

Article

A Multivariate Adaptive Regression Spline Approach for Prediction of Maximum Shear Modulus (G_{\max}) and Minimum Damping Ratio (ξ_{\min})

Pijush Samui^{1*} and Dwarkadas Pralhaddas Kothari²

¹ Centre for Disaster Mitigation and Management, VIT University, Vellore-632014, India

² Rasoni Group Of Institutions, Nagpur, India

* E-mail: pijush.phd@gmail.com

Abstract. This study uses Multivariate Adaptive Regression Spline (MARS) for determination Maximum Shear Modulus (G_{\max}) and Minimum Damping Ratio (ξ_{\min}) of synthetic reinforced soil. MARS employs confining pressure (σ , psi), rubber (r , %) and sand (s , %) as input variables. The output of the MARS is G_{\max} and ξ_{\min} . The developed MARS gives equations for determination of G_{\max} and ξ_{\min} . The results of MARS have been compared with the adaptive neuro-fuzzy inference system (ANFIS), multi-layer perceptron (MLP) and multiple regression analysis method (MRM). A sensitivity analysis has been also carried out to determine the effect of each input variable on G_{\max} and ξ_{\min} . This study shows that the developed MARS is a robust model for prediction of G_{\max} and ξ_{\min} .

Keywords: Maximum shear modulus, minimum damping ratio, multivariate adaptive regression spline, prediction, adaptive neuro-fuzzy inference system, multi-layer perceptron.

ENGINEERING JOURNAL Volume 16 Issue 5

Received 27 March 2012

Accepted 27 May 2012

Published 1 October 2012

Online at <http://www.engj.org/>

DOI:10.4186/ej.2012.16.5.69

1. Introduction

Geotechnical engineers use synthetic reinforced soil for different purposes such as pavement structures, canal lining, erosion control, slope stabilization, piles, walls, liquefaction, etc [1-5]. So, the determination of different dynamic properties {Maximum Shear Modulus (G_{\max}) and Minimum Damping Ratio (ξ_{\min})} of synthetic reinforced soil is an imperative task in geotechnical earthquake engineering. Laboratory determination of dynamic properties is a tedious and time consuming task [6]. Recently, Akbulut *et al.* [6] successfully used Adaptive Neuro-Fuzzy Inference (ANFIS) for determination of G_{\max} and ξ_{\min} of synthetic reinforced soil. However, the developed ANFIS has low generalization capability. It also did not give equations for determination of G_{\max} and ξ_{\min} .

This article adopts an alternative method based on Multivariate Adaptive Regression Spline (MARS) for determination of G_{\max} and ξ_{\min} of synthetic reinforced soil. MARS is a non-parametric adaptive regression procedure [7]. It can be considered as a generalisation of classification and regression trees (CART) [8]. It uses a lot of piecewise regression equations in the model. It has been successfully used for solving different problems in engineering [9-16]. This article uses the database collected by [6]. The results of MARS have been compared with the models developed by [6]. This study gives equations for prediction of G_{\max} and ξ_{\min} of synthetic reinforced soil. This paper is structured as follows: Section 2 describes the MARS model; Section 3 discusses the main results; finally Section 4 contains the conclusion of the paper.

2. Details of MARS

MARS divides the whole space of input variable into various sub-regions. It defines a different mathematical equation for each area. This equation relates each sub-region of input variable to the output variable. MARS uses the following two-sided truncated power functions as spline basis functions [17].

$$[-(x-t)]_+^q = \begin{cases} (t-x)^q, & \text{if } x < t \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$[+(x-t)]_+^q = \begin{cases} (x-t)^q, & \text{if } x \geq t \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where q is the power and t is knot.

The final MARS model has the following form:

$$\hat{y} = \hat{f}(x) = a_0 + \sum_{m=1}^M a_m B_m(x) \quad (3)$$

where y is the output variable, x is the input variable, a_0 is the coefficient of the constant term, M is the number of spline functions, and B_m and a_m is the m^{th} spline function and its coefficient [7] respectively. Confining pressure (σ , psi), rubber (r , %) and sand (s , %) have used as input of the MARS. The output of the MARS is G_{\max} and ξ_{\min} . So, $x = [\sigma, r, s]$ and $y = [G_{\max}, \xi_{\min}]$

MARS uses the following two steps:

Forward Algorithm: Basis functions are introduced to define Eq. (3). Many basis functions are added in Eq. (3) to get better performance. The developed MARS can show overfitting problem due to large number of basis functions.

Backward Algorithm: For preventing overfitting, redundant basis functions are deleted from Eq. (3). MARS adopts Generalised Cross-Validation (GCV) to delete the redundant basis functions [18]. The expression of GCV is given below:

$$GCV = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}(x_i)]^2}{\left[1 - \frac{C(B)}{N}\right]^2} \quad (4)$$

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

$$C(B) = (B+1) + dB \quad (5)$$

where d is a penalty for each BF included into the model and B is number of basis functions in Eq. (3). The details about d are given by [7].

3. Details of Present Analysis

This article employs the above MARS for prediction of G_{\max} and ξ_{\min} . The same training and testing dataset have been used as used by [6]. The data is normalized between 0 and one. Table 1 shows the different statistical parameters of the dataset.

Table 1. Statistical parameter of the dataset.

Variable	Mean	Standard Deviation	Skewness	Kurtosis
σ	16.66	10.52	0.23	1.50
s	80.00	6.27	0	2.16
r	10	6.27	0	2.16
ξ_{\min}	3.60	1.20	-0.01	1.97
G_{\max}	192.96	63.57	-0.32	2.50

A sensitivity analysis has been done to extract the cause and effect relationship between the inputs and outputs of the MARS model. The basic idea is that each input of the model is offset slightly and the corresponding change in the output is reported. The procedure has been taken from the work of [19]. According to [19], the sensitivity (S) of each input parameter has been calculated by the following formula:

$$S(\%) = \frac{1}{N} \sum_{j=1}^N \left(\frac{\% \text{ change in output}}{\% \text{ change in input}} \right)_j \times 100 \quad (6)$$

where N is the number of data points. The analysis has been carried out on the trained model by varying each of input parameter, one at a time, at a constant rate of 30%. The program of MARS has been constructed by using MATLAB.

4. Results and Discussion

The performance of the developed MARS has been accessed in terms of Coefficient of Determination (R^2). For good model, the value of R^2 should be close to one. Figure 1 shows the flow chart of the MARS.

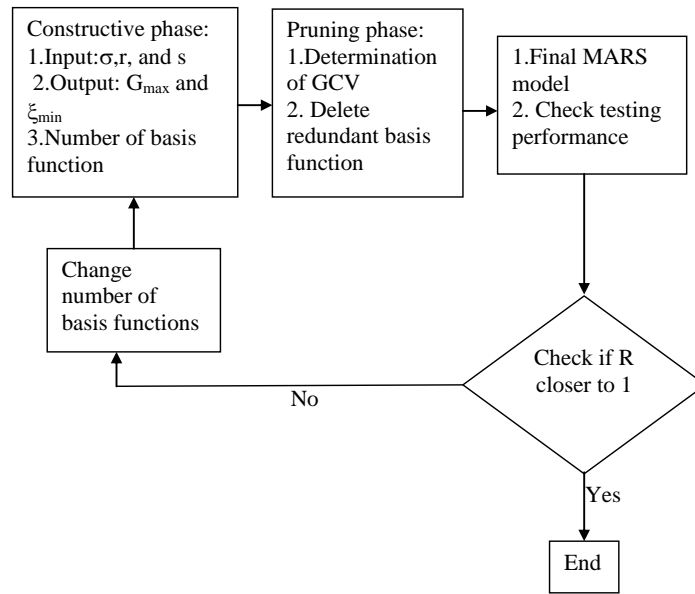


Fig 1. Flow chart of the MARS model.

Figure 2 depicts the effect of number of basis function on the testing performance. It is observed from Fig. 2 that 12 basis functions give best performance for prediction of G_{max} .

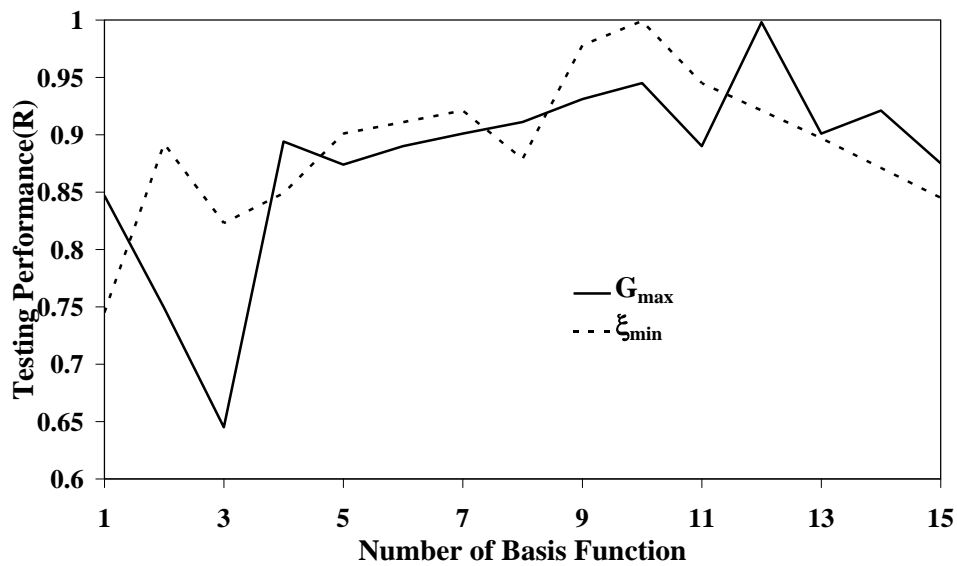


Fig 2. Effect of number basis function on the testing performance.

So, 12 basis functions have been introduced in forward algorithm. However, 5 basis functions have been deleted in backward algorithm. So, the final MARS model includes 7 basis functions. The expression of the final MARS model is given below (by putting $y=G_{max}$, $a_0=0.174$ and $M=7$ in Eq. (3)):

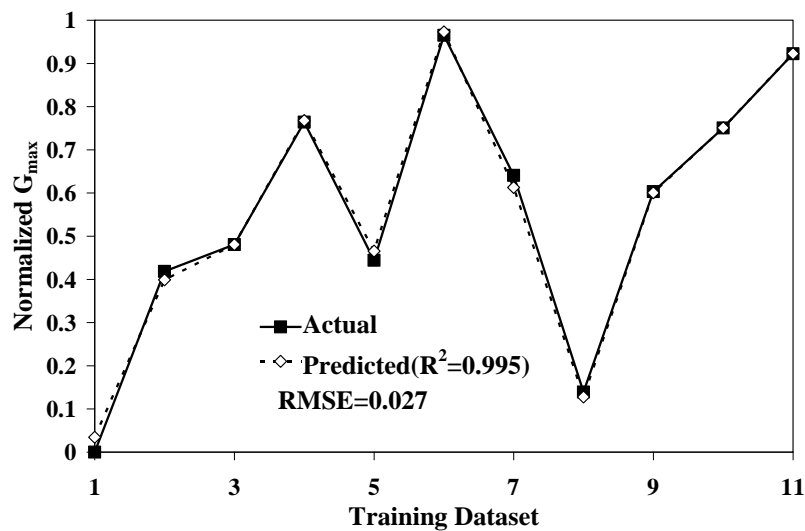
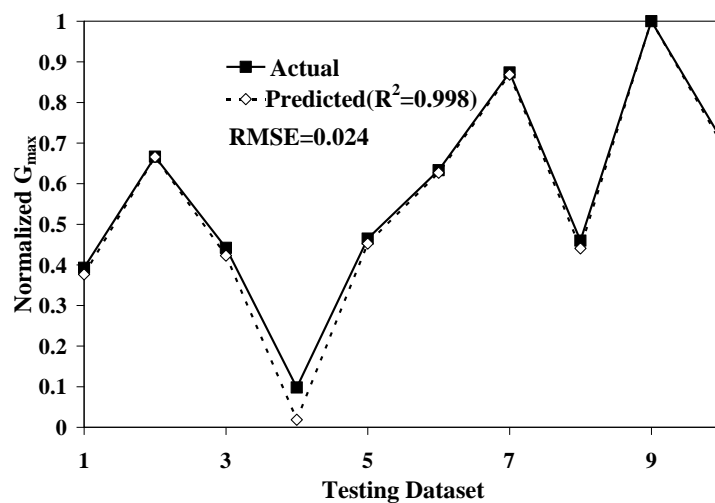
$$G_{max} = 0.174 + \sum_{m=1}^7 a_m B_m(x) \tag{7}$$

The expression of B_m and corresponding coefficients (a_m) is given in Table 2.

Table 2. Basis functions and their corresponding coefficient (a_m) for G_{\max} .

Basis Functions	Equation	Coefficient (a_m)
$B_1(x)$	$\max(0, r - 0.0343)$	0.761
$B_2(x)$	$\max(0, 0.0343 - r)$	-1.547
$B_3(x)$	$\max(0, 0.059 - \sigma)$	4.749
$B_4(x)$	$\max(0, s - 0.545)$	1.498
$B_5(x)$	$\max(0, 0.545 - s)$	-0.334
$B_6(x)$	$B_4(x) * \max(0, 0.545 - s)$	10.897
$B_7(x)$	$\max(0, \sigma - 0.059) * \max(0, s - 0.245)$	-4.729

The above Eq. (7) has been adopted for determination of the performance of training and testing dataset. Figure 3 illustrates the performance of the training dataset. As far as the MARS model training is concerned, the developed MARS has successfully captured the input and output relationship. The performance of testing dataset has been shown in Fig. 4.

Fig. 3. Performance of the training dataset for G_{\max} .Fig. 4. Performance of the testing dataset for G_{\max} .

It is observed from Figs. 3 and 4 that the value of R^2 is close to one. Therefore, the developed MARS shows good predictive ability for prediction of G_{\max} .

It is clear from Fig. 2 that 10 basis functions give best performance for prediction of ξ_{\min} . So, 10 basis functions have been introduced in forward algorithm. However, the final MARS model contains 6 basis functions. So, 4 basis functions have been deleted in backward algorithm. The expression of the final MARS model is given below (by putting $y = \xi_{\min}$, $a_0 = 0.189$ and $M = 6$ in Eq. (3)): Table 3 summarizes the expression of B_m and corresponding coefficients (a_m).

$$\xi_{\min} = 0.189 + \sum_{m=1}^6 a_m B_m(x) \quad (8)$$

Table 3. Basis functions and their corresponding coefficient (a_m) for ξ_{\min} .

Basis Functions	Equation	Coefficient(a_m)
$B_1(x)$	$\max(0, r - 0.0343)$	0.545
$B_2(x)$	$\max(0, 0.059 - \sigma)$	2.864
$B_3(x)$	$\max(0, s - 0.545)$	1.093
$B_4(x)$	$\max(0, 0.545 - s)$	-0.311
$B_5(x)$	$B_3(x) * \max(0, \sigma - 0.283)$	7.235
$B_6(x)$	$\max(0, \sigma - 0.059) * \max(0, s - 0.245)$	-3.310

The performance of training and testing dataset has been determined by using Eq. (8). Figures 5 and 6 depict the performance of training and testing dataset respectively. It is clear from Figs. 5 and 6 that the value of R^2 is close to one. Therefore, the developed MARS has ability for predicting ξ_{\min} .

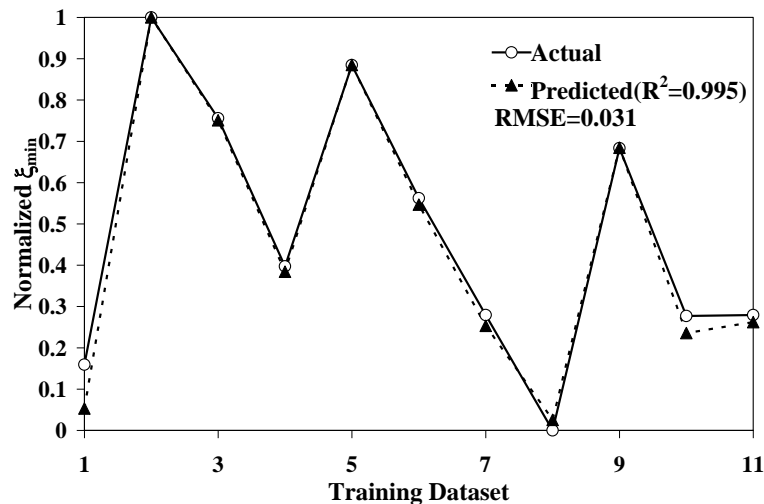


Fig. 5. Performance of the training dataset for ξ_{\min} .

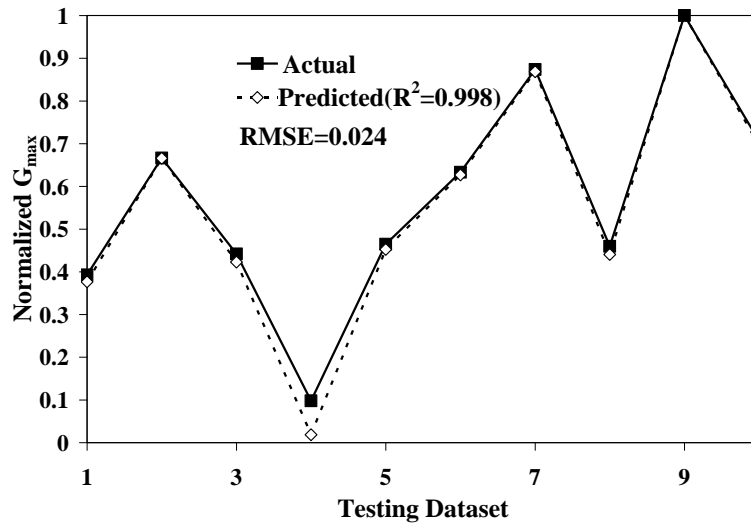


Fig. 6. Performance of the testing dataset for ξ_{\min} .

The results of MARS have been compared with ANFIS, multi-layer perceptron (MLP) and multiple regression analysis method (MRM) developed by [6]. The comparison has been carried out in terms of R^2

Figure 7 illustrates the bar chart of R^2 values of the different model.

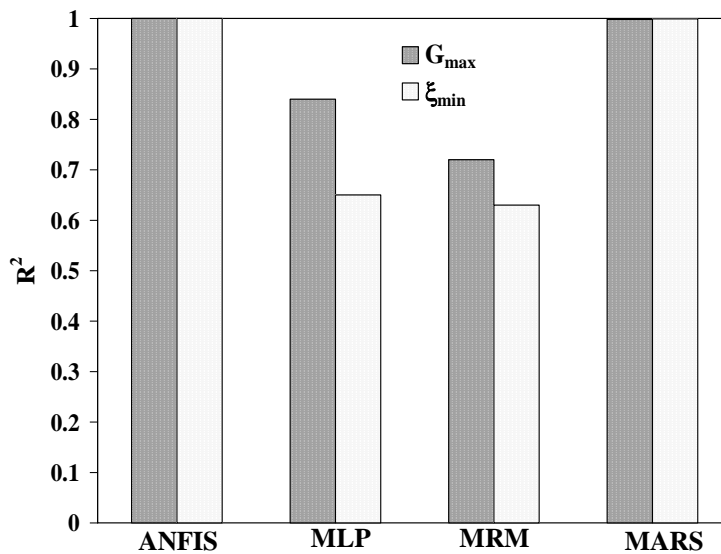


Fig. 7. Bar chart of R^2 for the different models.

The value of R^2 of ANFIS, MLP and MRM is given by [6]. It is observed from Fig. 7 that the performance of MARS and ANFIS is almost same. However, the developed MARS outperforms the MLP and MRM models. The developed MARS gives Eq. (6) and (7) for prediction of G_{\max} and ξ_{\min} . But, the developed ANFIS and MLP did not give any equation for prediction of G_{\max} and ξ_{\min} .

Figure 8 shows the results of sensitivity analysis. It is clear from Fig. 8 that r has maximum effect on G_{\max} and ξ_{\min} .

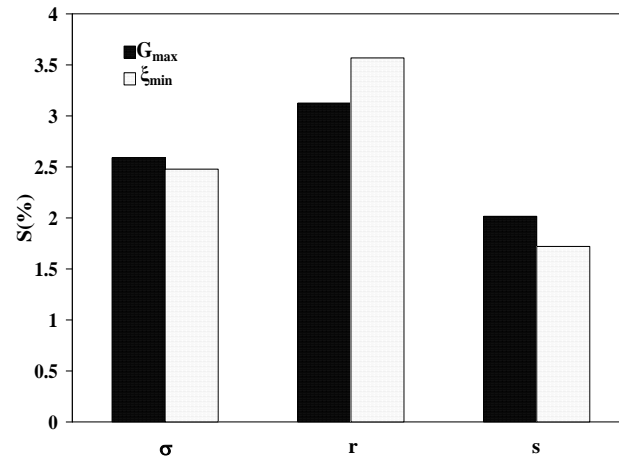


Fig 8. Sensitivity analysis of the input parameters.

5. Conclusion

This article successfully adopted MARS for prediction of G_{max} and ξ_{min} of synthetic reinforced soil. The developed MARS has shown good predictive abilities. The performance of MARS is comparable to ANFIS. However, the developed MARS outperforms the MLP and MRM models. Geotechnical engineers can use the developed equations for determining G_{max} and ξ_{min} . Sensitivity analysis shows that r has maximum impact on G_{max} and ξ_{min} . This study shows that MARS can be used as a robust tool for solving different problems in geotechnical earthquake engineering.

References

- [1] A. McGown, K. Z. Andrawes, and M. M. Al-Hasani, "Effect of inclusion properties on the behavior of a sand," *Geotechnique*, vol. 28, no. 3, pp. 237-346, 1978.
- [2] N. C. Consoli, P. D. M. Prietto, and L. A. Ulbrich, "The behavior of a fiber-reinforced cemented soil," *Ground Improv*, vol. 3, pp. 21-30, 1999.
- [3] A. J. Puppala and C. Musenda, "Effects of fiber reinforcement on strength and volume change in expansive soils," *Transport Res Rec*, vol. 1736, pp. 134-40, 2000.
- [4] J. J. Murray, J. D. Frost, and Y. Wang, "Behavior of a sandy silt reinforced with discontinuous recycled fiber inclusions," *Transport Res Rec*, vol. 1714, pp. 9-17, 2000.
- [5] R. L. Santoni, J. S. Tingle, and S. Webster, "Engineering properties of sand-fiber mixtures for road construction," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 127, no. 3, pp. 258-68, 2001.
- [6] S. Akbulbut, A. S. Hsiloglu, and S. Pamukun, "Data generation for shear modulus and damping ration in reinforced sands using adaptive neuro-fuzzy inference system," *Soil Dynam Earthquake Eng*, vol. 24, pp. 805-814, 2004.
- [7] J. H. Friedman, "Multivariate adaptive regression splines," *The Annals of Statistics*, vol. 19, pp. 1-14, 1991.
- [8] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*. New York: Springer-Verlag, 2003.
- [9] P. A. W. Lewis and J. G. Stevens, "Nonlinear modeling of time series using multivariate adaptive regression splines (mars)," *Journal of the American Statistical Association*, vol. 86, no. 416, pp. 864-877, 1991.
- [10] T. Ekman and G. Kubin, "Nonlinear prediction of mobile radio channels: measurements and mars model designs," in *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing*, 1999, vol. 5, pp. 2667-2670.
- [11] C. C. Yang, S. O. Prasher, R. Lacroix, and S. H. Kim, "Application of multivariate adaptive regression splines (mars) to simulate soil temperature," *Transactions of the ASAE*, vol. 47, no. 3, pp. 881-887, 2004.

- [12] S. Crino and D. E. Brown, "Global optimization with multivariate adaptive regression splines," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 2, pp. 333–340, 2007.
- [13] E. Deconinck, D. Coomans, and V. Y. Heyden, "Exploration of linear modelling techniques and their combination with multivariate adaptive regression splines to predict gastro-intestinal absorption of drugs," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 43, pp. 119–130, 2007.
- [14] N. O. Attoh-Okine, K. Cooger, and S. Mensah, "Multivariate adaptive regression (MARS) and hinged hyperplanes (HHP) for doweled pavement performance modeling," *Construction and Building Materials*, vol. 23, no. 9, pp. 3020-3023, 2009.
- [15] J. D. Andrés, P. Lorca, F. J. D. C. Juez, and F. S. Lasheras, "Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS)," *Expert Systems with Applications*, vol. 38, pp. 1866-1875, 2011.
- [16] F. Vidoli, "Evaluating the water sector in Italy through a two stage method using the conditional robust nonparametric frontier and multivariate adaptive regression splines," *European Journal of Operational Research*, vol. 212, pp. 583-595, 2011.
- [17] S. Sekulic and B. R. Kowalski, "MARS: A tutorial," *J. Chemom.*, vol. 6, pp. 199–216, 1992.
- [18] P. Craven and G. Wahba, "Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation," *Numer Math*, vol. 31, pp. 317-403, 1979.
- [19] S. Y. Liong, W. H. Lim, and G. N. Paudyal, "River stage forecasting in Bangladesh: neural network approach," *Journal of Computing in Civil Engineering*, vol. 14, no. 1, pp. 1-8, 2000.

