Int. J. Advance Soft Comput. Appl, Vol. 5, No. 3, November 2013
ISSN 2074-8523; Copyright © SCRG Publication, 2013

Determination of Uniaxial Compressive Strength and Modulus of Elasticity of Travertine using Machine Learning Techniques

Manoj Kumar, Pijush Samui, and Ajay Kumar Naithani

Scientist-I, National Institute of Rock Mechanics,
Kolar Gold Fields- 563 117, Karnataka, India
e-mail: manojgeologist84@gmail.com
Professor, Centre for Disaster Mitigation and Management,
VIT University, Vellore-632014, India
e-mail: pijush.phd@gmail.com
Scientist – III, Department of Engineering Geology,
National Institute of Rock Mechanics, P.O- Champion Reefs,
Kolar Gold Fields – 563117, Karnataka, India
e-mail: ajay_naithani@hotmail.com

Abstract

This article adopts machine learning techniques Relevance Vector Machine (RVM), Gaussian Process Regression (GPR) and Minimax Probability Machine Regression (MPMR)} for determination of Uniaxial Compressive Strength (UCS) and the Modulus of Elasticity (E) of Travertine samples. Point load index ($I_{s(50)}$), porosity (n), P-wave velocity (V_p), and Schmidt hammer rebound number (R_n) have been taken as inputs of the RVM, GPR and MPMR model. The outputs of RVM, MPMR and GPR are UCS and E. The developed RVM gives equations for prediction UCS and E. The performance of GPR, MPMR and RVM has been compared with the Artificial Neural Network (ANN) models. The simulation results show that the proposed methods give encouraging performance for prediction of UCS and E of Travertine samples.

Keywords: *Uniaxial Compressive Strength, Modulus of Elasticity, Relevance Vector Machine, Gaussian Process Regression, Minimax Probability Machine Regression, Artificial Neural Network, Travertine samples.*

1 Introduction

This document can be used as a template for Microsoft Word versions 6.0 or later. You may open this document then type over sections of the document or cut and paste to other document and then use adequate styles. The style will adjust your fonts and line spacing. Please set the template for A4 paper (21 x 29.7 cm). For emphasizing please use italics and do not use underline or bold. Please do not change the font sizes or line spacing to squeeze more text into a limited number of pages. Uniaxial Compressive Strength (UCS) and the modulus of elasticity (E) are key parameters in rock engineering. Experimental methods are difficult task for determination UCS and E. There are different methods available for prediction of UCS and E in the literatures [1-11]. These methods are based on index tests. Index tests are very easy to carry out. Sophisticated instrument is not required for conducting index tests. It can be easily done in field. However, the index tests have always uncertainty. It was found that there was no agreement between the equations suggested by these methods [11]. While some equations show the same trend, the others differ [11]. Dehghan et al.[12] successfully adopted Artificial Neural Network (ANN) for prediction UCS and E of Travertine samples. However, ANN has various drawbacks such as low generalization capability, black box approach, arriving at local minima, overtraining problem, etc [13;14]. The developed ANN did not give any equation. So, user can not use the developed model easily.

This article employs machine learning techniques {Gaussian Process Regression (GPR), Relevance Vector Machine (RVM), and Minimax Probability Machine Regression (MPMR)} for prediction of UCS and E of Travertine samples. GPR is non-parametric Bayesian regression [15;16]. It is constructed within a Bayesian framework. There are lots of applications of GPR in literatures [17-19]. RVM is developed by Tipping [20]. It is a Bayesian sparse kernel technique for regression and classification [21]. Researchers have successfully adopted RVM for solving different problems in engineering [22-24]. MPMR is developed based on Minimax Probability Machine Classification (MPMC) [25;26]. It is a discriminant classifier. It has been successfully applied in the different fields of engineering [27;28]. This article uses the database collected from the work of Dehghan et al.[12]. The database contains information about point load index ($I_{s(50)}$), porosity (n), P-wave velocity (V_p), Schmidt hammer rebound number (R_n), UCS and E. A comparative study has been presented between the developed GPR, RVM, MARS and ANN models developed by Dehghan et al.[12]. This article is organized as follows. The descriptions about GPR, RVM and MPMR have been presented in section 2. Section 3 describes the results and discussions. Major conclusions have been drawn in section 4.

2 Proposed Methods

GPR uses the following equation for prediction of output(y).

$$y_i = f_i(x_i) + \varepsilon_i \tag{1}$$

Where x is input, y is output and ε_i is Gaussian noise with zero mean and variance (σ_n^2) [29]. This article uses $I_{s(50)}$, n, V_p, and R_n as input variables. The output of GPR is UCS and E. So, $x = [I_{s(50)}, n, V_p, R_n]$ and $y = [UCS, E]$.

Output variable(y) takes the following form:

$$y \sim N(0, K(x,x) + \sigma_n^2 I) \tag{2}$$

Where K(x,x) is covariance matrix.

For a new input x_{N+1}, we can write,

$$\begin{pmatrix} y \\ y_{N+1} \end{pmatrix} \sim N(0, K_{N+1}) \tag{3}$$

Where,

$$K_{N+1} = \begin{bmatrix} K & k(x_{N+1}) \\ k(x_{N+1})^T & k(x_{N+1}) \end{bmatrix} \tag{4}$$

$k(x_{N+1})$ is the N × 1 vector of covariances between training inputs and the test input.

The distribution of y_{N+1} is Gaussian with mean and variance:

$$\mu(x_{N+1}) = k(x_{N+1})^T K^{-1} y \tag{5}$$

$$\sigma^2(x_{N+1}) = k(x_{N+1}) - k(x_{N+1})^T K^{-1} k(x_{N+1}) \tag{6}$$

To develop GPR, the total datasets have been divided into the following two groups:

Training Dataset: This is used to construct the GPR model. This article uses 23 out of 30 datasets as training dataset.

Testing Dataset: This is used to verify the developed GPR. The remaining 7 datasets have been adopted as testing dataset. Radial basis function ($K(x_i, x) = \exp\left[-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right]$, σ is width of radial basis function) has been used as covariance function. The data is normalized between 0 and 1. Table 1 shows the statistical parameters of the dataset.

Table 1: Statistical parameters of the dataset

Variables	Mean	Standard Deviation	Kurtosis	Skewness
$I_{s(50)}$(MPa)	3.17	0.41	0.07	2.64
V_p(km/s)	5.35	0.27	-0.17	2.15
R_n	27.87	1.43	0.30	1.99
n(%)	6.57	2.91	-0.37	1.59
E(GPa)	5.36	1.98	1.27	4.44
UCS(MPa)	38.99	12.58	0.90	2.94

In RVM, output(y) is determined by using the following equation.

$$y = \sum_{i=1}^{n} w_i K(x_i, x) + a_0 \tag{7}$$

where $a = [w_0, w_1, w_2, ..., w_n]$, x is input, w is weight and $K(x_i,x)$ is kernel function. This article uses $I_{s(50)}$, n, V_p, and R_n as input variables. The output of RVM is UCS and E. So, $x = [I_{s(50)}, n, V_p, R_n]$ and $y = [UCS, E]$.

Given a dataset of input-target pairs $\{x_n, y_n\}_n^N = 1$, we follow the standard formulation and assume p(y|x) is Gaussian $N(y|y(x),\sigma^2)$. The mean of this distribution for a given x is modelled by y(x) as defined in Eq.(7).The likelihood of dataset can be written as

$$p(y|w,\sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\|y - \Phi w\|^2\right\} \tag{8}$$

Where, $y_i = (y_1, ..., y_N)^T$, $\omega_i = (\omega_0, ..., \omega_N)$ and

$$\Phi^T = \begin{bmatrix} 1 & K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_n) \\ 1 & K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & K(x_n, x_1) & K(x_n, x_2) & \cdots & K(x_n, x_n) \end{bmatrix}$$

RVM assumes a Gaussian prior on the kernel weights. Gaussian prior has zero mean and variance (α_i^{-1} ,α is hyperparameters). Typically, new higher-level parameters are used to constrain an explicit zero-mean Gaussian prior probability distribution over the weights

$$p(w|\alpha) = \prod_{i=0}^{N} N(w_i|0, \alpha_i^{-1}) \tag{9}$$

Using Bayes' rule, the posterior over all unknowns could be computed, given the defined non-informative prior-distributions. To complete the specification of the prior-distribution, one must define hyperpriors over α and noise variance σ^2. These quantities are examples of scale parameters and suitable prior are Gamma Distributions [20]

$$p(\alpha) = \prod_{i=0}^{N} \text{Gamma}(\alpha_i|a, b), \tag{10}$$

$$p(\beta) = \prod_{i=0}^{N} \text{Gamma}(\beta|c, d) \tag{11}$$

Where, $\beta = \sigma^{-2}$.Thus, for α and σ it is gamma distribution and for w it is normal distribution and after the prior-distributions, Bayes rule is applied.

$$p(w, \alpha, \sigma^2|y) = \frac{p(y|w, \alpha, \sigma^2)p(w, \alpha, \sigma^2)}{p(y)} \tag{12}$$

Then, given a new test point (X_*), predictions are made for the corresponding target (y_*), in terms of the predictive distribution :

$$p(y_*|y) = \int p(y_*|w, \alpha, \sigma^2)p(w, \alpha, \sigma^2|y)dw\, d\alpha\, d\sigma^2 \tag{13}$$

However, computing the solution of the posterior in equation (13) directly is not possible, since we cannot perform the normalizing integral $p(y) = \int p(y|w, \alpha, \sigma^2)p(w, \alpha, \sigma^2|y)dw\, d\alpha\, d\sigma^2$. Instead, we decompose the posterior as in equation (13)

$$p(w, \alpha, \sigma^2|y) = p(w|y, \alpha, \sigma^2)p(\alpha, \sigma^2|y) \tag{14}$$

It should be noted that one can compute analytically the posterior distribution over the weights since its normalization integral is convolution of Gaussians [20]. Thus,to facilitate the solution. The posterior distribution of weights is given by

$$p\left(w|y,\alpha,\sigma^2\right)=\frac{p\left(y|w,\sigma^2\right)p(w,\alpha)}{p(y|\alpha,\sigma^2)} \tag{15}$$

Thus, the posterior over the weights is then obtained from Bayes'rule

$$p\left(w|y,\alpha,\sigma^2\right)=(2\pi)^{-(N+1)/2}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(w-\mu)^T\Sigma^{-1}(w-\mu)\right\} \tag{16}$$

Eq. (16) has an analytical solution where the posterior covariance and mean are

$$\Sigma=\left(\sigma^{-2}\Phi^{\mathrm{T}}\Phi+\mathrm{A}\right)^{-1} \tag{17}$$

$$\mu=\sigma^{-2}\Sigma\Phi^{\mathrm{T}}\mathrm{t} \tag{18}$$

where we have defined A = (α_0, $\alpha_{1\ldots}$ α_N) .

RVM removes weights from the model for which the corresponding α_i tends to infinite. The inputs corresponding to the remaining nonzero w are called relevance vectors. RVM uses relevance vectors for final prediction. RVM uses the same training dataset, testing dataset, and normalization as used by the GPR model. Radial basis function has been adopted as kernel function.

In MPMR, the relation between input(x) and output(y) is given by the following equation.

$$y=\sum_{i=1}^{N}\beta_i K(x_i,x)+b \tag{19}$$

Where N is number of datasets, K(x_i,x) is kernel function, β_i and b are output the MPMR algorithm. This study uses $I_{s(50)}$, n, V_p, and R_n as input variables. The output of MPMR is UCS and E. So, $x=\left[I_{s(50)},n,V_p,R_n\right]$ and $y=\left[UCS,E\right]$.

MPMR is an evolution of Minimax Probability Machine Classification (MPMC) by constructing a dichotomy classifier [25]. One data set is obtained by shifting all of the regression data +ε along the output variable axis. The other is obtained by shifting all of the regression data -ε along the output variable axis. The following two classes of points are crated from the original dataset.

$$u_i=\left(y_i+\varepsilon,x_{i1},x_{i2},...,x_{id}\right) \tag{20}$$

$$v_i=\left(y_i-\varepsilon,x_{i1},x_{i2},...,x_{id}\right) \tag{21}$$

The details of MPMR are given by Strohmann and Grudic [25]. MPMR adopts the same training dataset, testing dataset, normalization technique and kernel function as used by the RVM. The programs of GPR, RVM and MPMR have been implemented by using MATLAB.

3 Results and Discussion

For GPR, the design values of ε and σ have been determined by trial and error approach. For developing ANN model, the trial and error approach has been used by Dehghan et al.[12]. The developed GPR gives best performance at ε=0.01 and σ=0.4 for prediction of E. So, the design values of ε and σ are 0.01 and 0.4 respectively. Fig. 1 depicts the performance of training dataset for prediction of E.

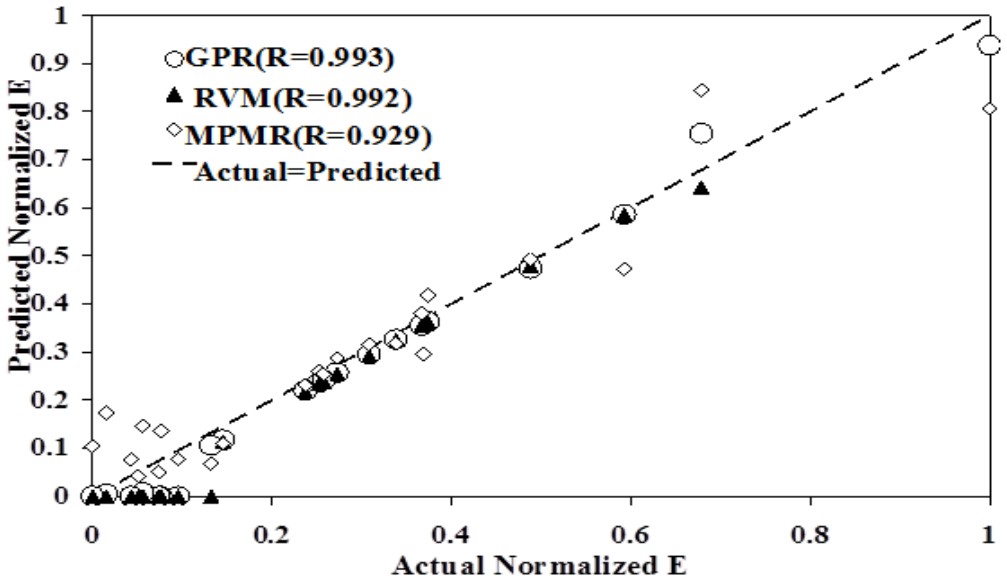

Fig. 1. Performance of training dataset for prediction of E

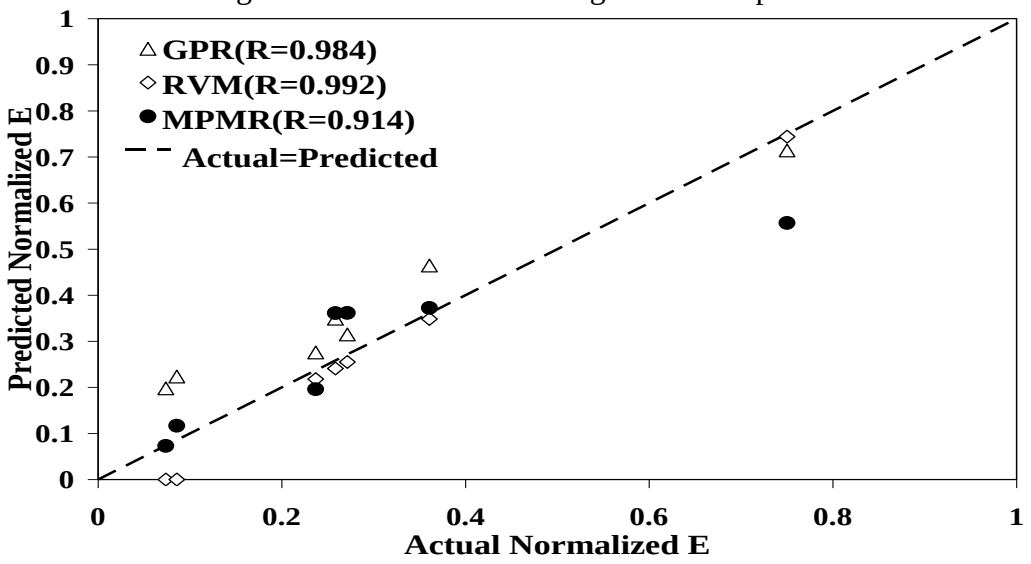

Fig. 2. Performance of testing dataset for prediction of E

The performance of testing dataset has been shown in Fig. 2. The value of Coefficient of Correlation(R) has been shown in Fig. 1 and 2. For a good model, the value of R should be close to one. For prediction UCS, the design values of ε and σ are 0.03 and 0.6 respectively.

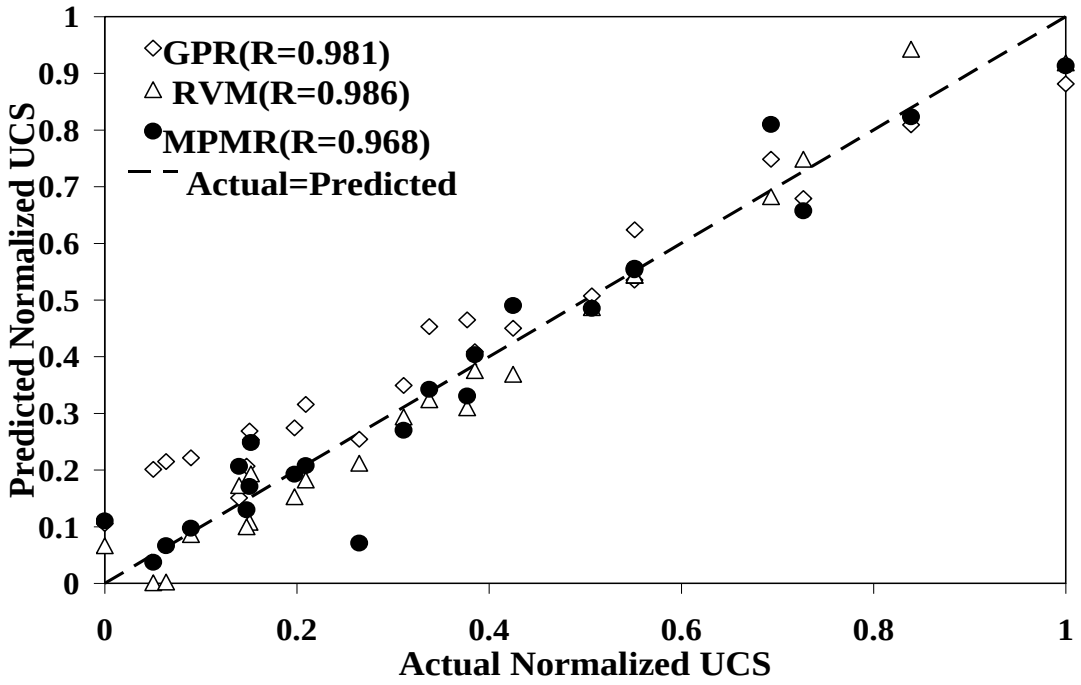

Fig. 3. Performance of training dataset for prediction of UCS

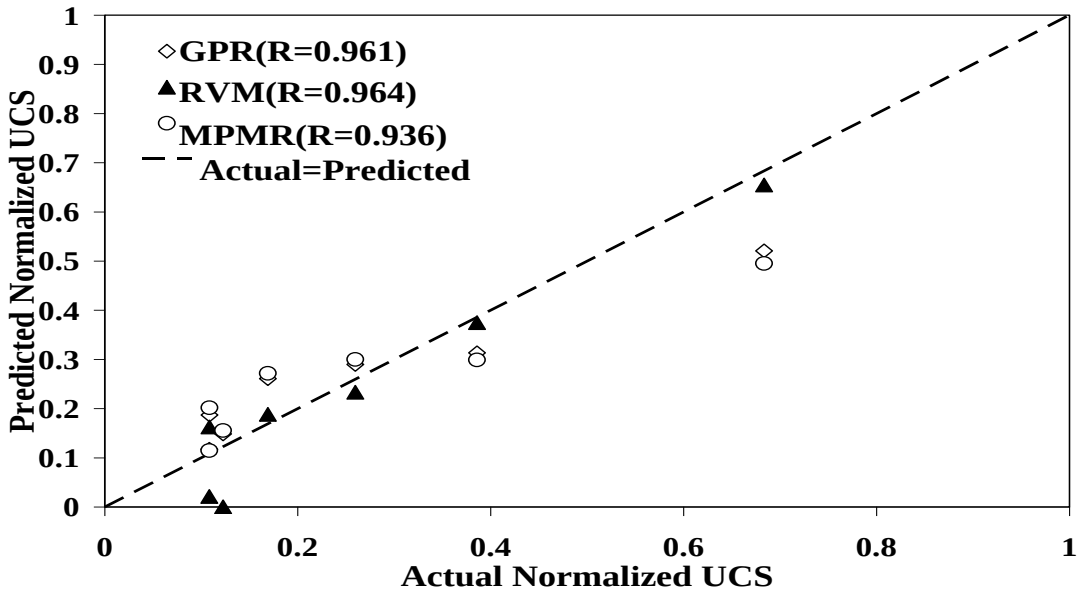

Fig. 4. Performance of testing dataset for prediction of UCS

The performance of training and testing dataset has been shown in Fig. 3 and 4 respectively. The value of R is close to one for prediction of E and UCS also. Therefore, the developed GPR proves his capability for prediction of E and UCS.

For RVM, the trial and error approach has been adopted for determination of design value of σ. The developed RVM gives best performance at σ=0.2 for prediction of E. The performance of training dataset has been shown in Fig. 1. Fig. 2 shows the performance of testing dataset. It is observed from Fig. 1 and 2 that the value of R is close to one for training as well as testing dataset. Therefore, the developed RVM proves his credibility for prediction of E. The developed RVM gives the following equation for prediction of E.

$$E = \sum_{i=1}^{23} w_i \exp\left[-\frac{(x_i - x)(x_i - x)^T}{0.08} \right] \tag{22}$$

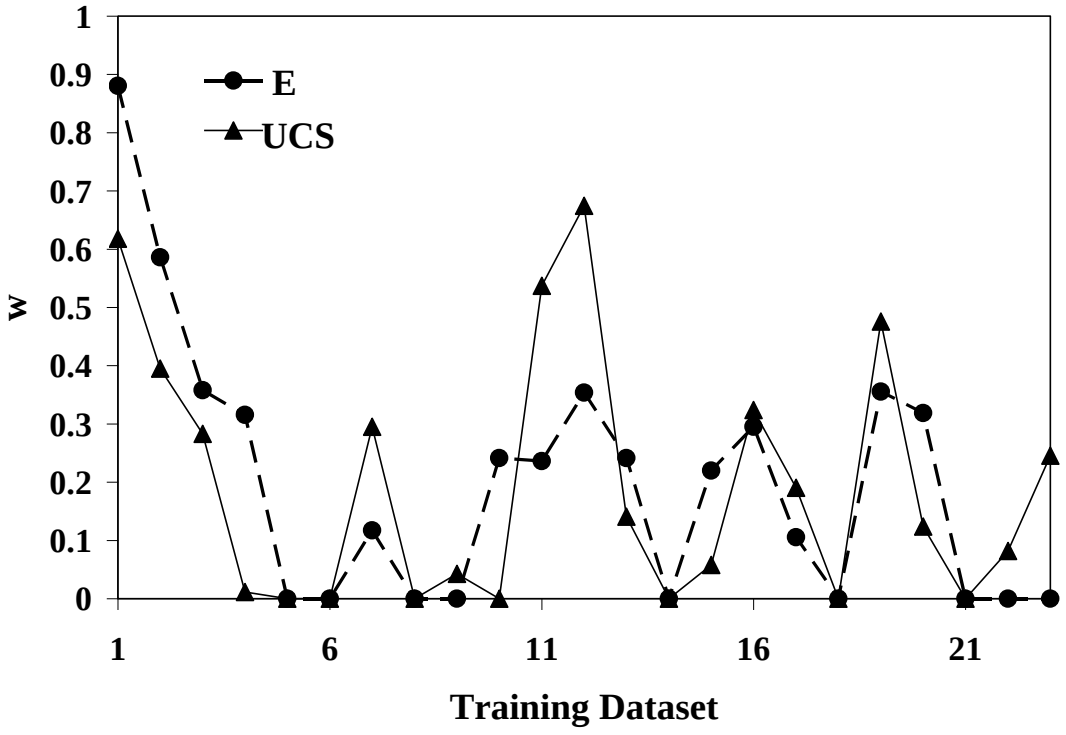

Fig. 5. Values of w

The value of w has been shown in Fig. 5. It is clear from Fig. 5 that 14 training datasets have non-zero w. So, the developed RVM produces 14 relevance vectors. For prediction of UCS, the design value of σ=0.42. The performance of training

and testing has been illustrated in Fig. 3 and 4 respectively. The following equation has been presented based on the developed RVM.

$$UCS = \sum_{i=1}^{23} w_i \exp\left[-\frac{(x_i - x)(x_i - x)^T}{0.3528} \right] \quad (23)$$

Table 2: Comparison between the developed models

Output	GRNN	FANN	GPR	RVM	MPMR
UCS(MPa)	0.927	0.964	0.984	0.992	0.914
E(GPa)	0.877	0.905	0.961	0.964	0.936

Fig. 5 illustrates the value of w. For prediction of UCS, the developed RVM creates 16 relevance vectors.

For prediction of E by using the MPMR model, the design values of ε and σ have been computed by trial and error approach. The design values of ε and σ are 0.003 and 0.1 respectively. The performance of training and testing has depicted in Fig. 1 and 2 respectively. For determination of UCS, the developed MPMR gives best performance at ε=0.001 and σ=0.3. Fig. 3 and 4 shows the performance of training and testing dataset respectively. The value of R is close to one for training as well as testing dataset. So, the developed MPMR predicts E and UCS reasonably well.

The developed GPR, RVM, and MPMR have been compared with the ANN models developed by Dehghan et al.[12]. Two ANN models {Generalized Regression Neural Network (GRNN) and Feed Forward Network (FANN)} were developed. Table 2 shows the comparison. The performance of RVM is slightly better than the ANN models. The developed GPR assumes Gaussian data distribution. However, MPRM, ANN and RVM do not assume any data distribution. The developed MPMR and GPR use only two tuning parameters. Whereas, RVM uses only one tuning parameter. ANN uses many tuning parameters (number of hidden layers, number of neurons, transfer function, number of epochs, etc.). For GPR, RVM and MPMR, the performance of training and testing dataset is almost same. So, the developed GPR, RVM and MPMR do not show any overtraining. Therefore, the developed GPR, RVM and MPMR show good generalization capability. In RVM, maximization of the type-II likelihood reduces the chance of overfitting. ANN has no control over future prediction. The developed MPMR has control over future prediction. In MPMR, the optimization problem is convex. So, the solution of this optimization problem is always global minimum. In ANN, the optimization is not always convex.

Therefore, the solution is not always global minimum. GPR and RVM do not involve any optimization problem.

4 Conclusion

This study describes GPR, RVM, and MPMR for prediction of E and UCS of Travertine samples. 30 datasets have been utilized to construct the model. There models (GPR, RVM and MPMR) produce excellent performance. The developed RVM gives equations for prediction of E and UCS. MPMR maximizes the minimum probability of future predictions being within some bound of the true regression function. A comparative study has been presented between the developed models and the ANN models. This study shows that the developed models can be adopted to model different problems in engineering.

References

[1] D'Andrea D.V, Fisher R.L, Fogelson D.E. 1964. Prediction of compression strength from other rock properties. Colo Sch Mines Q, 59(4B), 623.

[2] Inoue M., Ohomi M. 1981. Relation between uniaxial compressive strength and elastic wave velocity of soft rock.In: Proceedings of the International Symposium on Weak Rock. Tokyo.

[3] Cargill J.S., Shakoor A. 1990. Evaluation of empirical methods for measuring the uniaxial compressive strength of rock. *International Journal of Rock Mechanics, Mining Science and Geomechanical Abstracts*, 27(6), 495.

[4] Shakoor A., Bonelli R.E. 1991. Relationship between petrographic characteristics, engineering index properties and mechanical properties of selected sandstone. Bulletin of the Association of Engineering Geologists XXVIII, 1, 55.

[5] Deere D.U., Miller R.P. 1966. Engineering classification and index properties for intact rock. Air Force Weapons Lab, Tech Report, AFWL-TR, 65.

[6] Grasso P., Xu S. 1992. Mahtab A. Problems and promises of index testing of rock. Rock Mechanics, 879.

[7] Chau K.T., Wong R.H.C. 1996. Uniaxial compressive strength and point load strength. *Int J Rock Mech Min Sci,* 33,183.

[8] Kahraman S. 2001. Evaluation of simple methods for assessing the uniaxial compressive strength of rock. *Int J Rock Mech Min Sci,* 38, 981.

[9] Yilmaz I., Sendir H. 2002. Correlation of schmidt hardness with unconfined compressive strength and young modulus in gypsum from Sivas (Turkey). Engineering Geology, 66, 211.

[10] Tsiambaos G., Sabatakakis N. 2004. Considerations on strength of intact sedimentary rocks. Eng Geol, 72, 261.

[11] Fener M., Kahraman S., Bilgil A., Gunaydin O. 2005. A comparative evaluation of indirect methods to estimate the compressive strength of rocks. Rock Mech Rock Eggnog, 38(4), 329.

[12] Dehghan S., Sattari Gh., Chehreh Chelgani S., Aliabadi M.A. 2010. Prediction of uniaxial compressive strength and modulus of elasticity for travertine samples using regression and artificial neural networks. Mining Science and Technology, 20, 41–46.

[13] Park D., Rilett L.R. 1999. Forecasting freeway link ravel times with a multi-layer feed forward neural network. Computer Aided Civil And Znfa Structure Engineering, 14, 358-367.

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

[15] Neal R.M. 1997. Monte Carlo implementation of Gaussian process models for Bayesian regression and classification. Technical Report No.9702, Department of Statistics, University of Toronto, Canada.

[16] MacKay D.J.C. 1988. Introduction to Gaussian processes. In: Neural Networks and Machine Learning (C. M. Bishop, Ed.), Springer, 133-165.

[17] Yi G., Shi J.Q., Choi T. 2011. Penalized Gaussian process regression and classification for high-dimensional nonlinear data, Biometrics, 67(4), 1285-1294.

[18] Clifton L., Clifton D.A., Pimentel M.A.F., Watkinson P.J., Tarassenko L. 2012. Gaussian process regression in vital-sign early warning systems, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Art. No. 6347400, 6161-6164.

[19] Chen J., Chan L.L.T., Cheng Y.-C. 2013. Gaussian process regression based optimal design of combustion systems using flame images, Applied Energy, 111, 153-160.

[20] Tipping ME. 2000. The relevance vector machine. Advances in Neural Information Processing Systems, 12, 625-658.

[21] Tipping M.E. 2001. Sparse Bayesian learning and the relevance vector machine, *J. Mach. Learn.*, 1, 211-244.

[22] Han M., Sun L.-L., Hong X.-J., Han J. 2011. Automatic classification of epileptic EEG signals based on AR model and relevance vector machine, Chinese Journal of Biomedical Engineering, 30(6), 864-870.

[23] Wang T., Li Y., Wang H., Wang P. 2012. A RVM reliability analysis model for wireless sensor of networked synchronization control system, Journal of Information and Computational Science, 9(17), 5345-5351.

[24] Zhou Z., Wang Z., Sun X. 2013. Face recognition based on optimal kernel minimax probability machine, Journal of Theoretical and Applied Information Technology, 48(3), 1645-1651.

[25] Strohmann T.R., Grudic, G.Z. 2002. A Formulation for minimax probability machine regression. In Advances in Neural Information Processing Systems (NIPS) 14, Dietterich T.G., Becker S., Ghahramani Z., editors, Cambridge, MA, MIT Press.

[26] Lanckriet G.R.G., Ghaoui L.E., Bhattacharyya C., Jordan M.I. 2002. Minimax probability machine. In Dietterich T.G., Becker S., Ghahramani Z., editors, Advances in Neural Information Processing Systems 14, Cambridge, MA, MIT Press.

[27] Yang L., Wang L., Sun Y., Zhang R. 2010. Simultaneous feature selection and classification via Minimax Probability Machine, International Journal of Computational Intelligence Systems, 3(6), 754-760.

[28] Zhao H.-C. 2013. Research on mechanical mechanics with fault diagnosis method for gearbox based on relevance vector machine, Advanced Materials Research 703, 208-211.

[29] Rasmussen C.E., Williams C.K. 2006. Gaussian Processes for Machine Learning. Massachusetts Institute of Technology, MIT-Press.