on developing probabilistic methodology for estimation of seismic soil liquefaction induced ground deformation.

REFERENCES:

1. Andrus, R.D., and Stokoe, K.H. 1997. Liquefaction resistance based on shear wave velocity. *In* Proceedings of the NCEER Workshop on Evaluation of Liquefaction Resistance in Soils. Technical ReportNCEER-97-0022, National Centre for Earthquake Engineering Research, State University of New York at Buffalo, Buffalo, N.Y.,pp. 89–128

2. Baziar, M.H. and Jafarian, Y.(2007).Assessment of liquefaction triggering using strain energy concept and ANN model, capacity energy. *Journal of Soil Dynamics and Earthquake Engineering*, 27, 1056-1072.

3. Berrill, J.B., and Davis, R.O. 1985. Energy dissipation and seismic liquefaction of sands: revised model. *Soils and Foundations*, **25**(2): 106–118.

4. Das, S. K. and Basudhar P. K. ,(2008). "Prediction of residual friction angle of clays using artificial neural network". *Engineering Geology*,100 (3-4), 142- 145.

5. Das, S. K. and Muduli, P. K. (2011). Evaluation of liquefaction potential of soil using extreme learning machine. *Proceedings of the 13th International conference of the International Association for Computer Methods and Advances in Geomechanics, Melbourne, Austrailia, Khalili, N. and Oeser, M., eds.*,1, 548-552.

6. Friedman, J. (1991). " Multivariate adaptive regression splines",*Ann. Stat*. 19, 1–141.

7. Gandomi, A.H., and A.H. Alavi (2012b), " A new multi-gene genetic programming approach to nonlinear system modeling*". Part II: geotechnical and earthquake engineering problems, Neural Comput. Applic.* 21(1), 189-201.

8. Goh, A. T. C. (1994). "Seismic liquefaction potential assessed by neural networks". *Journal of Geotechnical Engineering*, 120 (9), 1467-1480.

9. Goh, A.T.C., and S.H. Goh (2007), “Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data”, *Comput. Geotech*. 34, 5, 410-421.
10. Goh, T.C. and Goh, S.H.(2007). Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data. *Journal of Computers and Geomechanics*,.34, 410-421.
11. Hanna, A.M., D. Ural, and G. Saygili (2007), “Neural network model for liquefaction potential in soil deposits using Turkey
12. Hwang, J.H., and C.W. Yang (2001), “Verification of critical cyclic strength curve by Taiwan Chi-Chi earthquake data”, *Soil Dyn. Earthq. Eng*. 21,(3), 237-257.
13. Ishihara, K. (1993). “Liquefaction and flow failure during earthquakes.”*Geotechnique*, Vol. 43, No. 3, 351-415.
14. Juang CH, Yuan H, Lee DH,Lin PS(2003). “Simplified Cone Penetration Test basedmethod for evaluating liquefaction resistance of soils.”*JGeotechGeoenvironEng* ,Vol.129(1):66-80.
15. 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. *Journal of Soil Dynamics and Earthquake Engineering*, 22, 241-258.
16. Juang, C.H., Chen,C.J., JiangT., andAndrus, R.D. ,(2000),“ Risk -based liquefaction potential evaluation using standard penetration tests”, *Can. Geotech. J*. 37(6), 1195-1208.
17. Juang, C.H., Fang, S.Y. and Khor, E.H. (2006). “First order reliability method for probabilistic liquefaction triggering analysis using CPT.” *Journal of Geotechechnical and Geoenvironmental Engineering (ASCE)*, Vol.132 (3), 337-350.
18. Koza, J.R. (1992), “Genetic Programming: On the Programming of Computers by Means of Natural Selection”, *The MIT Press, Cambridge, MA.*
19. Krammer, S. L. (1996). *Geotechnical earthquake engineering.* Singapur: Pearson Education.

20. Krammer, S.L. (1996) .*Geotechnical earthquake engineering*. India:Pearson Education.

21. Ku, C. S., Lee, D.H. and Wu, J.H. (2004). Evaluation of soil liquefaction in the Chi-Chi, Twain earthquake using CPT. *Journal of Soil Dynamics and Earthquake Engg. 24, 659-673.*

22. MathWorks Inc. (2010), "MatLab User's Manual", Version 7.10, *The MathWorks Inc., Natick.*

23. Mayne, P.W. (2007). Cone penetration testing state-of-practice, Final report. *NCHRP project: 20-05; Task: 37-14; synthesis on cone penetration test.*

24. Moss, R.E.S., (2003). "*CPT-based probabilistic assessment of seismic soil liquefaction initiation*." PhD dissertation, Univ. of California,Berkeley, Calif.

25. Moss, R.E.S., Seed, R.B., Kayen, R.E., Stewart, J.P. and Tokimatsu, K. (2005). "Probabilistic liquefaction triggering based on the cone penetration test." *Geotechnical Special Publication*, 133, E. M. Rathje, ed., (CD Rom) ASCE, Reston, Va.

26. Muduli,P.K.,and Das,S.K.,(2013), "Evaluation of Liquefaction potential of soil based on Standard Penetration Test using multi-gene genetic programming", *Acta Geophysica*, 10.2478/s11600-013-0181-6 .

27. Olsen, R. S. (1997). "Cyclic liquefaction based on the cone penetration test." *Proceedings of the NCEER Workshop of Evaluation of liquefaction resistance of soils.Technical report No.NCEER-97-0022, Youd, T. L. and Idriss, I. M., eds., Buffalo. NY: National center for Earthquake Engineering Research. State University of New York at Buffalo. -225-276*

28. Pal, M. (2006), " Support vector machines-based modelling of seismic liquefaction potential", *Int. J. Numer. An. Met. Geomech*. 30,(10), 983-996.

29. Pal,M. (2006). Support vector machines-based modeling of seismic liquefaction potential. *Journal for Numerical and Analytical Methods in Geomechanics,. 24, 1-27.*

30. Robertson, P.K. and Campanella, R. G. (1985). *Liquefaction potential of sands using the CPT. Journal of Geotechnical Engineering,* 111(3) , 38-403.

31. Samui, P., and T.G. Sitharam (2011), "Machine learning modelling for predicting soil liquefaction susceptibility", *Nat. Hazard. Earth Sys. Sci*. 11, 1-9.

32. Samui, P., Das, S.K., Kim, D.(2011). "Uplift capacity of suction caisson in clay using multivariate adaptive regression spline". *Ocean Engineering*. 38(17–18), 2123-2127.

33. Samui,P.(2007). Seismic Liquefaction potential assessment by using relevance vector machine, *Journal of Earthquake Engineering and Engineering Vibration,* 6(4), 331-336.

34. Samui,P.and Sitharam,T.G. (2011). Determination of liquefaction susceptibility of soil based on field test and artificial intelligence, *International Journal of Earth Science and Engineering,* 4(2), 216-222.

35. Searson, D.P., D.E. Leahy, and M.J. Willis (2010), " GPTIPS: an open source genetic programming toolbox from multigene symbolic regression". *In: Proc. Int. Multi Conf. of Engineers and Computer Scientists*, 17-19 .

36. Seed, H.B. and de Alba, P. (1986). Use of SPT and CPT tests for evaluating liquefaction resistance of sands. *Proceedings of specialty conference on use of in-situ testing in geotechnical engineering, Geotechnical special publication No. 6, ASCE, pp. 281-302.*

37. Seed, H.B. and Idriss, I.M. (1971). Simplified procedure for evaluating soil liquefaction potential. *Journal of the Soil Mechanics and Foundations Division, ASCE,* 97(SM9) ,1249-1273.

38. Seed, H.B., De Alba, P. (1986). " Use of SPT and CPT tests for evaluating liquefaction resistance of sands".*Proceedings of specialty conference on use of in-situ testing in geotechnical engineering, Geotechnical special publication,ASCE* . 6, 281-302.

39. Stark, T.D. and Olson,S.M. (1995). Liquefaction resistance using CPT and field case histories. *Journal of Geotechnical Engineering,* 121(12), 856-869.

40. Tokimatsu, K., and Yoshimi, Y. 1983. Empirical correlation of soil liquefaction based on SPT N-value and fines content. Soils and Foundations, **23**(4): 56–74.

41. Vapnik, V.N.(1998). "Statistical learning theory." Wiley, New York.