Article

Determination of Strain Energy for Triggering Liquefaction Based on Gaussian Process Regression

Jayaraman Karthikeyan, and Pijush Samui*

Centre for Disaster Mitigation and Management, VIT University, Vellore 632014, India.
*Email: pijush.phd@gmail.com

Abstract. The determination of seismic liquefaction potential of soil is an imperative task in earthquake engineering. This article adopts Gaussian process regression (GPR) for determination of amount of strain energy required to induce liquefaction. Effective mean confining pressure (σ'_{mean}), initial relative density after consolidation (D_r), percentage of fines content (FC), coefficient of uniformity (C_u), and mean grain size (D_{50}) are considered as input of the GPR model. The developed GPR gives the variance of predicted output. The results of GPR have been compared with the artificial neural network (ANN). The results of this article show the suitability of the proposed approach for determination of stain energy for triggering liquefaction.

Keywords: Earthquake, liquefaction, Gaussian process regression, variance, artificial neural network.

1. Introduction

Liquefaction is a phenomenon whereby a granular material transforms from a solid state to a liquefied state as a consequence of increase in pore water pressure. Liquefaction of soil causes lot of damages during earthquake. It has been seen from different earthquakes such as Niigata and Alaska (1964), Loma Prieta (1989), Kobe (1995), Chi-Chi (1999), etc. Geotechnical engineers use different methods for determination of liquefaction potential of soil based on in situ techniques [1-11]. Energy-based pore pressure buildup models have been also developed for prediction of liquefaction potential of soil [12-15]. Baziar and Jafarian (2007) [16] successfully used Artificial Neural Network (ANN) for determination of amount of strain energy required to induce liquefaction. They adopted 217 cyclic triaxial data [17], 61 cyclic torsional shear data [15, 18] and six cyclic simple shear tests data for developing the ANN model. The developed ANN outperformed the regression model. However, the developed ANN has some limitations such as black box approach, arriving at local minima, low generalization capability, absence of probabilistic output, etc [19, 20].

This study adopts Gaussian process regression (GPR) for prediction of amount of strain energy required to induce liquefaction. In GPR, the learning of data is modeled as Bayesian estimation problem. It is assumed that the parameters of GPR are random variables. GPR has been successfully adopted for solving different problems in engineering [21-24]. This article uses the database collected by Baziar and Jafarian (2007) [16]. The database contains information about initial effective mean confining pressure (σ' mean), initial relative density after consolidation (D_r), percentage of fines content (FC), coefficient of uniformity (C_u), mean grain size (D_{50}), coefficient of curvature (C_c), and measured strain energy density required for triggering liquefaction (W). The developed GPR model has been compared with the ANN model. The developed GPR has been used to determine the variance of predicted W. The rest of this paper is organized in this way. Section 2 presents the basic principles of GPR and present analysis of the GPR. Section 3 describes the results of GPR model. Section 4 draws some concluding remarks.

2. Details of GPR

Let us consider the following set of samples

$$L = \left\{ x_i, y_i \right\}_{i=1}^{D} \quad x_i \in R^N; \ y_i \in R \tag{1}$$

where x is input variable, y is output, R_N is N-dimensional vector space and R is one dimensional vector space. This article uses σ'_{mean}, D_r, FC, C_u, and D_{50} as input variables. The output of GPR is Log$W$. So, $x = \left[\sigma'_{mean,} D_r, FC, C_u, D_{50} \right]$ and $y = \left[\log W \right]$.

GPR uses the following expression for prediction of y

$$y_i = f(x_i) + \varepsilon_i \tag{2}$$

where $f(x_i)$ is latent function and ε_i is Gaussian noise. GPR treats $f(x_i)$ as a random variable. The joint distribution of y is given by the following equation

$$P(y) = N\left(0, K(x, x) + \sigma^2 I \right) \tag{3}$$

where $K(x, x)$ is kernel function and I is the identity matrix. The predictive distribution of y_{D+1} corresponding to a new given input x_{D+1} is given by the following expression

$$\begin{pmatrix} y \\ y_{D+1} \end{pmatrix} \sim N\left(0, K_{D+1} \right) \tag{4}$$

where K_{D+1} is covariance matrix and its expression is given by

$$K_{D+1} = \begin{bmatrix} K(x,x) + \sigma^2 I & K(x_{D+1}) \\ K(x_{D+1})^T & K(x_{D+1}) \end{bmatrix} \tag{5}$$

The distribution of y_{D+1} is Gaussian with mean and variance [25]:

$$\mu = k(x_{D+1})^T \left[K(x,x) + \sigma^2 I \right]^{-1} y \tag{6}$$

$$\Sigma = K(x_{D+1}) - K(x_{D+1})\left[K(x,x) + \sigma^2 I \right]^{-1} K(x_{D+1}) \tag{7}$$

To develop GPR, a covariance function is required. The details of GPR are given by Williams and Rasmussen (1996) [26].

To develop GPR for prediction of Log$W$, the total datasets have been divided into the following two groups:

- Training Dataset: This is used to develop the GPR. This article uses 198 datasets out of 284 datasets as training dataset.
- Testing Dataset: This is used to verify the developed GPR model. The remaining 86 datasets have been used as testing dataset.

The data are normalized between 0 and 1. The normalization has been done by using the following equation.

$$d_{normalized} = \frac{(d - d_{min})}{(d_{max} - d_{min})} \tag{8}$$

where d is any data (input or output), d_{min} is the minimum value of the entire dataset, d_{max} is the maximum value of the entire dataset, and $d_{normalized}$ is the normalized value of the data. The radial basis function $\left(\exp\left\{ -\frac{(x_i - x)(x_i - x)^T}{2s^2} \right\} \right.$, where s is the width of radial basis function) has been used as covariance function. GPR has been developed by using Matlab.

3. Results and Discussion

The success of GPR depends on the value of ε and s. The design values of ε and s have been determined by trial and error approach. The developed GPR gives best performance at $\varepsilon = 0.02$ and $s = 0.4$. Therefore, the design values of ε and s are 0.02 and 0.4 respectively. The performance of training and testing dataset has been determined by using the design values of ε and s.

Figures 1 and 2 depict the performance of training and testing dataset respectively. The performance of the developed GPR has been assessed in terms of coefficient of determination (R^2). For good model, the value of R^2 is close to one. It is clear from figures 1 and 2 that the value of R^2 is close to one for training dataset as well as testing dataset. So, the developed GPR predicts Log$W$ reasonable well. The developed GPR gives the variance of the predicted Log$W$.
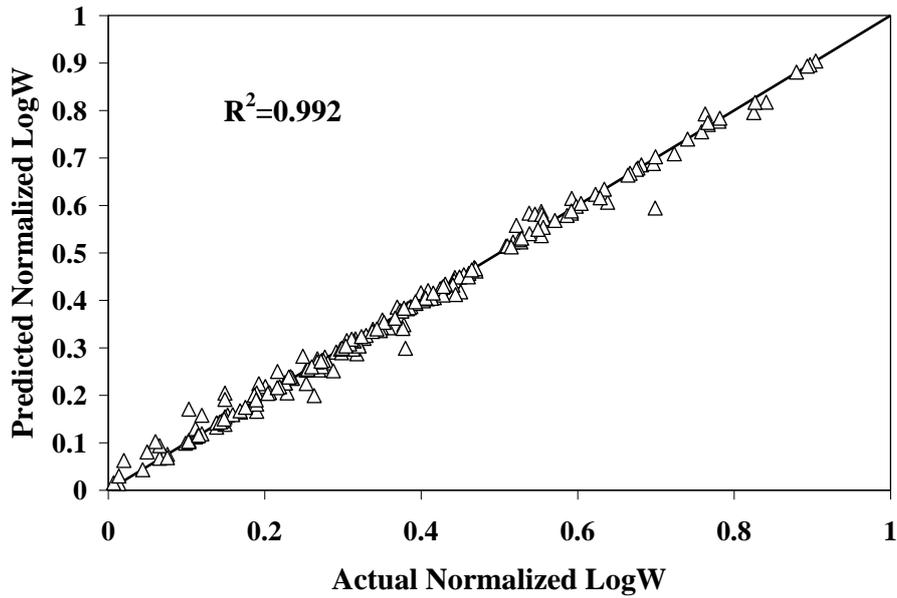
Fig. 1. Performance of training dataset.
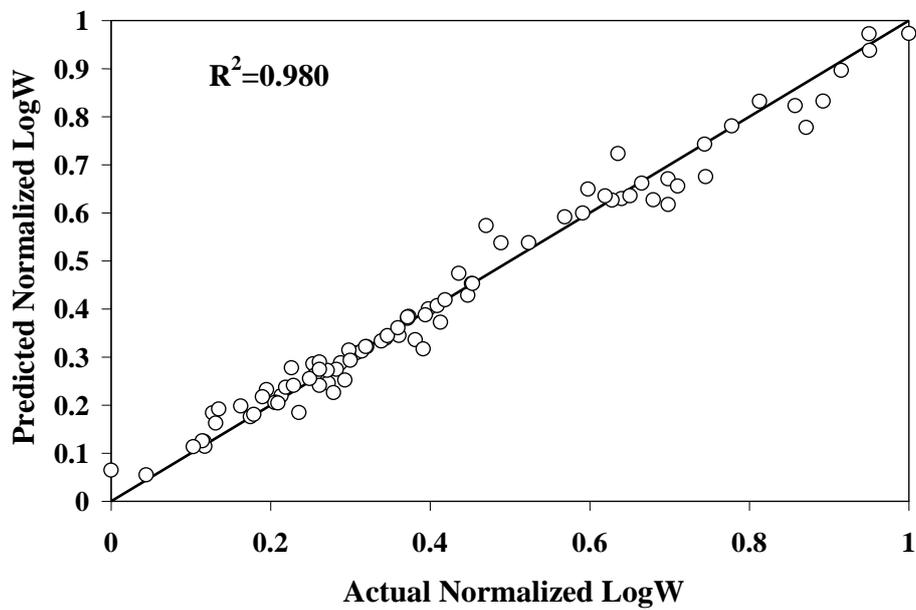
Fig. 2. Performance of testing dataset.

Figures 3 and 4 show the variances of training and testing datasets, respectively. The obtained variance gives the uncertainty of prediction.
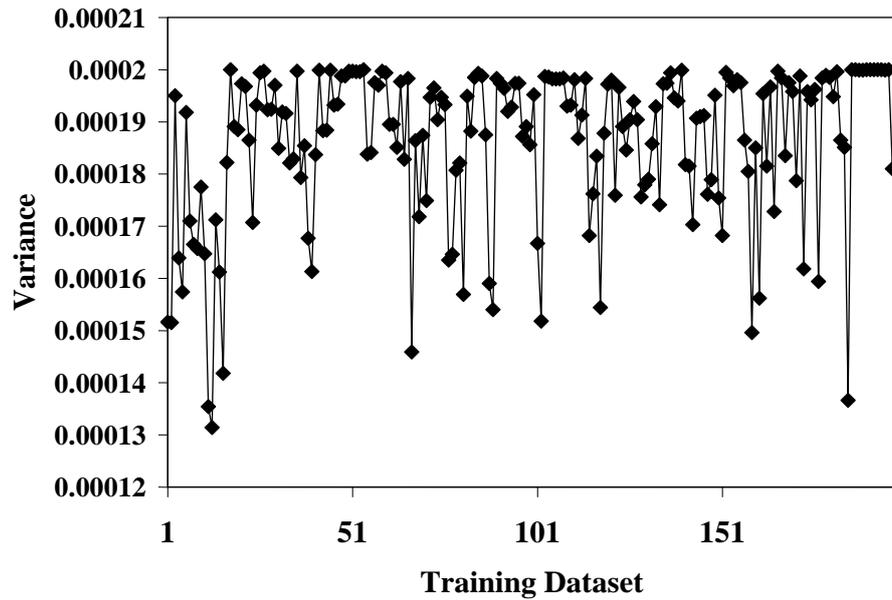
Fig. 3. Variance of training dataset.
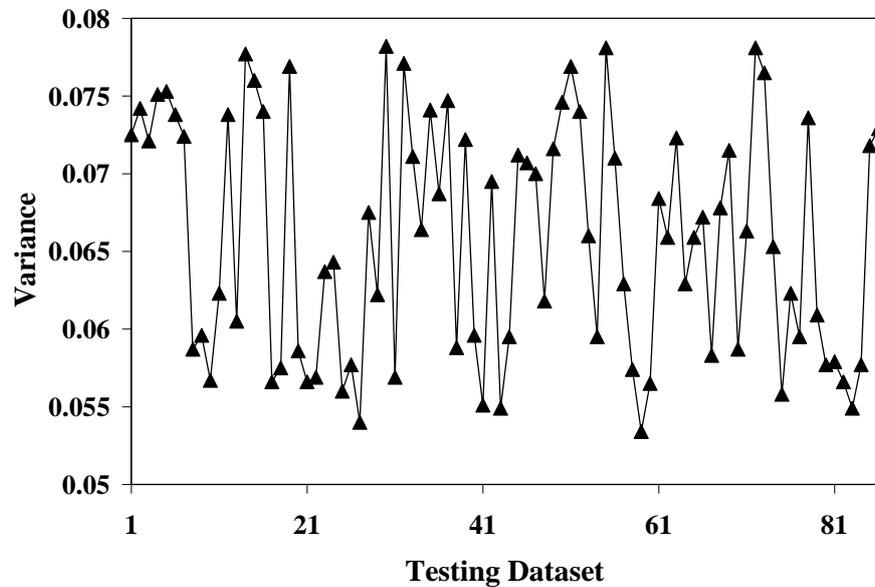
Fig. 4. Variance of testing dataset.

A comparative study has been carried out between the developed GPR and ANN model developed by Baziar and Jafarian (2007) [16]. Comparison has been done in terms of R^2 value.

Figure 5 shows the bar chart of the value of R^2 of the ANN and GPR models. It is clear from Fig. 5 that the performance of GPR is better than the ANN model. ANN was developed by feed-forward multilayer perceptron (MLP) with back-propagation learning rules. There is one hidden layer with 10 neurons in the ANN model. 6,500 epochs have been used to get the best performance of the ANN. The values of mean square error (MSE) and mean absolute error (MAE) have been determined for training and testing dataset. Table 1 shows the value of MSE and MAE for the ANN and GPR models. It is clear from Table 1 that the developed GPR outperforms the ANN. The developed GPR has been applied for the centrifuge ground level liquefaction tests data [27].
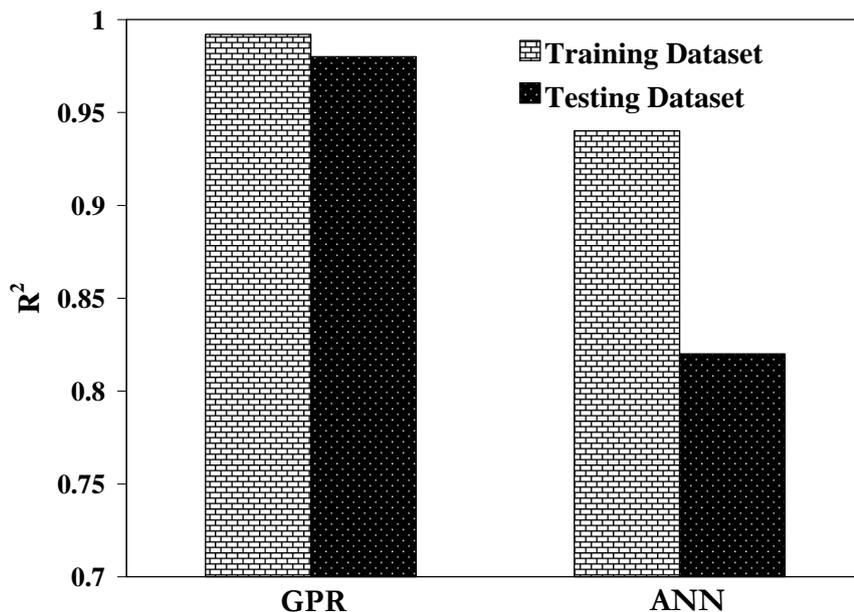
Fig. 5. Bar chart of R^2 of the ANN and GPR models.

Table 1. Comparison between the ANN and GPR models.

Models	Training Performance (MSE)	Testing Performance (MSE)	Training Performance (MAE)	Testing Performance (MAE)
GPR	0.99%	1.82%	3.25%	5.82%
ANN	1.33%	3.7%	8.46%	14.9%

Figure 6 depicts the performance of GPR for the centrifuge ground level liquefaction tests data. It is observed from Fig. 6 that the value of R^2 is close to one. So, the developed GPR shows good generalization capability. The main limitation of the developed GPR is the Gaussian processes are not appropriate priors for all problems.
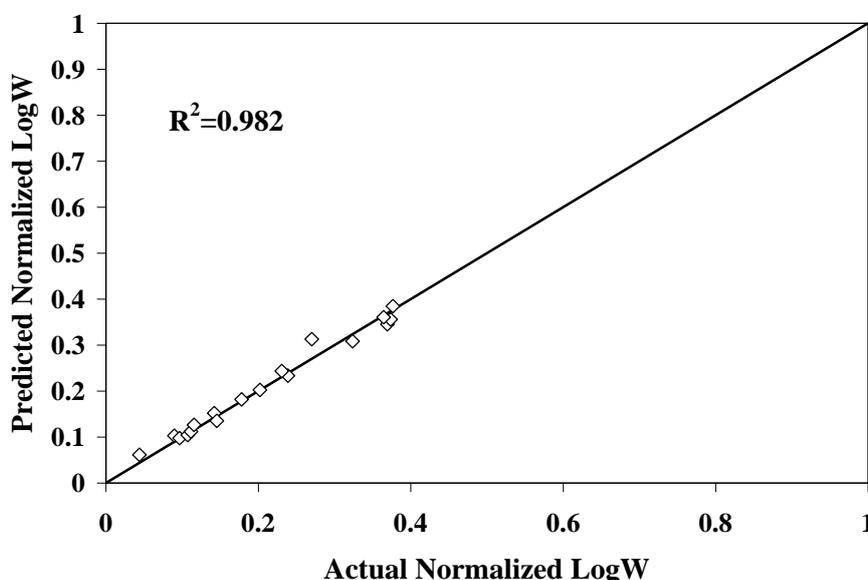
Fig. 6. Performance of the GPR for the centrifuge ground level liquefaction tests data (Dief, 2000) [27].

4. Conclusion

This study examines the capability of GPR for prediction of amount of strain energy required to induce liquefaction. Radial basis function has been used as covariance function for developing the GPR model. The developed GPR gives good performance. It gives better performance than the ANN model. The predicted variance can be used to determine the corresponding risk. This article shows that the developed GPR can be effectively used for the other dataset. GPR can be examined for solving different problems in earthquake engineering.

References

[1] H. B. Seed, and I. M. Idriss, "Analysis of soil liquefaction: Niigata earthquake," *Journal of Soil Mech. and Foun. Div,* ASCE, vol. 93, No. 3, pp. 83-108, 1967.

[2] H. B. Seed, and I. M. Idriss, "Simplified procedure for evaluating soil liquefaction potential," *Journal of Soil Mechanics and Foundation Division*, ASCE, 1971, vol. 97, no. 9, pp. 1249-1273.

[3] H. B. Seed, I. M. Idriss, and I. Arango, "Evaluation of liquefaction potential using field performance data," *Journal of Geotech. Eng. Div.*, ASCE, vol. 109, no. 3, pp. 458-482, 1983.

[4] H. B. Seed, K. Tokimatsu, L. F. Harder, and R. M. Chung, "Influence of SPT procedures in soil liquefaction resistance evaluation," Rep. no. UCB/EERC-84/15, Earthquake Eng. Res. Ctr., Univ. of California, Berkeley, California. 1984.

[5] P. K. Robertson, and R. G. Campanella, "Liquefaction potential of sands using the cone penetration test," *J. of the Geotech. Eng. Div.,* ASCE, 1985, vol. 111, no. GT3, pp. 384-403.

[6] H. B. Seed, and Alba. P. De, "Use of SPT and CPT tests for evaluating the liquefaction resistance of sands," Use of in situ tests in geotechnical engineering, Geotechnical Special Publication 6, ASCE. 1986.

[7] T. D. Stark, and S. M. Olson, "Liquefaction resistance using CPT and field case histories," *Journal of Geotech. Eng,* ASCE, vol. 121, no. 12, pp. 856-869, 1995.

[8] R. S. Olsen, "Cyclic liquefaction based on the cone penetrometer test," Proc., NCEER Workshop on Evaluation of Liquefaction Resistance of Soils, Tech. Rep. NCEER-97-0022, T. L. Youd and I. M. Idriss, eds., National Center for Earthquake Engineering Research, Buffalo, 1997, pp. 225-276.

[9] P. K. Robertson, and C. E. Wride, "Evaluating cyclic liquefaction potential using the cone penetration test," *Canadian Geotech. Journal,* vol. 35, no. 3, pp. 442-459, 1998.

[10] R. D. Andrus, K. H. Stokoe, and R. M. Chung, "Draft guidelines for evaluating liquefaction resistance using shear wave velocity measurements and simplified procedure," NISTIR 6277, National Institute of Standards and Technology, Gaithersburg, 1999.

[11] R. D. Andrus, and K. H. Stokoe, "Liquefaction resistance of soils from shear wave velocity," *Journal of Geotech. and Geoenviron. Eng,* ASCE, vol. 126, no. 11, pp. 1015-1025, 2000.

[12] M. Mostaghel, and K. Habibaghi, "Cyclic liquefaction strength of sands," *Earthquake Eng. Struct. Dyn.,* 1979, pp. 7, 213-33.

[13] J. B. Berrill, and R. O. Davis, "Energy dissipation and seismic liquefaction of sands: revised model," *Soils Found*, vol. 25, no. 2, pp. 106-18, 1985.

[14] K. T. Law, Y. L. Cao, and G. N. He, "An energy approach for assessing seismic liquefaction potential," *Can. Geotech.,* vol. 27, pp. 320-9, 1990.

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

[16] M. H. Baziar, and Y. Jafarian, "Assessment of liquefaction triggering using strain energy concept and ANN model: capacity energy," *Soil Dynamics and Earthquake Eng.*, vol. 27, pp. 1056-1072, 2007.

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

[18] I. Towhata, and K. Ishihara, "Shear work and pore water pressure in undrained shear," *Soils and Foundations*, vol. 25, pp. 73-84, 1985.

[19] D. Park, and L. R. Rilett, "Forecasting link travel times with a multilayer feed-forward neural network," *Comp. Aided Civil and Infrastructure Eng.*, vol. 14, no. 5, pp. 357-367, 1999.

[20] V. Kecman, *Leaming and Soft Computing: Support Vector Machines, Neural Networks, And Fuzzy Logic Models*, MIT press, Cambridge, Massachusetts, London, England, 2001.

[21] O. Stegle, S. V. Fallert, D. J. MacKay, and S. Brage, "Gaussian process robust regression for noisy heart rate data," *IEEE Transactions on Biomed. Eng.*, vol. 55, no. 9, pp. 2143-2151, 2008.

[22] F. Sciascio, A. N. Amicarelli, "Biomass estimation in batch bio-technological processes by Bayesian Gaussian process regression," *Comput. Chem. Eng.*, vol. 32, pp. 3264-3273, 2008.

[23] J. Yuan, K. Wang, T. Yu, and M. Fang, "Reliable multi-objective optimization of high speed WEDM process based on Gaussian process regression", *Int. Journal Mach. Tools Manuf.*, vol. 48, pp. 47-60, 2008.

[24] M. Pal, and S. Deswal, "Modeling pile capacity using Gaussian process regression," *Computers and Geotech.*, vol. 37, pp. 942-947, 2010.

[25] C. K. I. Williams, "Prediction with Gaussian processes: From linear regression and beyond," *In: Jordan MI, editor. Learning and inference in graphical models,* Kluwer Academic Press, pp. 599-621, 1998.

[26] K. I. Williams, and C. E. Rasmussen, "Gaussian processes for regression," MIT Press. 1996.

[27] H. M. Dief, "Evaluating the liquefaction potential of soils by the energy method in the centrifuge," *PhD dissertation*, Dept., of Civil Eng., Case Western Reserve Univ., Cleveland, OH, 2000.