Neural Comput & Applic
DOI 10.1007/s00521-012-1043-x

ORIGINAL ARTICLE

Least square support vector machine and multivariate adaptive regression spline for modeling lateral load capacity of piles

Pijush Samui · Dookie Kim

Received: 29 September 2011 / Accepted: 22 June 2012

Abstract This article adopts least square support vector machine (LSSVM) and multivariate adaptive regression spline (MARS) for prediction of lateral load capacity (Q) of pile foundation. LSSVM is firmly based on the theory of statistical learning, uses regression technique. MARS is a nonparametric regression technique that models complex relationships. Diameter of pile (D), depth of pile embedment (L), eccentricity of load (e), and undrained shear strength of soil (S_u) have been used as input parameters of LSSVM and MARS. Equations have been presented from the developed MARS and LSSVM. This study also presents a comparative study between the developed MARS and LSSVM.

Keywords Least square support vector machine · Multivariate adaptive regression spline · Soft computing · Prediction · Pile foundation

1 Introduction

Piles are structural members of timber, concrete, and/or steel that are used to transmit surface loads to lower levels in the soil mass. The determination of lateral load capacity (Q) of pile is an imperative task in geotechnical engineering. Geotechnical engineers use the different methods [1–6] for the determination of Q of pile. However, every method has own limitation. Geotechnical engineers use different soft computing techniques such as artificial neural network (ANN) [7, 8], Genetic Programming (GP) [9–11]. This study adopts least square support vector machine (LSSVM) and multivariate adaptive regression spline (MARS) for prediction of Q of pile foundation. This study uses the database collected by Rao and Kumar [12]. LSSVM is a modified version of support vector machine (SVM) proposed by Suykens and Vandewalle [13]. The formulation of LSSVM is related to regularization networks [14, 15]. It involves equality instead of inequality constraints and works with a least squares cost function. The introduction of equality constraint simplifies the problem in such a way that the solution is characterized by a linear system, more precisely a KKT (Karush–Kuhn–Tucker) system [16]. MARS is a nonparametric regression procedure that makes no specific assumption about the underlying functional relationship between the output and input variables [17]. It is an adaptive procedure because the selection of basis functions depends on the data set and problem. This article has the following aims:

- To examine the capability of LSSVM and MARS for prediction of Q of pile foundation
- To develop equations for the determination of Q of pile foundation based on the LSSVM and MARS
- To carry out a comparative study between the developed LSSVM, MARS, and other available methods for prediction of Q of pile foundation.

P. Samui (✉)
Centre for Disaster Mitigation and Management,
VIT University, Vellore 632014, India
e-mail: pijush.phd@gmail.com

D. Kim
Department of Civil Engineering, Kunsan National University,
Kunsan, Jeonbuk, South Korea
e-mail: kim2kie@chol.com

2 Details of LSSVM

The basic formulation of the standard LSSVM [18] for function estimation is briefly described in this section.

Published online: 06 July 2012

Consider a given training set of N data points $\{x_k, y_k\}_{k=1}^{N}$ with input data $x_k \in R^N$ and output $y_k \in r$, where R^N is the N-dimensional vector space and r is the one-dimensional vector space. This study uses diameter of pile (D), depth of pile embedment (L), eccentricity of load (e), and undrained shear strength of soil (S_u) as input parameters of the LSSVM. The output of LSSVM is Q.

So, $x = [D,\ L,\ e,\ S_u]$ and $y = [Q]$. The following regression model is used.

$$y(x) = w^T \varphi(x) + b, \tag{1}$$

where the nonlinear mapping $\varphi(.)$ maps the input data into a higher dimensional feature space; $w \in R^n$; $b \in r$; w = an adjustable weight vector; b = the scalar threshold.

The following optimization problem is formulated:

$$\text{Minimize}: \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2$$

$$\text{Subject to: } y(x) = w^T \varphi(x_k) + b + e_k, \quad k = 1, \ldots, N. \tag{2}$$

Lagrange function is adopted to solve the above optimization problem.

$$L(w, b, e; \alpha) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i \{ w^T \varphi(x_i) + b + e_i - y_i \} \tag{3}$$

with α_k Lagrange multipliers. The solution of the above Eq. (3) can be obtained by partially differentiating with respect to each variable

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{k=1}^{N} \alpha_k \varphi(x_k)$$
$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{k=1}^{N} \alpha_k = 0$$
$$\frac{\partial L}{\partial e_k} = 0 \Rightarrow \alpha_k = \gamma e_k, \quad k = 1, \ldots, N.$$
$$\frac{\partial L}{\partial \alpha_k} = 0 \Rightarrow w^T \varphi(x_k) + b + e_k - y_k = 0, \quad k = 1, \ldots, N. \tag{4}$$

When the variables w and e are removed, the equation can be rewritten as a linear function group

$$\begin{bmatrix} 0 & 1^T \\ 1 & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{5}$$

where $y = y_1, \ldots, y_N$, $1 = [1, \ldots, 1]$, $\alpha = [\alpha_1, \ldots, \alpha_N]$ and Mercer's theorem [15, 19], is applied within the Ω matrix,

$$\Omega = \varphi(x_k)^T \varphi(x_l) = k(x_k, x_l), \quad k, l = 1, \ldots, N$$

where $k(x_k,\ x_l)$ is the kernel function. Choosing $\gamma > 0$, ensures the matrix $\Phi = \begin{bmatrix} 0 & 1^T \\ 1 & \Omega + \gamma^{-1} I \end{bmatrix}$ is invertible. Then, the analytical of α and b is given by

$$\begin{bmatrix} b \\ \alpha \end{bmatrix} = \Phi^{-1} \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{6}$$

Radial basis function has been used as a kernel in this analysis. Radial basis function is given by

$$K(x_k, x_l) = \exp\left\{ -\frac{(x_k - x_l)(x_k - x_l)^T}{2\sigma^2} \right\} \quad k, l = 1, \ldots, N \tag{7}$$

where σ is the width of radial basis function.

The resulting LSSVM model for Q prediction becomes then

$$Q = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{8}$$

The above described LSSVM has been adopted for prediction of Q of pile foundation. The data set contains information about D, L, e, S_u, and Q. Table 1 shows the different statistical parameters of the data sets. The data set are divided into the following two groups:

Training Data set: This is required to develop the LSSVM model. 31 out of 41 data sets have been used as training data set.

Testing Data set: This is required to determine the performance of the developed LSSVM model. The remaining 13 data sets have been used as testing data set.

The data are normalized between 0 and 1. The design values of γ and σ have been determined by trial-and-error approach during the training of LSSVM. The program of LSSVM has been constructed by using MATLAB.

Table 1 Basis function and their corresponding equations

Basis function	Equation
BF1	max (0, $e - 0.783$)
BF2	max (0, $0.783 - e$)
BF3	BF1*max (0, $L - 0.31$)
BF4	BF1*max (0, $0.31 - L$)
BF5	BF1*max (0, $D - 0.576$)
BF6	BF1*max (0, $0.576 - D$)
BF7	BF1*max (0, $S_u - 0.25$)
BF8	max (0, $S_u - 0.7$)
BF9	max ($0.7 - S_u$)

3 Details of MARS

The algorithm of MARS was developed by Friedman [17]. It uses expansions in piecewise linear basis functions of the following form.

$$c^{+}(x,\tau) = [+(x-\tau)]_{+}, c^{-}(x,\tau) = [-(x-\tau)]_{+} \quad (9)$$

where $[q]_{+} := \max\{0, q\}$ and τ is an univariate knot. Each function is piecewise linear, with a knot at the value τ, and it is called a reflected pair.

The relation between input and output can be written in the following way:

$$Y = f(X) + \varepsilon \quad (10)$$

where Y is a output variable, $X = (X_1, X_2, \ldots, X_p)^T$ is a vector of predictors, and ε is an additive stochastic component, which is assumed to have zero mean and finite variance.

In MARS, the above equation is written in the following way:

$$Y = \theta_0 + \sum_{m=1}^{M} \theta_m \psi_m(X) + \varepsilon \quad (11)$$

where ψ_m ($m = 1, 2, \ldots, M$) are basis functions from or products of two or more such functions, ψ_m is taken from a set of M linearly independent basis elements, and θ_m are the unknown coefficients for the mth basis function ($m = 1, 2, \ldots, M$) or for the constant 1 ($m = 0$). In this study, $X = [D, L, e, S_u]$ and $Y = [Q]$. The expression of mth basis function is given below:

$$\psi_m(x) := \prod_{j=1}^{K_m} \left[s_{k_j^m} \cdot \left(x_{k_j^m} - \tau_{k_j^m} \right) \right]_{+} \quad (12)$$

where K_m is the number of truncated linear functions multiplied in the mth basis function, $x_{k_j^m}$ is the input variable corresponding to the jth truncated linear function in the mth basis function and $\tau_{k_j^m}$ is the knot value corresponding to the variable $x_{k_j^m}$, and $s_{k_j^m}$ is the selected sign +1 or −1.

MARS algorithm contains the following steps [17]:

A forward process: In forward process, basis functions are selected to define Eq. (11).

A backward process: In backward process, the ineffective basis functions are deleted from model. Basis functions are deleted based on the generalized cross-validation (GCV) criterion Craven and Wahba [20]. The GCV criterion is defined in the following way:

$$\mathrm{GCV} = \frac{\frac{1}{N}\sum_{i=1}^{N} [y_i - f(x_i)]^2}{\left[1 - \frac{C(B)}{N}\right]^2} \quad (13)$$

where N is the number of data and $C(B)$ is a complexity penalty that increases with the number of BF in the model and which is defined as:

$$C(B) = (B+1) + dB \quad (14)$$

where d is a penalty for each basis function (BF) included into the model. It can be also regarded as a smoothing parameter. Friedman [17] provided more details about the selection of the d.

The details of MARS are given by Friedman [17]. The above described MARS has been used to determine Q of pile foundation. The same training data set, testing data set, and normalization technique have been used in the MARS as used in the LSSVM model. The program of MARS has been constructed by using MATLAB.

4 Results and discussion

The performance of the developed LSSVM and MARS has been accessed by the coefficient of correlation (R) value. The value of R has been computed by using the following equation:

$$R = \frac{\sum_{i=1}^{n} \left(Q_{ai} - \overline{Q}_a\right)\left(Q_{pi} - \overline{Q}_p\right)}{\sqrt{\sum_{i=1}^{n} \left(Q_{ai} - \overline{Q}_a\right)} \sqrt{\sum_{i=1}^{n} \left(Q_{pi} - \overline{Q}_p\right)}} \quad (15)$$

where Q_{ai} and Q_{pi} are the actual and predicted Q values, respectively, $\overline{Q}_a$ and $\overline{Q}_p$ are mean of actual and predicted Q values corresponding to n patterns. We have used different kernel functions (radial basis function, polynomial, linear, spline, etc.). However, the performance of radial basis function is best. For good model, the value of R is close to one. For LSSVM, the design values of γ and σ are 70 and 4, respectively. The performance of the LSSVM has been depicted in Fig. 1. It is observed from Fig. 1 that the value of R is close to one for training as well as testing data set. Therefore, the developed LSSVM predicts Q reasonable well. The developed LSSVM gives the following equation (by putting $K(x_i, x) = \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right\}$, $\sigma = 4$, $N = 31$ and $b = 0.8442$ in Eq. 8) for prediction of Q of pile foundation.

$$Q = \sum_{k=1}^{31} \alpha_k \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{32}\right\} + 0.8442 \quad (16)$$

Figure 2 shows the values of α.

For the MARS model, 15 basis functions have been used in forward process. However, six basis functions have been deleted in backward process. So, the final model contains 9 basis functions. The developed MARS gives the following equation for prediction of Q.

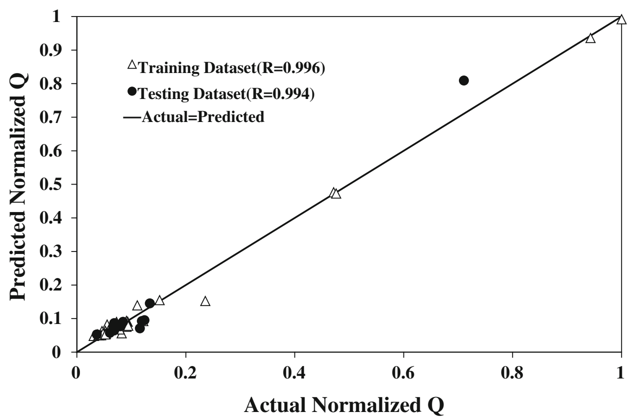

Fig. 1 Performance of the LSSVM model

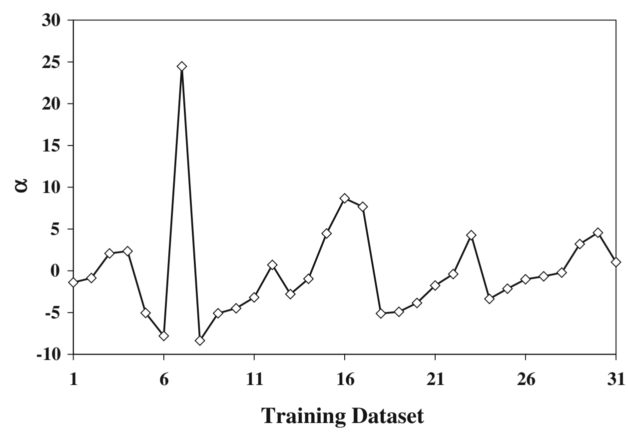

Fig. 2 Values of α for the LSSVM model

$$Q = 0.115 + 0.397 * \mathrm{BF1} - 0.042 * \mathrm{BF2} + 1.643 * \mathrm{BF3} - 6.250 * \mathrm{BF4} - 0.651 * \mathrm{BF5} + 3.103 * \mathrm{BF6} + 1.485 * \mathrm{BF7} - 0.154 * \mathrm{BF8} - 0.115 * \mathrm{BF9} \quad (17)$$

Table 1 shows equation of the different basis functions.

The performance of training and testing data set has been determined by the above Eq. (17). Figure 3 depicts the performance of the training and testing data set for the MARS. It is clear from Fig. 3 that the value of R is close to one for training as well as testing data set. So, the developed MARS has capability for prediction of Q of pile foundation.

A comparative study has been carried out between the developed LSSVM, MARS, and the other available methods [1, 3]. Hansen [1] proposed the following equation for prediction of Q.

$$Q = (\gamma x K_q + cK_c)MD \quad (18)$$

where M is an empirical modification factor = 0.85 (dimensionless); D is the pile width or diameter (length units); γ is the moist unit weight of foundation soil (force per volume units); x is the depth measured from the ground surface (length units); c is the cohesion of the foundation soil (force per area units); K_q is a coefficient for the frictional component of net soil resistance under 3D conditions (dimensionless); and K_c is a coefficient for the cohesive component of net soil resistance under 3D conditions (dimensionless). Broms [3] developed the following equation for prediction of Q

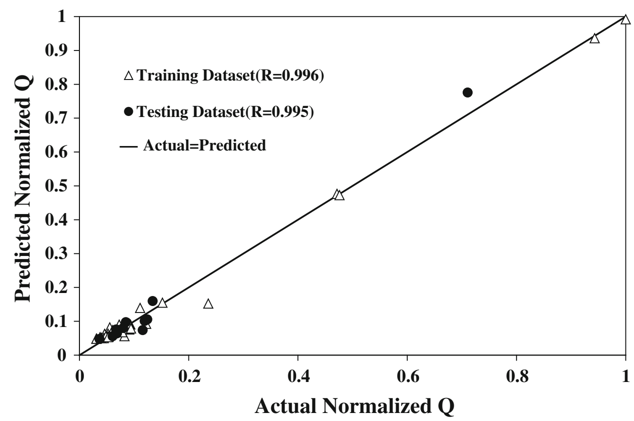

Fig. 3 Performance of the MARS model

$$Q = 9fS_{\mathrm{u}}D \quad (19)$$

where f is required depth of soil with S_{u}.

Comparison has been made in terms root mean square error (RMSE) and mean absolute error (MAE). The values of RMSE and MAE have been determined by using the following relation.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (Q_{ai} - Q_{pi})^2}{n}} \quad (20)$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left|Q_{ai} - Q_{pi}\right|}{n} \quad (21)$$

where Q_{ai} and Q_{pi} are the actual and predicted Q values, respectively, and n is the number of data.

Figure 4 shows the bar chart of RMSE and MAE for the different models. It is observed from Fig. 4 that the performance of MARS is better than the other models. The developed MARS does not require any prior assumption regarding the statistical distribution of the data. The developed LSSVM and MARS have been compared with the ANN model developed by Das and Basudhar [7]. Comparison has been made in terms of R value. The reason is that [7] did not give RMSE and MAE value for the ANN model. Figure 5 shows the bar chart of R value for the ANN, LSSVM, and MARS. It is observed from Fig. 5 that the performance of LSSVM and MARS is better than the ANN model.

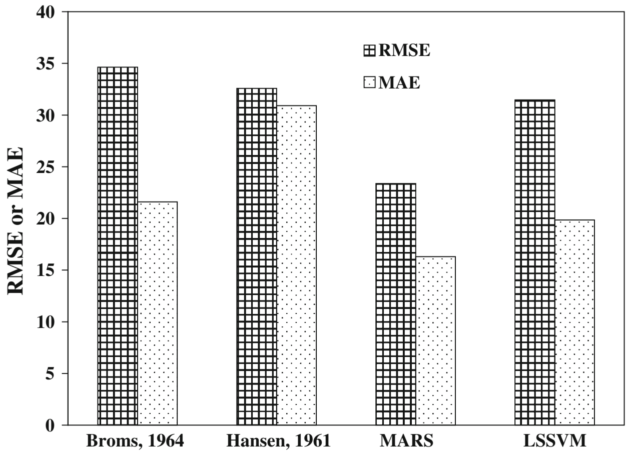

Fig. 4 Values of RMSE and MAE of the different models

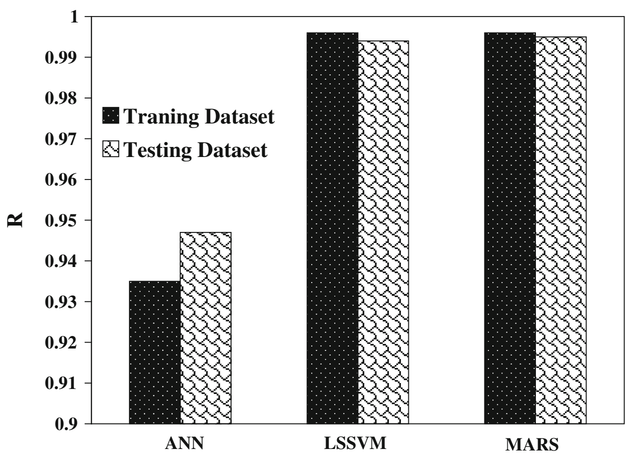

Fig. 5 Comparison between ANN, LSSVM, and RVM

5 Conclusion

This article presents two soft computing techniques (LSSVM and MARS) for prediction of Q of pile foundation. The main aim of this study is the development of equations for the determination of Q of pile foundation. 41 data have been used to develop MARS and LSSVM models. The performance of the LSSVM and MARS is encouraging. User can use the developed equations for the determination of Q of pile foundation. The developed equations are new contribution to the literature of the modeling of lateral load capacity of pile foundation. The developed LSSVM and MARS outperform the other available models. Based on the results, the developed LSSVM and MARS seem to be interesting tools for prediction of Q of pile foundation.

References

1. Hansen B (1961) The ultimate resistance of rigid piles against transversal force. Danish Geotechnical Institute, Copenhagen, Bulletin No. 12, pp 5–9
2. Matlock H, Reese LC (1962) Generalized solutions for laterally loaded piles. Trans ASCE 127:1220–1248
3. Broms BB (1964) Lateral resistance of piles in cohesive soils. J Soil Mech Found Eng ASCE 90(SM2):27–63
4. Meyerhof GG (1976) Bearing capacity and settlement of pile foundations. J Geotech Eng ASCE 102(3):196–228
5. Poulos HG, Davis EH (1980) Pile foundation analysis and design. Wiley, New York
6. Portugal JC, Seco e Pinto PS (1993) Analysis and design of pile under lateral loads. In: Proceedings of the 11th international geotechnical seminar on deep foundation on bored and auger piles, Belgium, pp 309–313
7. Das SK, Basudhar PK (2006) Undrained lateral load capacity of piles in clay using artificial neural network. Comput Geotech 33(8):454–459
8. Alavi AH, Gandomi AH (2011) Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing. Comput Struct. doi: 10.1016/j-compstruc-2011-08-019
9. Gandomi AH, Alavi AH (2011) Multi-stage genetic programming: a new strategy to nonlinear system modeling. Inf Sci. doi:10.1016/j-ins-2011-07-026
10. Gandomi AH, Alavi AH (2012) A new multi-gene genetic programming approach to nonlinear system modeling. Part II: geotechnical and earthquake engineering problems. Neural Comput Appl 21(1):189–201
11. Gandomi AH, Alavi AH (2012) A new multi-gene genetic programming approach to nonlinear system modeling. Part I: materials and structural engineering problems. Neural Comput Appl 21(1):171–187
12. Rao KM, Suresh Kumar V (1996) Measured and predicted response of laterally loaded piles. In: Proceedings of the sixth international conference and exhibition on piling and deep foundations, India, pp 1.6.1–1.6.7
13. Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9:293–300
14. Evgeniou T, Pontil M, Poggio T (2001) Regularization networks and support vector machines. Adv Comput Math 13:1–50
15. Smola A, Schölkopf B, Müller KR (1998) The connection between regularization operators and support vector kernels. Neural Netw 11:637–649
16. Fletcher R (1987) Practical methods of optimization. Wiley, Chichester, NY
17. Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19(1):1–141
18. Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J (2002) Least squares support vector machines. World Scientific, Singapore
19. Vapnik VN (1998) Statistical learning theory. Wiley, New York
20. Craven P, Wahba G (1979) Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. Numer Math 31:317–403