



Developing multivariate adaptive regression splines model for predicting the undrained shear strength of clayey soil from cone penetration test data

Dai Yu¹

Received: 1 September 2021 / Accepted: 22 December 2021 / Published online: 16 January 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

In geotechnical planning methods, the undrained shear strength of clayey soil is very important as one of the engineering features. Over the past years, several theoretical and empirical methods have been developed to estimate the undrained shear strength based on soil properties using in-situ tests such as cone and piezocone penetration tests. However, most of these methods involve correlation assumptions that can result in inconsistent accuracy. In this study, multivariate adaptive regression splines (MARS) model with different degrees of interactions was developed for predicting the undrained shear strength of soil from cone penetration test data. To this aim, the model had five variables named cone tip resistance, sleeve friction, liquid limit, plastic limit, and overburden weight as inputs and undrained shear strength of soil as output. In all proposed models, the estimated USS values demonstrate acceptable agreement with experimental records, representing the workability of proposed equations for predicting the USS values with high accuracy. Comparison of three developed equations supplied that MARS-O4 has a better result than MARS-O3, followed by MARS-O2. Furthermore, by apprising the PI and OBJ indexes, the MARS-O4 model outperforms the other two models, with lower PI and OBJ values equal to 0.1464 and 169.14. Therefore, the 4th interaction equation of MARS for predicting the undrained shear strength of soil can be recognized as the proposed regression model.

Keywords Cone penetration test · Undrained shear strength of soil · Prediction · Multivariate adaptive regression splines (MARS)

1 Introduction

The potency of soil to tolerate shear stress is referred to by the undrained shear strength (USS) of soil. It is one primary value in the computation of various geotechnical phenomena, such as settlement (Esmaili-Falak et al. 2017, 2018). Moreover, to design foundations (deep and shallow foundations, both of them), the USS of soil is also of superior matter. Hence, circumspect appraisal of the shear strength is indispensable. Various analytical and experimental works can be performed on different types of soils to determine the properties of soils (Sarkhani Benemaran 2017; Poorjafar et al. 2021; Esmaili-Falak 2017; Esmaili-Falak et al. 2020). After the years, scholars had extended plenty of analytical and experiential

procedures to specify the shear strength (Sarkhani Benemaran et al. 2020), for instance, bearing capacity (BCM), strain path (SPM), cavity expansion (CEM), and finite element (FEM) procedures (Huang et al. 2004). In addition, mixtures of mentioned procedures were examined as well to get an advanced commentary of shear strength, such as cavity expansion-finite element (Abu-Farsakh et al. 2003), cavity expansion-strain path (Yu and Whittle 1999), cavity expansion-bearing capacity (Salgado et al. 1997), and strain path-finite element (Teh and Houlsby 1991).

Nevertheless, many of these techniques synthesized streamlining presumptions about soil condition, margin circumstances, and failure benchmark (Esmaili Falak et al. 2020). Hence, detecting theoretic procedures requires endorsement from in-situ and experimental soil parameters. In this regard, the unconsolidated undrained triaxial test can be very helpful. However, unfortunately, managing triaxial tests need more time as well as is costly. Furthermore, an

✉ Dai Yu
ziranrain@163.com

¹ Department of Ship and Marine Engineering, JiangSu Shipping College, Nantong 226010, China

unavoidable disorder in transport, handling, and collecting soil specimens means the test conclusions are controversial.

In-situ cone penetration tests can be effectually applied for soil recognition, appraisal of soil virtues, such as shear strength, and numbers of other geotechnical usages. In-situ CPT is trusty, quick, and economical compared to the old soil properties tests related to boring and experimental tests. The CPT test can supply persistent profile information with profundity relevant to soil stiffness and strength variables, which is good to estimate the USS of soil. Hence, some experimental procedures were extended in the last decades to approximate the USS from in-situ cone penetration test parameters (Lunne 1982; Senneset 1982). Nonetheless, many numbers of these techniques bring some presumptions and discernments in choosing suitable correlation coefficients factors like the cone tip factor, N_{kt} among the CPT profiles information and USS that can impact the computation of shear strength. This can affect in incompatible precision of appraising the USS for various site situations.

In the recent decades, using artificial-based neural networks has enhanced and carried out with successful results in various civil (Masoumi et al. 2020) and geotechnical engineering problems, for example, shallow foundation settlement, liquefaction, the behavior of frozen soils, settlement behavior of pile foundations, swelling pressure of expansive soil and so forth (Sarkhani Benemaran et al. 2020; Das and Basudhar 2006; Ikizler et al. 2010; Neaupane and Achet 2004; Nejad and Jakska 2017; Shahin et al. 2009; Esmaeili-Falak et al. 2019; Nassr et al. 2018). Regression is also one of the methods to determine the relationship between input and output variables such as MARS (Zhang et al. 2021; Sahraei et al. 2021; Raja and Shukla 2021).

Neural network application to estimate the USS of soil from the CPT test is anticipated to solve the deficiencies mentioned above in traditional techniques because there is no correlation hypothesis or judgment (Samui and Kurup 2012). To some extent, the neural network method tends to iterate the human brain learnings from a prior instance and is learned by specific mathematical methods. This aim can be gained from reiterative steps by regulating the weights, node numbers, and a number of layers. The difficulty of the neural network could be changed by altering the transfer function or the model form (Shahin et al. 2002). After identifying the most precise neural network model after training well, the developed algorithm could be explored for estimating the USS for other sites. An ANN is utilized to develop a model with more firm forecasting of USS from CPT records in a study (Abu-Farsakh and Mojumder 2020). First, a dataset was created of soil boring records and experimental tests from 70 sites located in Louisiana. Then, various neural network models were trained by cone tip resistance, sleeve friction, and other assessable soil characteristics. The conclusions were next compared with an experiential reference technique of

specifying USS from CPT. The outputs specified that the neural network techniques outperformed the reference method, which approves this model's applicability in estimating the USS.

Another study applied data-driven extreme gradient boosting and random forest methods optimized with the Bayesian optimization method to find the relationships between the USS and soil parameters. To this aim, five variables containing the pre-consolidation stress, effective vertical stress, liquid limit, plastic limit, and natural water content are selected. It is shown that XGBoost- and RF-based models outperform others. Along with this, the XGBoost-based model provides properties significance ranks, which introduces it as an efficient tool for predicting geotechnical parameters and enhancing the model's interpretability (Zhang et al. 2021).

The prime objective of this paper is to find out the applicability of utilizing the multivariate adaptive regression splines (MARS) model for predicting the USS of soil from cone penetration test records to generate models which are to be used in practical applications. Moreover, various degrees of interactions of models are examined to have comprehensive, accurate, and reliable outputs. To gain this aim, the model had five input variables named cone tip resistance, sleeve friction, liquid limit, plastic limit, overburden weight, and USS of soil as output. To evaluate the accuracy of the proposed models, six statistical performance indices were considered.

2 Dataset and methodologies

2.1 Description of the dataset

To predict the undrained shear strength of soil from the cone penetration test, CPT data and corresponding bore log data were collected from 70 different sites in Louisiana (Fig. 1) (Mojumder 2020). Five different variables that can affect the value of the USS were considered as input variables. These variables included: cone tip resistance (CTR), sleeve friction (SF), liquid limit (LL), plastic limit (PL), and overburden weight (OBW). The statistics and histograms of the variables used for developing the model along with their normal distribution curves are given in Table 1 and Fig. 2, respectively.

2.2 Multivariate adaptive regression splines (MARS)

Multivariate asymmetric regression process is defined as multivariate adaptive regression spline (MARS) (Friedman 1991; Sekulic and Kowalski 1992; Friedman and Roosen 1995). Its most important duty is the amount's forecast of a continual affiliate variable, $y(n \times 1)$, using a series of inputs

Fig. 1 Locations of cone penetration test (CPT) sites

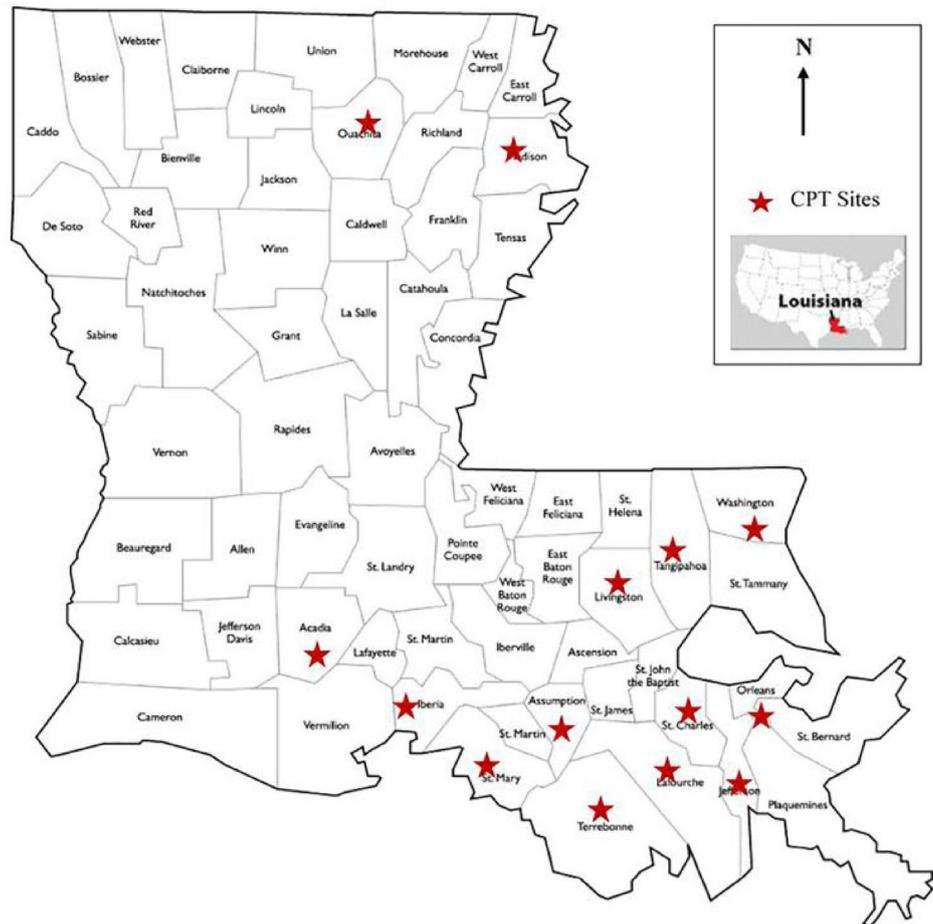


Table 1 The statistical values of the input and output variables

Variable	Inputs					Output
	CTR	SF	LL	PL	OBW	USS
Minimum	1.59	0.02	24.00	12.00	0.03	100.00
Maximum	32.86	1.40	133.00	50.00	3.64	5670.00
St. deviation	6.4164	0.28867	19.836	6.9231	0.82181	966.1815
Average	12.03	0.42	63.49	27.48	1.57	1271.91
Median	10.10	0.36	61.00	26.00	1.52	965.00
Skewness	1.0306	1.2994	0.6217	0.6192	0.2817	2.0557
Kurtosis	0.6327	1.3710	0.2248	- 0.0772	- 0.5301	5.1846

that is independent, $X(n \times p)$. The MARS could be presented as:

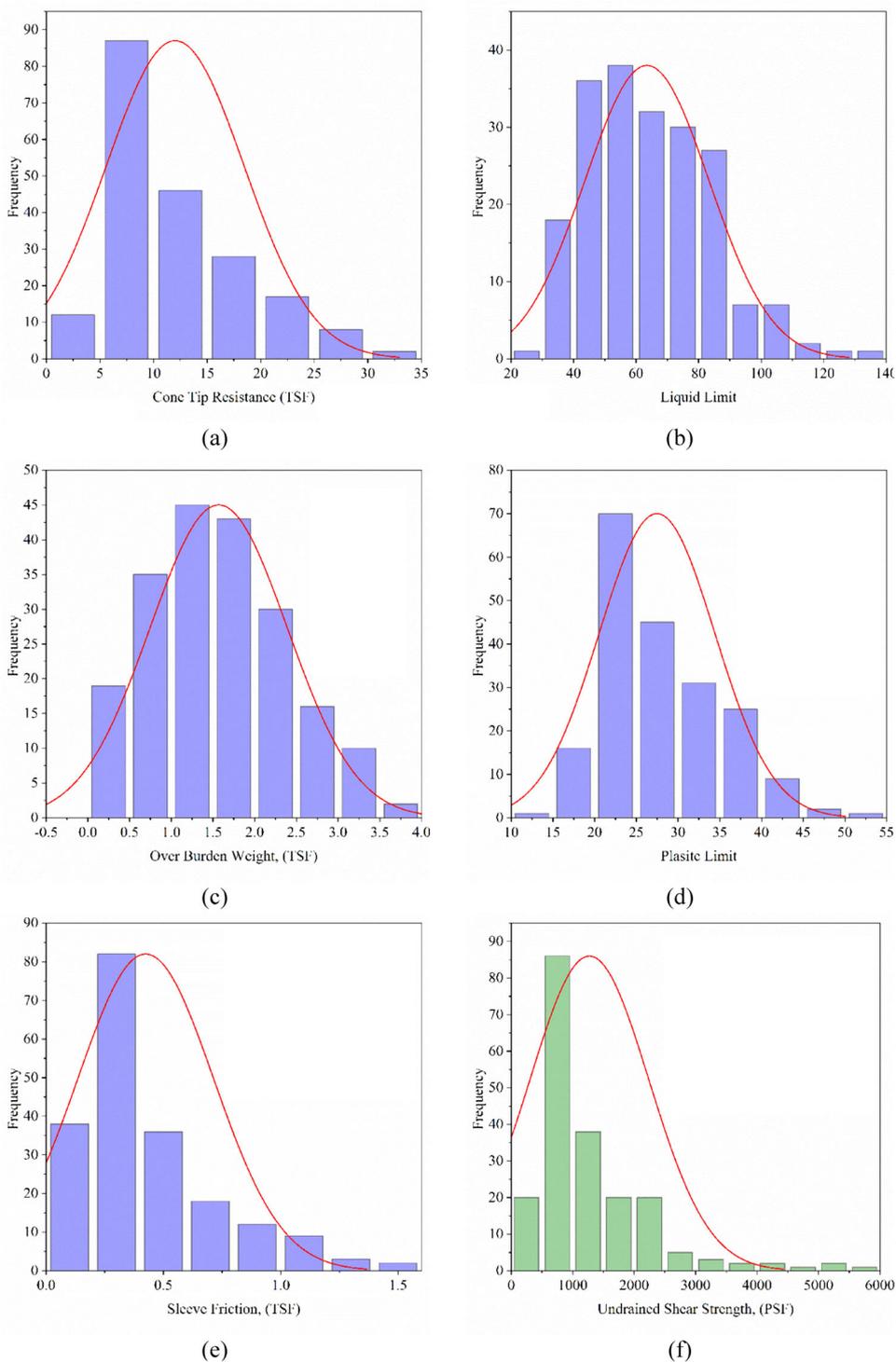
$$y = f(x) + e \tag{1}$$

where f is a sum of basis functions which is weighted that depend on x , and e presented the error vector in $(n \times 1)$ dimension. The multivariate adaptive regression spline method is a generalization of classification and regression trees (Hastie et al. 2001) but prevailing the constraints of classification and regression tree (CART). This regression model does not

request any preference assumption in the case of the fundamental operational connection between inputs and output variables. Apart from that, this connection is appointed from a totality of coefficients and piecewise multinomial of degree basis functions (q) fully related to the regression records (x, y) .

This regression method is created by coordinating the basis function to various spaces of the inputs. MARS uses

Fig. 2 Histograms of input (a–e) and output (f) variables and their normal distribution curves



bilateral shorten power operates as spline basis function (BF), shown in the below relations (Friedman and Roosen 1995):

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

$$[+(x - t)]_+^q = \begin{cases} (t - x)^q & \text{if } x \geq t \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In these relations, the power to which the splines are picked up specifies the degree of the monotony of the resulting appraise all defined $q (\geq 0)$. Notice that when q is equal to one, just simple lineal splines are appraised. For example, Fig. 3 shows a pair of splines for q equal to one at node 3.5. This figure shows a mirrored pair of hinge functions with a knot at for example 3.5.

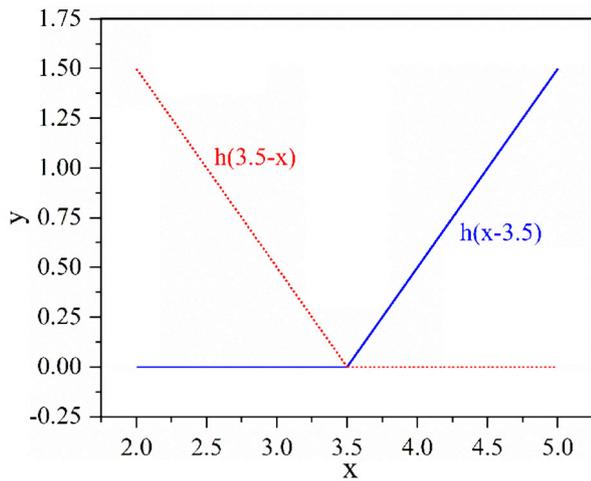


Fig. 3 A schematic exhibition of a spline BF. Dashed line ($x < t, -(x-t)$), solid line ($x > t, +(x-t)$)

MARS of a dependent variable y with M basis function could be represented as Eq. 4 (Xu et al. 2004; Cheng and Cao 2014; Benemaran and Esmaeili-Falak 2020):

$$\hat{y} = \hat{f}_m(x) = c_0 + \sum_{m=1}^M c_m B_m(x) \tag{4}$$

being \hat{y} the affiliate variable forecasted by the MARS, c_0 , $B_m(x)$ and c_m shows a constant, the m th basis function and the coefficient of m th basis function, respectively. Specifically, both parameters are demonstrated in the model, and the node situations for each variable have to be optimized. Moreover, this regression method utilizes generalized cross-validation (GCV) to describe the basic function contained in the model.

Indeed, the mean squared residual error apportioned by a penalty factor specifying the generalized cross-validation which pertains to the model intricacy, and it is described as Eq. 5 (Friedman 1991; Friedman and Roosen 1995):

$$GCV(M) = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}_M(x_i))^2}{(1 - C(M)/n)^2} \tag{5}$$

So that an intricacy penalty that increases with a number of basis function in the model is shown as $C(M)$ (Benemaran and Esmaeili-Falak 2020):

$$C(M) = (M + 1) + dM \tag{6}$$

where M and d value the basis function in Eq. (4) and a penalty factor for all one of BF included in the model. In addition, the must utilize the GCV make clear before with the parameters N-subsets and the residual sum of squares

to achieve correct results. The MARS model was performed with the ARESLab toolbox in MATLAB (Jekabsons 2011).

2.3 Performance evaluators

Different statistical evaluators were used to appraisal the performance of developed models for predicting the USS. Coefficient of determination (R^2), root mean squared error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), performance index (PI), and OBJ were used as precision measurements (Eqs. 7–12):

$$R^2 = \left(\frac{\sum_{p=1}^P (t_p - \bar{t})(y_p - \bar{y})}{\sqrt{\left[\sum_{p=1}^P (t_p - \bar{t})^2 \right] \left[\sum_{p=1}^P (y_p - \bar{y})^2 \right]}} \right)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{P} \sum_{p=1}^P (y_p - t_p)^2} \tag{8}$$

$$MAPE = \frac{1}{P} \sum_{p=1}^P \left| \frac{t_p - y_p}{t_p} \right| \tag{9}$$

$$MAE = \frac{1}{P} \sum_{p=1}^P |y_p - t_p| \tag{10}$$

$$PI = \frac{1}{|\bar{t}|} \frac{RMSE}{\sqrt{R^2 + 1}} \tag{11}$$

$$OBJ = \frac{RMSE + MAE}{R^2 + 1} \tag{12}$$

where, y_p represent the predicted values of the P th pattern, t_p depicts the target values of the P th pattern, \bar{t} shows the averages of the target values, \bar{y} is the averages of the predicted values, and P is the number of dataset.

3 Result and discussion

The details of the basis functions and corresponding equations are shown in Table 2 for the degree of interactions of 2, 3, and 4 MARS equations. The explicable MARS approach different orders (from 2 to 4) formulations to estimate the USS results are created in Eqs. (13–15), respectively. Basis functions of 2, 3, and 4 order MARS models from 3 to 37 were assessed and measured. With an incremental number of basis functions, the performance indicators for data grew to the proper orientation. They gained the supreme outcomes when the number of basis functions was 17, 29, and 31 for MARS-O2, O3, and O4, respectively. By increasing the degree of interactions of equations from 2 to 4, the values of R^2 raised from 0.8339 to 0.8619. on the other hand, RMSE

Table 2 Basis functions and corresponding equations of MARS approach for USS

BF	Equation		
	Order 2 (Eq. 13)	Order 3 (Eq. 14)	Order 4 (Eq. 15)
BF1	max (0, CTR-15.27)	max (0, CTR-15.27)	max (0, CTR-15.27)
BF2	max (0, 15.27-CTR)	max (0, 15.27-CTR)	max (0, 15.27-CTR)
BF3	BF1 × max (0, SF-0.71)	BF1 × max (0, SF-0.71)	BF1 × max (0, SF-0.71)
BF4	max (0, OBW-2.81)	BF3 × max (0, 73-LL)	BF3 × max (0, LL-73)
BF5	max (0, 2.81-OBW)	max (0, 2.81-OBW)	BF3 × max (0, 73-LL)
BF6	max (0, 23-CTR) × BF4	BF3 × max (0, OBW-1.76)	max (0, 2.81-OBW)
BF7	BF1 × max (0, LL-24)	BF3 × max (0, 1.76-OBW)	BF1 × max (0, SF-0.71) × max (0, OBW-1.76)
BF8	max (0, SF-0.63) × BF5	BF3 × max (0, PL-12)	BF1 × max (0, SF-0.71) × max (0, 1.76-OBW)
BF9	max (0, SF-0.48)	max (0, CTR-21.78) × BF5	BF1 × max (0, SF-0.71) × max (0, 73-LL) × max (0, PL-12)
BF10	max (0, CTR-21.78) × BF9	max (0, 21.78-CTR) × BF5	max (0, CTR-21.78) × BF6
BF11	max (0, CTR-21.78) × BF5	BF3 × max (0, OBW-2.51)	BF7 × max (PL-12)
BF12		max (0, CTR-20.1) × max (0, 2.81-OBW)	max (0, CTR-20.1) × BF6
BF13		BF10 × max (0, SF-0.63)	max (0, CTR-17.45) × BF6

shows a decline of about 34.8, and MAE values decreased from 296.85 to 270.84.

Order 2:

$$\begin{aligned}
 USS = & 1213.536 + 137.701 \times BF1 - 123.36 \times BF2 \\
 & + 615.968 \times BF3 + 2791.082 \times BF4 + 322.3104 \\
 & \times BF5 - 252.727 \times BF6 - 1.294 \times BF7 \\
 & - 2204.53 \times BF8 + 1416.494 \times BF9 \\
 & - 774.561 \times BF10 + 204.69 \times BF11
 \end{aligned}
 \tag{13}$$

Order 3:

$$\begin{aligned}
 USS = & 1041.429 + 108.146 \times BF1 - 74.742 \times BF2 \\
 & - 484.94 \times BF3 + 25.556 \times BF4 + 689.83 \times BF5 \\
 & - 323.353 \times BF6 - 826.5 \times BF7 + 34.329 \times BF8 \\
 & + 536.98 \times BF9 - 32.992 \times BF10 + 646.669 \\
 & \times BF11 - 371.9 \times BF12 - 114.455 \times BF13
 \end{aligned}
 \tag{14}$$

Order 4:

$$\begin{aligned}
 USS = & 1250.073 + 78.435 \times BF1 - 117.763 \times BF2 \\
 & - 456.211 \times BF3 + 65.9 \times BF4 + 31.39 \times BF5 \\
 & + 274.632 \times BF6 - 1164.09 \times BF7 - 1319.083 \\
 & \times BF8 + 1.255 \times BF9 + 708.68 \times BF10 + 75.149 \\
 & \times BF11 - 782.71 \times BF12 + 275.08 \times BF13
 \end{aligned}
 \tag{15}$$

The result of developed models for predicting USS value is presented as follows. Figure 4 specifies acceptable potential

in the modeling phase. Comparing the measured records from experimental efforts with those predicted by MARS-O2, O3, and O4 models are supplied in Fig. 4. It can be observed that the developed models have R^2 larger than 0.834 and 0.9236. It means that the correlation between measured and predicted values from developed models is in good correlation so that it shows the highest accuracy in the regression process. Besides, to compare the productivity of the applied models, six statistical evaluators (R^2 , RMSE, MAE, MAPE, PI, and OBJ) were utilized. The results are shown in Table 3. MARS-O2 has the worst values regarding MARS models, which its R^2 stood at 0.8339, and PI 0.1618. All indices are better by increasing the MARS interaction. For instance, RMSE declines from 393.79 to 358.99. By apprising the PI and OBJ indexes as the whole model evaluator, which considers other indexes altogether, the MARS-O4 model outperforms the other two models, with lower PI and OBJ values equal to 0.1464 and 338.27. Therefore, the 4th interaction equation of MARS for predicting USS of soil can be recognized as the proposed regression model.

The performance assessment results of the implemented models are portrayed in Fig. 5, showing a graphical comparison between the distribution of errors. Also, an acceptable fit between measured and predicted USS values are obtainable from the time series plots presented in Fig. 5. As can be seen, in all proposed models, the estimated USS values demonstrate acceptable agreement with experimental records, representing the workability of proposed equations for predicting the USS values with high accuracy. Comparison of three developed equations supplied that MARS-O4 has a better result than MARS-O3, followed by MARS-O2. Based on error distribution figures, the MARS-O4 model

Fig. 4 The scatter plot of between observed and predicted USS

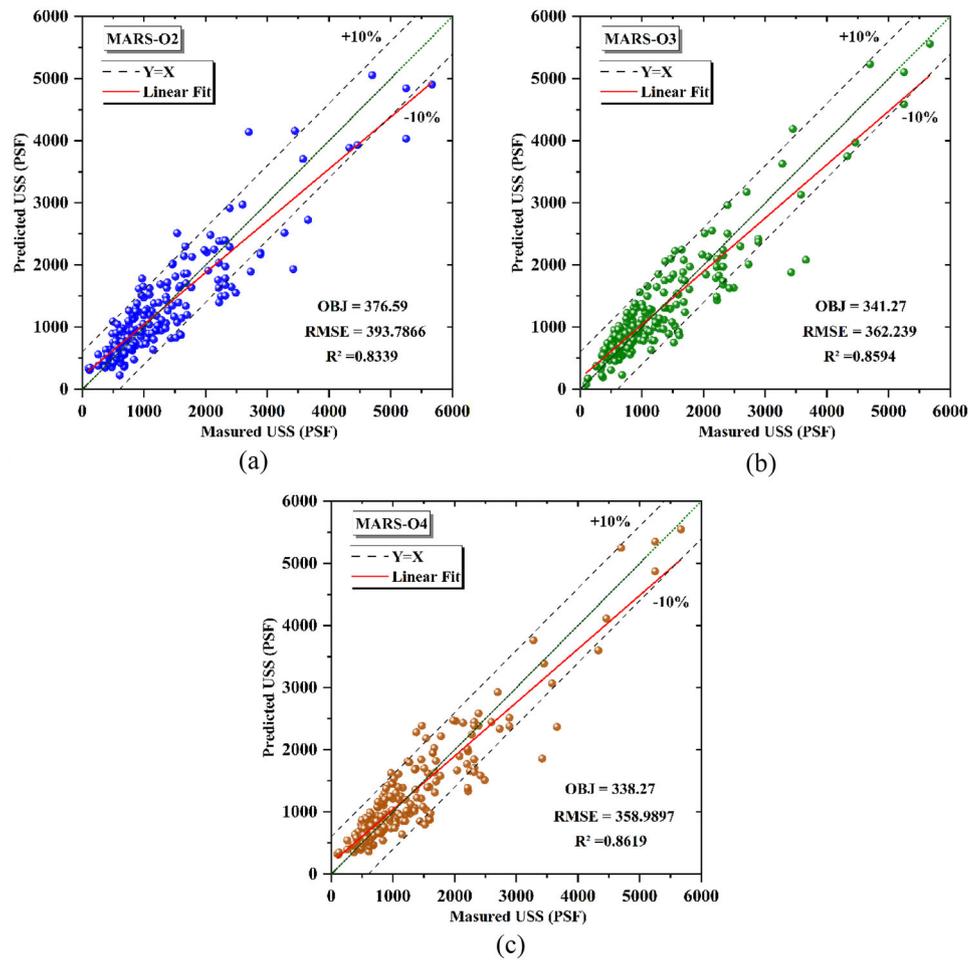


Table 3 The results of created models

Models		MARS-O2	MARS-O3	MARS-O4
Number of basis function		17	29	31
Network results	R ²	0.8339	0.8594	0.8619
	RMSE	393.79	362.24	358.99
	MAE (PSF)	296.85	272.33	270.84
	MAPE (PSF)	0.2014	0.4274	0.3241
	PI	0.1618	0.1478	0.1464
	OBJ	376.59	341.27	338.27

results in the lowest error percentage in the USS predicting process, providing roughly accurate predictions than those of the rest developed methods specified.

3.1 Sensitivity analysis

An evaluation of the sensitivity of the models was conducted to assess the most determinative input parameters to com-

pute the USS. Various input data were built by removing a single input parameter simultaneously, and the test data set reported the amounts of three statistical performance criteria as R^2 , RMSE, and MAE. The best model for the sensitivity analysis is chosen using the statistical performance criteria. In the present study, the MARS-O4 model is selected due to its remarkable performance. The results are as Table 4, which is shown that the cone tip resistance (CTR) is the most influential parameter for predicting the USS using the mentioned model, with a decline of about 0.1831 for R^2 . From this perspective, although they are not as effective as CTR, the overburden weight (OBW) and sleeve friction (SF) parameters are in the following ranks, respectively. It is worth considering that eliminating input variables may only cause a minimal performance loss for the model, but in the present study, because the analysis was based on experimental measurements, eliminating variables could decline the model’s generalizability. Considering the multicollinearity problem has not a significant impact on the fit of a model. It commonly does not impress remarkably on predictions, and the present study does not prefer deleting any variable.

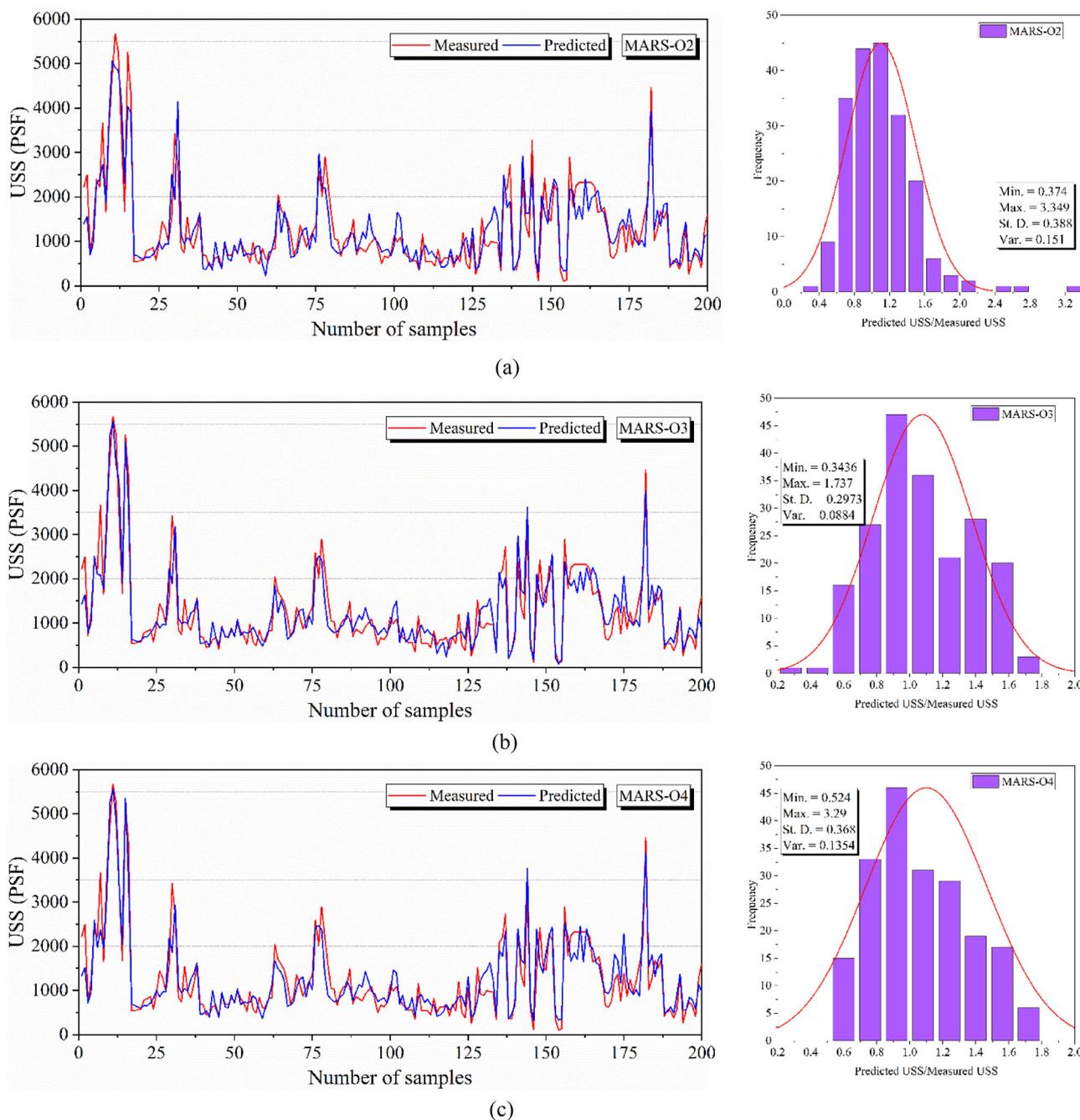


Fig. 5 USS prediction using MARS models with distribution

4 Conclusion

In this study, Multivariate Adaptive Regression Splines (MARS) model with different degrees of interactions were proposed for predicting the undrained shear strength of soil from cone penetration test data. To this aim, the model had five variables named cone tip resistance, sleeve friction, liquid limit, plastic limit, and overburden weight. To evaluate

the accuracy of the developed model, six performance indices were considered.

MARS-O2 has the worst values regarding MARS models, which its R^2 stood at 0.8339, and PI 0.1618. All indices better by increasing the MARS interaction. The MARS-O4 model outperforms the other two models by apprising the PI and OBJ indexes, with lower PI and OBJ values equal to 0.1464 and 338.27. Therefore, the 4th interaction equation of MARS

Table 4 Sensitivity analysis of input variables using MARS-O4 model

Model	Index	Main	Removed parameter				
			CTR	SF	LL	PL	OBW
MARS-O4	R^2	0.8619	0.6788	0.8264	0.8403	0.8573	0.8225
	RMSE	358.99	547.58	402.6	386.13	365.01	407.08
	MAE	270.84	372.672	287.2	282.45	275.84	295.82

Bold values indicate that the effect of CTR is more than other input variables

for predicting USS of soil can be recognized as the proposed regression model.

In all proposed models, the estimated USS values demonstrate acceptable agreement with experimental records, representing the workability of proposed equations for predicting the USS values with high accuracy. Comparison of three developed equations supplied that MARS-O4 has a better result than MARS-O3, followed by MARS-O2. Based on error distribution figures, the MARS-O4 model results in the lowest error percentage in the USS predicting process, providing roughly accurate predictions than those of the rest developed methods specified.

Acknowledgements Science and Technology Planning Project of Nantong City, JiangSu Province (MS22020021), College Students Innovation and Entrepreneurship Training Program of JiangSu Province (202012703018Y), Scientific Research Project of JiangSu Shipping College (HYKY/2020B01).

References

- Abu-Farsakh MY, Mojumder MAH (2020) Exploring artificial neural network to evaluate the undrained shear strength of soil from cone penetration test data. *Transp Res Rec* 2674(4):11–22
- Abu-Farsakh M, Tumay M, Voyiadjis G (2003) Numerical parametric study of piezocone penetration test in clays. *Int J Geomech* 3(2):170–181
- Benemaran RS, Esmaili-Falak M (2020) Optimization of cost and mechanical properties of concrete with admixtures using MARS and PSO. *Comput Concr* 26(4):309–316. <https://doi.org/10.12989/cac.2020.26.4.309>
- Cheng M-Y, Cao M-T (2014) Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Appl Soft Comput* 22:178–188
- Das SK, Basudhar PK (2006) Undrained lateral load capacity of piles in clay using artificial neural network. *Comput Geotech* 33(8):454–459
- Esmaili Falak M, Sarkhani Benemaran R, Seifi R (2020) Improvement of the mechanical and durability parameters of construction concrete of the qotursuyi spa. *Concr Res* 13(2):119–134. <https://doi.org/10.22124/JCR.2020.14518.1395>
- Esmaili-Falak M, Katebi H, Javadi A, Rahimi S (2017) Experimental investigation of stress and strain characteristics of frozen sandy soils—A case study of Tabriz subway. *Modares Civ Eng J* 17(5):13–23
- Esmaili-Falak M, Katebi H, Javadi A (2018) Experimental study of the mechanical behavior of frozen soils—A case study of tabriz subway. *Period Polytech Civ Eng* 62(1):117–125
- Esmaili-Falak M, Katebi H, Vadiati M, Adamowski J (2019) Predicting triaxial compressive strength and Young's modulus of frozen sand using artificial intelligence methods. *J Cold Regions Eng* 33(3):4019007. [https://doi.org/10.1061/\(ASCE\)CR.1943-5495.0000188](https://doi.org/10.1061/(ASCE)CR.1943-5495.0000188)
- Esmaili-Falak M, Katebi H, Javadi AA (2020) Effect of freezing on stress-strain characteristics of granular and cohesive soils. *J Cold Regions Eng* 34(2):05020001. [https://doi.org/10.1061/\(ASCE\)CR.1943-5495.0000205](https://doi.org/10.1061/(ASCE)CR.1943-5495.0000205)
- Esmaili-Falak M (2017) Effect of system's geometry on the stability of frozen wall in excavation of saturated granular soils. Doctoral dissertation, University of Tabriz
- Friedman JH (1991) Multivariate adaptive regression splines. *Ann Stat* 2:1–67
- Friedman JH, Roosen CB (1995) An introduction to multivariate adaptive regression splines. Sage Publications Sage CA, Thousand Oaks
- Hastie T, Tibshirani R, Friedman JH (2003) The elements of statistical learning, corrected. In: Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (eds) Distributed and overlapping representations of faces and objects in ventral temporal cortex, vol 293. Springer, Berlin, p 24252430
- Huang W, Sheng D, Sloan SW, Yu HS (2004) Finite element analysis of cone penetration in cohesionless soil. *Comput Geotech* 31(7):517–528
- Ikizler SB, Aytakin M, Vekli M, Kocabaş F (2010) Prediction of swelling pressures of expansive soils using artificial neural networks. *Adv Eng Softw* 41(4):647–655
- Jekabsons G (2016) ARESLab: adaptive regression splines toolbox for Matlab/Octave, 2011. <http://www.cs.rtu.lv/jekabsons>
- Lunne T (1982) Role of CPT in North Sea foundation engineering
- Masoumi F, Najjar-Ghabel S, Safarzadeh A, Sadaghat B (2020) Automatic calibration of the groundwater simulation model with high parameter dimensionality using sequential uncertainty fitting approach. *Water Supply* 20(8):3487–3501
- Mojumder MAH (2020) Evaluation of undrained shear strength of soil, ultimate pile capacity and pile set-up parameter from cone penetration test (CPT) using artificial neural network (ANN)
- Nassar A, Esmaili-Falak M, Katebi H, Javadi A (2018) A new approach to modeling the behavior of frozen soils. *Eng Geol* 246:82–90. <https://doi.org/10.1016/j.enggeo.2018.09.018>
- Neaupane KM, Achet SH (2004) Use of backpropagation neural network for landslide monitoring: a case study in the higher Himalaya. *Eng Geol* 74(3–4):213–226
- Nejad FP, Jaksa MB (2017) Load-settlement behavior modeling of single piles using artificial neural networks and CPT data. *Comput Geotech* 89:9–21
- Poorjafar A, Esmaili-Falak M, Katebi H (2021) Pile-soil interaction determined by laterally loaded fixed head pile group. *Geomech Eng* 26(1):13–25. <https://doi.org/10.12989/gae.2021.26.1.013>
- Raja MNA, Shukla SK (2021) Multivariate adaptive regression splines model for reinforced soil foundations. *Geosynth Int* 21:1–23

- Sahraei MA, Duman H, Çodur MY, Eyduran E (2021) Prediction of transportation energy demand: multivariate adaptive regression splines. *Energy* 224:120090
- Salgado R, Boulanger RW, Mitchell JK (1997) Lateral stress effects on CPT liquefaction resistance correlations. *J Geotech Geoenviron Eng* 123(8):726–735
- Samui P, Kurup P (2012) Multivariate adaptive regression spline and least square support vector machine for prediction of undrained shear strength of clay. *Int J Appl Metaheuristic Comput (IJAMC)* 3(2):33–42
- Sarkhani Benemaran R (2017) Experimental and analytical study of pile-stabilized layered slopes. Tabriz University, Tabriz
- Sarkhani Benemaran R, Esmaili-Falak M, Katebi H (2020) Physical and numerical modelling of pile-stabilized saturated layered slopes. *Proc Inst Civ Eng Geotech Eng*. <https://doi.org/10.1680/jgeen.20.00152>
- Sekulic S, Kowalski BR (1992) MARS: a tutorial. *J Chemom* 6(4):199–216
- Senneset K (1982) Strength and deformation parameters from cone penetration tests
- Shahin MA, Maier HR, Jaksa MB (2002) Predicting settlement of shallow foundations using neural networks. *J Geotech Geoenviron Eng* 128(9):785–793
- Shahin MA, Jaksa MB, Maier HR (2009) Recent advances and future challenges for artificial neural systems in geotechnical engineering applications. *Adv Artif Neural Syst* 2009:2
- Teh CI, Houlsby GT (1991) An analytical study of the cone penetration test in clay. *Geotechnique* 41(1):17–34
- Xu Q-S et al (2004) Multivariate adaptive regression splines—Studies of HIV reverse transcriptase inhibitors. *Chemom Intell Lab Syst* 72(1):27–34
- Yu HS, Whittle AJ (1999) Combining strain path analysis and cavity expansion theory to estimate cone resistance in clay. Unpublished notes
- Zhang W, Wu C, Li Y, Wang L, Samui P (2021) Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. *Georisk Assess Manag Risk Eng Syst Geohazards* 15(1):27–40
- Zhang W, Wu C, Zhong H, Li Y, Wang L (2021) Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geosci Front* 12(1):469–477. <https://doi.org/10.1016/j.gsf.2020.03.007>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.