



A hybrid computing model to predict rock strength index properties using support vector regression

Abbas Abbaszadeh Shahri¹ · Fardad Maghsoudi Moud² · Seyed Poorya Mirfallah Lialestani³

Received: 1 January 2020 / Accepted: 26 May 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

The uniaxial compressive strength (UCS) and elasticity modulus (E) are two of the most quoted rock strength parameters in engineering application. Due to approved technical difficulties indirect measurements, the tendency for determining these parameters through predictive models using simpler and cheaper tests in practical oriented applications have widely been highlighted. In this paper, a new hybridized multi-objective support vector regression (MSVR) model integrated with the firefly metaheuristic algorithm (FMA) was developed to touch upon a computational method in rock engineering purposes. The optimum internal parameters were adjusted through parametric investigation using 222 physical and mechanical rock properties corresponding to a variety of quarried stones from all over Iran. The accuracy and robustness of models were evaluated using different error indices, the area under curve for receiver operation characteristics (AUC_{ROC}) and F_1 -score criteria. Comparing to MSVR, the predictability level of UCS and E showed 8.35% and 5.47% improvement in hybrid MSVR-FMA. The superior and more promising results imply that hybrid MSVR-FMA as a flexible alternative can be applied for rock strength prediction in designing of construction projects. Using tow sensitivity analyses, the point load index and P -wave velocity were distinguished as the main effective factors on predicted UCS and E .

Keywords Hybrid model · Predictability level · Multi-objective · Quarried stone · Rock strength · Firefly algorithm

1 Introduction

The natural stones as one of the efficient but oldest recognized materials provide many possibilities in different civil and construction applications (e.g., building industry, road base, paving, concrete, and asphalt). However, due to heterogeneity of these materials plethora and quite variable engineering characteristics can be observed.

Strength properties, durability, attractiveness (appearance and color), cost, economy and quarrying susceptibility are the primary common criteria in selecting the appropriate building stones. In specific applications, some other

properties such as hardness, toughness, specific gravity, porosity and water absorption, dressing, seasoning, workability, fire, and chemical resistance also may need to be considered. Using laboratory tests suitability of stones for building purposes can be evaluated. However, many scholars indicated that laboratory tests for determining uniaxial compressive strength (UCS) and elasticity modulus (E) as two of the most important rock mechanical characteristics is a challenging task (e.g., [2, 11, 42]). Thereby, in practice prediction of UCS and E using statistical regression of simple, inexpensive and non-destructive tests are preferred and notified (e.g., [1, 2, 36, 42, 43, 45, 53]). However, such correlations due to inconsistency of various rock types have shown different degrees of success. Furthermore, the regression analyses among the variables do not imply causality [50], and the strong relationship between variables can be the result of the influence of other unmeasured parameters [46]. Therefore, to interpret predictive statistical model different incompetence (e.g. assumptions, subjective judgment of unobserved data, effect of auxiliary factors, uncertainty of experimental tests, inaccurate prediction in wide expanded range of data) should be considered [2, 4, 27, 29].

✉ Abbas Abbaszadeh Shahri
a_abbaszadeh@iauh.ac.ir

¹ Department of Civil Engineering, Roudehen Branch, Islamic Azad University, Tehran, Iran

² Department of Earth Systems Analysis, Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands

³ Departmetn of Mining, Industrial and TIC Engineering, Universitat Politecnica de Catalunya, Barcelona, Spain

The demerits of statistical techniques in producing more efficient and accurate predictive models can be covered using different subcategories of soft computing approaches. The literature reviews highlighted that soft computing techniques such as artificial neural networks (e.g. [5, 35, 36]), support vector machine [6], random forest [38], genetic programming [7, 14], ANFIS [53, 57], Gene expression programming [15] and hybrid systems [11, 31] are able to predict more promising results for the UCS and E than the conventional statistical methods.

The support vector regression (SVR) [24] is a developed novel kind of supervised-learning support vector machine (SVM) for both classification and regression purposes that can map the inputs to an n -dimensional feature space. This model using nonlinear kernel functions simultaneously can maximize predictive accuracy and avoids overfitting [58]. Similar to SVM, the main idea in SVR is always minimizing the error and individualizing the hyperplane which maximizes the margin. However, using small subset of training points in SVR gives enormous computational advantages than SVM which does not depend on the dimensionality of the input space and thus provides excellent generalization capability, with high prediction accuracy [13]. This implies that the possible poor performance of ANNs (e.g. few labeled data points, trapping into local minimal, overfitting) can be treated using SVR to achieve more precise results [49].

In the recent years, different metaheuristic algorithms have been used for possible enhancement in the performance and predictability level of intelligence models (e.g. [8, 11, 61, 65]). Designing supervised learning systems generally is a multi-objective optimization problem [55] which aims to find appropriate trade-offs between several objectives in complex models. However, in practice it is advised to make the number of function evaluations as few as possible in finding an optimal solution [65]. Moreover, the value of design variables (objectives) are obtained by real or computational experiments, where the form of objective functions is not given explicitly in terms of objectives [66]. However, the time dependent (dynamic) multi-objective optimization due to relying on different moments is a very difficult task [40]. Therefore, in the current paper a hybridized multi-objective support vector regression (MSVR) incorporated

to firefly metaheuristic algorithm (FMA) for prediction of UCS and E was developed.

The population based stochastic FMA is a swarm intelligence method inspired by the flashing behavior of fireflies [61]. This trial and error procedure efficiently and simultaneously can be applied for solving the hardest optimization problems to find both global and local optima [64]. The performance of hybrid MSVR-FMA was examined by different error criteria and then compared with MSVR. The models were run using 222 datasets of different building stones including rock class, density (γ), porosity (n), P -wave velocity (V_p), water absorption (w) and point load index (I_s) from almost all over quarry locations of Iran. It was demonstrated that by applying the FMA, the success of correct classification rate for UCS and E from 81.2% and 79.5% were progressed to 88.6% and 84.1%, respectively. The comparison of different error criteria showed that MSVR-FMA as an accurate enough model can efficiently be applied to estimate the UCS and E . The main effective factors on predicted values were then recognized using different sensitivity analyses.

2 Mathematical configuration of MSVR

In classification, SVR is characterized by the use of kernels, sparse solution and control of the margin, and the number of support vectors. Hence, the output of SVR is found from the mapped support vectors through feature space and calculated weights using Lagrange multipliers and assigned biases (Fig. 1). As the output of SVR is a real number, thus in regression purposes a tolerance margin (ϵ) known as ϵ -insensitive loss function is set in the approximation [58]. This provides a symmetrically flexible tube of minimal radius around the estimated function in which the absolute values of errors less than a certain threshold are ignored both above and below the estimate. Consequently, the points outside the tube are penalized, but those within the tube, either above or below the function receive no penalty.

Considering to dimensionality of input and output spaces (d and Q), the output vector ($y_i \in R^Q$) subjected to a given set of training input data $\{(x_i, y_i)\}_{i=1, 2, \dots, n}$ ($x_i \in R^d$) is derived from minimizing of:

$$L_p(W, b) = L_p([w_1, \dots, w_Q], [b_1, \dots, b_Q]) = \underbrace{\frac{1}{2} \sum_{j=1}^Q w_j^2}_{\text{structural risk}} + C \sum_{i=1}^n \overbrace{L_v(y_i - (\phi(x_i)^T + w_j + b_j))_\epsilon}^{L(u_i)} \quad (1)$$

empirical risk

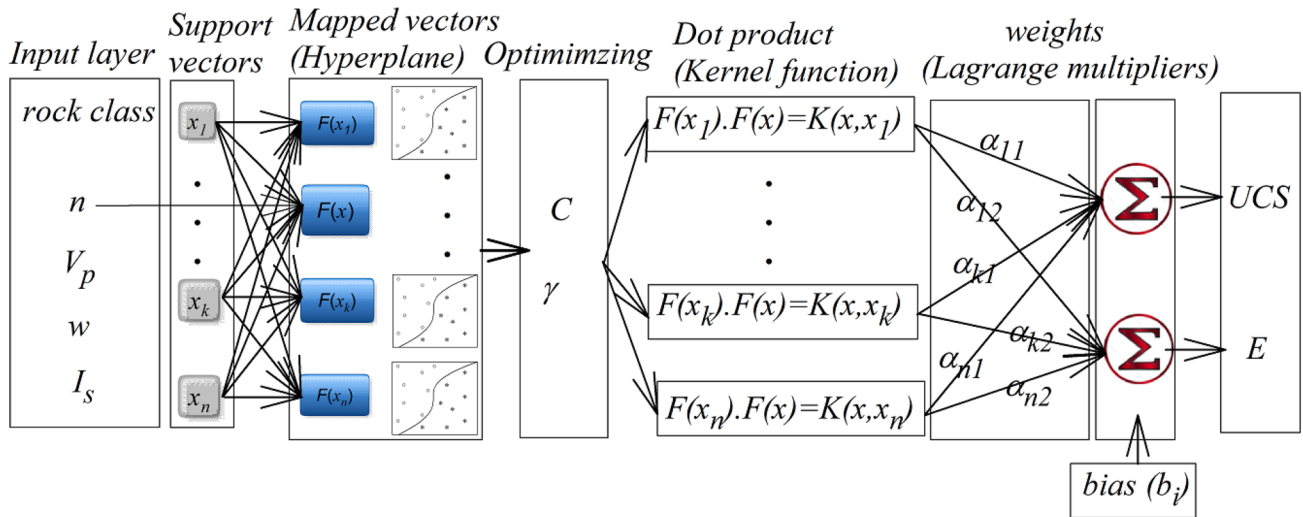


Fig. 1 Processing procedure in MSVR structure using support vector algorithm

$$u_i = \sqrt{e_i^T e_i}; \quad e_i^T = y_i^T - \phi^T(x_i) \mathbf{W} - \mathbf{b}^T \quad (2)$$

where; $L_p(\mathbf{W}, \mathbf{b})$ is the Lagrangian optimization function. w_j ($\mathbf{W} \in \mathbb{R}^{Q \times d}$) represents an $m \times m$ weighted matrix corresponds the model parameter and thus each $w_i \in \mathbb{R}^d$ is the predictor for y_i . The b_j ($\mathbf{b} \in \mathbb{R}^Q$; $j = \{1, \dots, Q\}$) and $e_k \in \mathbb{R}$ denote

On k th iteration (\mathbf{W}^k and \mathbf{b}^k), optimizing problem and constructing the quadratic approximation is expressed by:

$$L'_p(\mathbf{W}, \mathbf{b}) = \frac{1}{2} \sum_{j=1}^Q w_j^2 + C \left(\left[\sum_{i=1}^n L(u_i^k) + \frac{dL(u)}{du} \right]_{u_i^k} \frac{(e_i^k)^T}{u_i^k} [e_i - e_i^k] \right) \quad (4)$$

$$L''_p(\mathbf{W}, \mathbf{b}) = \frac{1}{2} \sum_{j=1}^Q w_j^2 + C \left(\left[\sum_{i=1}^n L(u_i^k) + \frac{dL(u)}{du} \right]_{u_i^k} \frac{u_i^2 - (u_i^k)^2}{2u_i^k} \right) = \frac{1}{2} \sum_{j=1}^Q w_j^2 + \frac{1}{2} \sum_{i=1}^n a_i u_i^2 + C\tau \quad (5)$$

the bias matrix term and error variables. The w and b can be obtained using the iterative reweighted trial error least squares procedure (IRWLS) to lead a matrix for each component that to be estimated [47, 48]. The term structure risk (regularization term) is used to control the smoothness or complexity of the function. The user specified constant $C > 0$ determines the trade-off between the empirical error and the amount up to deviations larger than ε [16]. The parameter ε should be tuned and is equivalent to the approximation accuracy in the training process and shows that the datasets in the range of $[\varepsilon, -\varepsilon]$ do not contribute to the empirical error [13]. $\phi(\cdot)$ is the feature space factor to provide nonlinear transformation to a higher dimension. The term $L(u_i)$ as the loss function using Taylor expansion is defined as:

$$|y_i - f(x_i)|_\varepsilon = \begin{cases} 0 & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{otherwise} \end{cases} \quad (3)$$

where $C\tau$ as a sum of constant term is independent either on \mathbf{W} or \mathbf{b} . Applying $\mathbf{W} = \mathbf{W}^k$ and $\mathbf{b} = \mathbf{b}^k$, provide the value and gradient for $L_p(\mathbf{W}, \mathbf{b})$ and $L'_p(\mathbf{W}, \mathbf{b})$. Thereby, $\nabla L'_p(\mathbf{W}^k, \mathbf{b}^k) = \nabla L_p(\mathbf{W}^k, \mathbf{b}^k) \cdot L'_p(\mathbf{W}^k, \mathbf{b}^k)$ is a lower bound of $L_p(\mathbf{W}, \mathbf{b})$ where $L_p(\mathbf{W}, \mathbf{b}) > L'_p(\mathbf{W}, \mathbf{b})$. The a_i , u_i^k and $(e_i^k)^T$ then can be calculated using:

$$a_i = \frac{C}{u_i^k} \frac{dL(u)}{du} \bigg|_{u_i^k} = \begin{cases} 0 & u_i^k < \varepsilon \\ \frac{2c(u_i^k - \varepsilon)}{u_i^k} & u_i^k \geq \varepsilon \end{cases} \quad (6)$$

$$u_i^k = e_i^k = \sqrt{(e_i^k)^T e_i^k}; \quad (e_i^k)^T = y_i^T - \phi(x_i)^T \mathbf{W}^k - (\mathbf{b}^k)^T \quad (7)$$

Linear combination of the training datasets can provide the best solution for optimizing the learning problem within the inner product of feature space kernel [56]:

$$w_j = \sum_{i=1}^n \phi(x_i) \beta_j = \Phi^T \beta_j, \quad 0 \leq \beta \leq C \quad (8)$$

$$\begin{bmatrix} K + D_a^{-1} & 1 \\ a^T K & 1^T a \end{bmatrix} \begin{bmatrix} \beta_j \\ b_j \end{bmatrix} = \begin{bmatrix} y_j \\ a^T y_j \end{bmatrix} \quad (9)$$

$$K(x_i, x_j) = \exp \left(-\frac{x_i - x_j^2}{2\sigma^2} \right) = \exp \left(-\gamma \|x_i - x_j\| \right) \text{ and } \gamma > 0 \quad (10)$$

Where $1 = [1, 1, \dots, 1]^T$ is an n dimensional column vector and $a = [a_1, \dots, a_n]^T$ shows an identity matrix. $(K)_{ij} = k(x_i, x_j)$ is the kernel matrix of two vectors x_i and x_j that can be easily evaluated [23]. The $\|\cdot\|$ correspond the Euclidean norm for vectors and γ denotes the variance of the radial basis function (RBF) kernel which controls the sensitivity of the kernel function. β_j is the parameter which should be computed by searching algorithm and depends on Lagrange multipliers. Hereby, the training datasets via the kernel function are moved into a higher dimension space where various kernel functions may produce different support vectors (Fig. 1). Therefore, the j th output of each new incoming vector x can be expressed as:

$$y_j = \Phi^T(x) \Phi^T \beta_j, \quad \Phi = [\phi(x_1), \dots, \phi(x_n)] \rightarrow y_j = \sum_{i=1}^n K(x, x_i) \beta_i^j + b^j \quad (11)$$

where $y_j = [y_{1j}, \dots, y_{nj}]^T$ is the outputs. Consequently, the final output (y) is computed by:

$$y = \Phi^T(x) \Phi^T \beta = K_x \beta, \quad k(x_i, x_j) = \phi^T(x_i) \phi(x_j) \quad (12)$$

where $\phi(x_i)$ and $\phi(x_j)$ are the projection of the x_i and x_j in feature space. The number of support vectors and biases are noted by n and b^j respectively. K_x is a vector that contains the kernel of the input vector x and the training datasets. RBF kernel (Eq. 10) has shown more promising results compared than other proposed kernels [33].

3 Firefly metaheuristic algorithm (FMA)

The FMA as a swarm intelligence population-based algorithm inspired by flashing behavior of fireflies [61] effectively can be applied to solve the hardest global and local optimization problems [64]. During the recent years the applicability of this algorithm has been modified. Gao et al. [25] improved this algorithm using particle filter and presented a powerful tool in solving visual tracking problems. Sayadi et al. [52] developed a powerful version of discrete FMA to deal with non-deterministic

polynomial-time scheduling problems. An efficient binary coded FMA to investigate the network reliability was proposed by Chandrasekaran and Simon [17]. Coelho et al. [21] proposed a chaotic FMA that outperformed other algorithms [22]. The studies of Yang [62] on this version of FMA showed that under different ranges of parameter, an enhanced performance using tuning can be achieved. The Lagrangian FMA is another proposed variant [51] for solving the unit commitment problem for a power system. An interesting multi-objective discrete FMA version or the economic emission load dispatch problem was proposed by Apostolopoulos and Vlacho [9]. Meanwhile, Arsuaga-Rios and Vega-Rodriguez [10] independently proposed another multi-objective FMA tool for minimizing energy consumption in grid computing. This version further was developed to solve multi-objective production scheduling systems [34]. Furthermore, a discrete variant of FMA for the multi-objective hybrid problems [37], an extended FMA for converting the single objective to multi-objective optimization in continuous design problems [63], and an enhanced multi-objective FMA in complex networks [8] were also presented.

As presented in Fig. 2, the primary concept for a firefly's flash is based on signal system to attract other fireflies which can be figured out using brightness (I), attractiveness (β) of fireflies i and j in the adjacent distance (r_{ij}), absorption coefficient (γ) and tradeoff constant to determine the random behavior of movement (α). The fireflies in this system subjected to trial-error procedure tend to move towards the brighter one and aim to find a new solution using the updated distance between two considered fireflies.

The I of each firefly represents the solution, s , as a proportion of the objective function [$I(s) \propto f(s)$]. The β is also

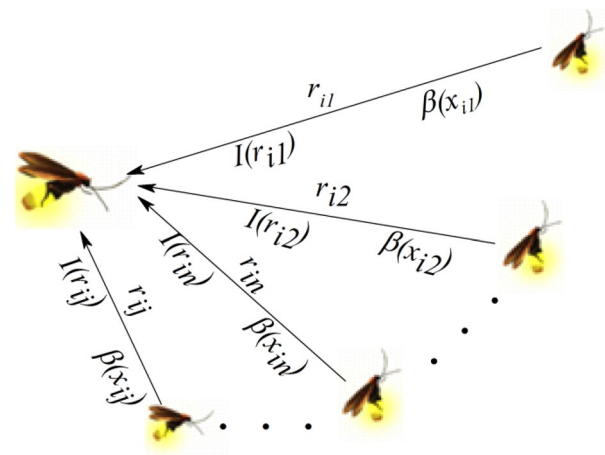


Fig. 2 The configuration of FMA and corresponding applied parameters

proportional to the intensity of visible light for adjacent fireflies in each distance coordinate, $I(r)$, as:

$$I(r) = I_0 e^{-\gamma r^2} \quad (13)$$

$$\beta = \beta_0 e^{-\gamma r^2} \quad (14)$$

The distance between any two s_i and s_j or i and j fireflies in an n -dimensional problem is expressed as the Euclidean or the Cartesian distance by:

$$r_{ij} = s_i - s_j = \sqrt{\sum_{k=1}^n (s_{ik} - s_{jk})^2} \quad (15)$$

where I_0 denotes the light intensity of the source. γ is the absorption coefficient with a decisive impact on the convergence speed that can theoretically capture any value from interval $\gamma \in [0, \infty)$ but in most optimizing problems typically varies within $[0.1-10]$. β_0 is the attractiveness at $r_{ij}=0$.

In each iteration of FMA, the fitness function (FT) of the optimal solution of each firefly will own its brightness. Therefore, searching for better FT corresponding to higher brightness level produces new solutions. This embedded iterative process will renew several times comparing to previous results, and only one new solution based on FT is kept. This iterative process can be expressed as:

$$s_i^{\text{new}} = s_i^t + \beta_0 e^{-\gamma r_{ij}^2} (s_j - s_i) + \alpha_t (\text{rand} - 0.5) \quad (16)$$

where α_t denotes the tradeoff constant to determine the random behavior of movement and varies in $[0, 1]$ interval. The rand function as a random number of solutions I and β_0 at zero distance is normally set to 1. s_j is a solution with lower FT than s_i and $(s_j - s_i)$ represents the updated step size.

Considering to variation of γ within $[0, \infty)$ interval, when $\gamma \rightarrow 0$ then $\beta_0 = \beta$ that express the standard particle swarm optimization (PSO). In the situation that $\gamma \rightarrow \infty$, the second term falls out from Eq. (16) which not only indicate random walk movement but also is essentially a parallel version of simulated annealing. Consequently, the FMA generally is controlled by three parameters γ , β , and α where in $\beta_0 = 0$, the movement is a simple random walk.

Depending on compared FT, the new solution (one or more than one or no new solution) between firefly i and other fireflies in the current population is described with:

$$s_i^{\text{new}} = \begin{cases} s_i & s_i = s_{\text{best}} \\ s_i^{\text{new}} & s_j = s_{\text{best}} \\ s_{ij}^{\text{new}} & \text{with FT}_{\text{best}} \text{ otherwise} \end{cases} \quad (17)$$

Therefore, if the considered solution i is also the global best solution, no new solution will be generated. If the best global solution in the population (n) belongs to firefly j , then only one new solution s_i^{new} is achieved, else, at least two better solutions than s_i in $n-1$ is available where the lowest FT (FT_{best}) is retained and others are discarded.

4 Acquired database

A database including 222 sets of rock class, density (γ), porosity (n), P -wave velocity (V_p), water absorption (w) and point load index (I_s) from 49 different quarry locations in Iran was assembled (Tables 1 and 2). The statistically analyzed datasets as well as calculated 95% confidence intervals of mean and median of provided datasets are presented in Table 3 and Fig. 3. According to suggested classification by the International society of rock mechanics [30], the majority of compiled datasets fall in the medium to high strength categories (Table 4). The components of processed database due to different units were then normalized within the range of $[0, 1]$ to produce dimensionless sets and improve the learning speed and model stability. These sets further were randomized into training (55%), testing (25%) and validation (20%). The rock classes including sedimentary, igneous, diagenetic and metamorphic were coded from 1 to 4, respectively.

5 Hybridized MSVR and system results

The structure of MSVR is developed through the input, intermediate and output layers subjected to a series of training experiments. As presented in Fig. 4, the MSVR was trained using

Table 1 Specification of acquired datasets

Rock class	Rock type	Number of samples
Igneous	Andesite, Granite, Dacite, Terakite, Vitric tuff, Ignemberite, Gabbro, Syenite	47
Sedimentary	Limestone, Travertine, Onyx travertine	148
Metamorphic-diagenetic	Dolomitic limestone, Marble	27

Table 2 Sample of compiled datasets

Rock type	Rock class	Location	γ (g/cm ³)	n (%)	V _p (km/s)	Is (Mpa)	w (%)	UCS (Mpa)	E (Gpa)
Travertine	1	Azarshahr	2.48	7.41	3.95	3.59	1.43	33.60	4.50
Onyx travertine	1	Hamedan	2.32	7.18	3.75	2.37	2.87	52.96	5.1
Travertine	1	Atashkooh	2.47	4.20	4.60	4.20	1.47	49.30	6.61
Travertine	1	Abbasabad	2.43	4.86	4.15	3.64	2.69	41.30	5.56
Travertine	1	Firuzkuh	2.38	3	5.01	5.11	0.88	59.9	5.11
Andesite	2	Yazd	2.72	0.87	5.79	8.82	0.32	126.00	46.40
Granite	2	Nehbandan	2.8	0.32	6.10	10.67	0.23	145.00	36.50
Granite	2	Hamedan	2.71	1.36	4.39	10.78	0.46	125.64	52.76
Syenite	2	Iranshahr	2.54	1.88	5.30	6.27	0.59	91.00	20.00
Gabbro	2	Natanz	2.82	0.26	6.11	14.98	0.23	147.00	81.40
Vitric tuff	2	Farahzad-Tehran	2.18	1.15	3.72	7.20	5.61	101.00	32.4
Marble	4	Sirjan	2.69	1.10	5.21	3.30	0.16	64.00	12.46
Dolomitic limestone	3	Dehbid	2.71	0.35	5.89	5.64	0.20	103.50	17.44
Marble	4	Neyriz	2.7	0.37	5.83	5.53	0.18	101.80	13.58
Limestone	1	Sarvestan	2.69	0.87	4.85	2.66	4.36	47.00	7.33
Limestone	1	Kharama	2.70	0.27	5.87	4.03	1.94	77.00	12.27
Limestone	1	Firoozabad	2.56	9.76	3.42	2.01	6.38	32.90	8.010

Is, point load index; V_p, P-wave velocity; n , porosity; w , water absorption; γ , density

Table 3 Descriptive statistics of acquired datasets

Variable	Mean	Mean st.e	St. Dev	Min	Median	Max	Skewness
Rock class	1.51	0.058	0.846	1	1	4	1.65
γ (g/cm ³)	2.5919	0.0116	0.1733	2.18	2.585	3.06	0.39
n %	6.405	0.409	6.09	0.15	4.435	31.4	1.11
V _p (m/s)	4.7661	0.0651	0.9698	2.154	5.01	6.82	-0.62
w (%)	4.274	0.199	2.968	0.3	3.36	15.12	1.54
Is (MPa)	2.859	0.266	3.958	0.07	1.38	16.16	1.94
UCS (Mpa)	58.86	2.52	37.51	4.75	47.23	193	0.8
E (Gpa)	14.2	1.15	17.1	0.5	6.81	89.4	2.1

Mean st.e. Mean standard error; St. Dev Standard Deviation

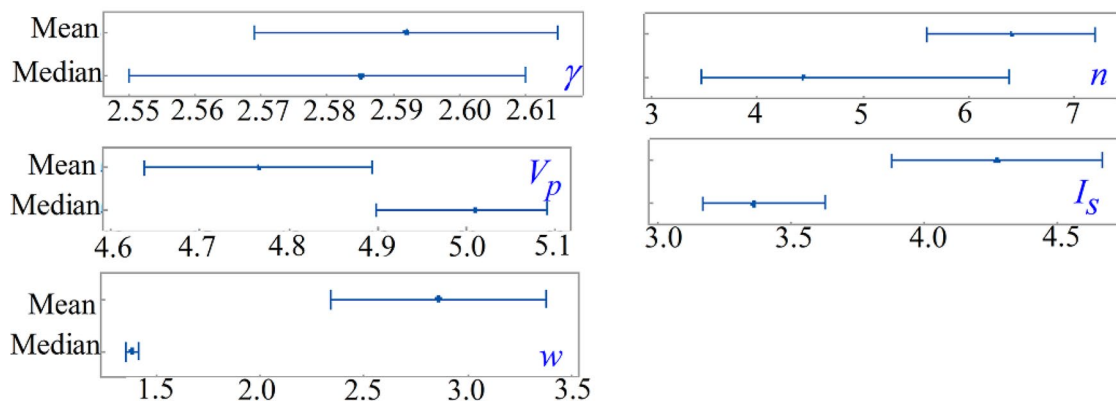
**Fig. 3.** 95% confidence intervals of mean and median for the employed variables

Table 4 Strength classification based on ISRM [30] for provided database

Strength description	UCS (MPa)	No of samples	Percentage
Extremely low strength	< 1	0	0
Very low strength	1–5	1	0.45
Low strength	5–25	28	12.61
Medium strength	25–50	88	39.64
High strength	50–100	70	31.53
Very high strength	100–250	35	15.77
Extremely high strength	> 250	0	0

iterative reweighted trial error least squares procedure (IRWLS) [47, 48]. Weighting is based on the output of the true objective function and thus the reweighting scheme is considered as a feedback control. The accuracy of MSVR output depends on the appropriate regularized C , ε as well as γ and σ (Min and Lee [39], but no unified procedure for estimating these parameters are accepted. To tune the optimum C and γ , numerous combinations of these parameters with step sizes of $2^{0.2}$ and $2^{0.1}$ over \log_2 using the LIBSVM code in Matlab (Chang and Lin [18] were examined. The FT of optimum parameters in the training process then was evaluated using separate validation datasets or cross-validation technique [19, 32] using:

$$f(m) = \text{RMSE}_{\text{training data}}^{\text{optimization}} + \text{RMSE}_{\text{validation data}}^{\text{optimization}} \quad (18)$$

Table 5 Adjusting the FMA parameters

Parameter	Initial value	Final value	Incre- ment step size
α	0.01	1.0	0.05
β_0	0.1	1.0	0.05
γ	0	1.0	0.05
n	10	100	5

where RMSE expresses the root mean square error.

Due to ability of the FMA in control the parameters for effective balancing [20], it was applied to improve the quality of the initial population and optimizing the C and σ . Refer to Fig. 4, the main loop of FMA is controlled by the maximum number of generations (Max Gen). This loop using a gen-counter parameter (t), calculates the new values for the randomization parameter (α) through the functions $\Delta = 1 - 10^{-4}/0.9^{1/\text{Max Gen}}$ and $\alpha^{(t+1)} = 1 - \Delta \cdot \alpha^{(t)}$. Δ determines the step size of changing parameter $\alpha^{(t+1)}$ and is descended with the increasing of t . Then, the new solution $s_i^{(t)}$ is evaluated based on a fitness function $f(s^{(t)})$. With respect to the fitness function, $f(s_i^{(t)})$ is ordered ascending the solutions $s_i^{(t)}$ for n populations, where $s_i^{(t)} = S(x_i^{(t)})$ and thus the best solution $s^* = s_0^{(t)}$ is determined in the population $P^{(t)}$. The FMA parameters ($n, \alpha, \beta_0, \gamma$) were obtained considering

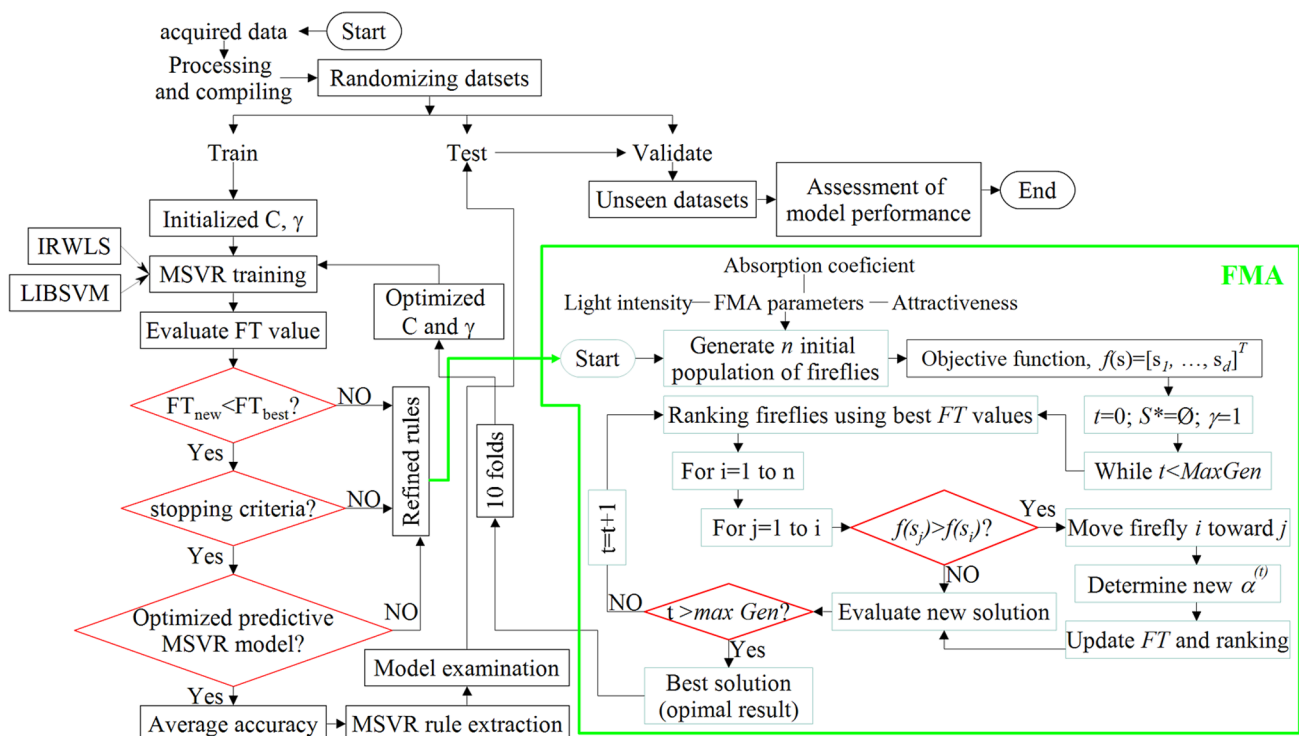


Fig. 4 Flowchart of model construction and assessing the optimized structure

Table 6 A brief result of analyzing process to tune FMA parameters

$\beta_0=0.5, \gamma=1$		$\alpha=0, \gamma=1$		$\alpha=0.1, \beta_0=0.5$	
α	RMSE	β_0	RMSE	γ	RMSE
0.05	0.143	0.05	0.114	0.05	0.169
0.10	0.120	0.10	0.119	0.1	0.170
0.15	0.102	0.15	0.128	0.15	0.168
0.20	0.075	0.20	0.125	0.2	0.169
0.25	0.143	0.25	0.120	0.25	0.169
0.30	0.168	0.30	0.091	0.3	0.171
0.35	0.185	0.35	0.072	0.35	0.166
0.40	0.179	0.40	0.058	0.4	0.161
0.45	0.160	0.45	0.038	0.45	0.161
0.50	0.157	0.50	0.027	0.5	0.157
0.55	0.168	0.55	0.067	0.55	0.159
0.60	0.169	0.60	0.125	0.6	0.150
0.65	0.187	0.65	0.143	0.65	0.145
0.70	0.162	0.70	0.139	0.7	0.148
0.75	0.170	0.75	0.129	0.75	0.158
0.80	0.158	0.81	0.147	0.8	0.159
0.85	0.178	0.85	0.158	0.85	0.161
0.90	0.193	0.90	0.157	0.9	0.134
0.95	0.194	0.95	0.137	0.95	0.080
1.00	0.161	1.00	0.143	1.00	0.025

variations in the weight results, meanwhile, one was tested the other one was fixed. The number of fireflies was obtained according to the convergence history of the iteration process. Tables 5 and 6 summarize the information about initial, final values and rate of variation in each parameter as well as a sample of series efforts in parametric analyzing. The number of iteration and corresponding fireflies were found through the convergence history of different populations (Fig. 5). The results showed that 1000, 30, 1, 0.2, 0.05, 0.2 and 0.5 corresponding to the number of iterations, number of fireflies, γ ,

β , Δ , α and β_0 can be selected as the most appropriate tuned parameters in FMA.

In this study, for a stable learning process and reduce the computational effort the *MSVR* was managed using a quadratic loss function with the value of 0.1 and *RBF* kernel function subjected to tenfold cross-validation. In the cross-validation method (Fig. 6), the entire training dataset is randomly split into roughly equal subset folds. For K times, each of the folds can be chosen for test data and the remaining is used as training sets. The errors should be less than ε and any deviation larger than this is not accepted. As reflected in Table 7, the optimum values for C and γ were then selected from the lowest error and highest correlation coefficient (R^2) in tenfold cross-validation. Accordingly, the performance of the hybridized *MSVR*-FMA using adjusted parameters comparing to measured values was checked and presented in Fig. 7.

6 Validation and discussion

The correct classification rate (CCR) is a leading assessment metric in discriminant analysis. This criterion can be extracted from the confusion matrix [54] as an unambiguous

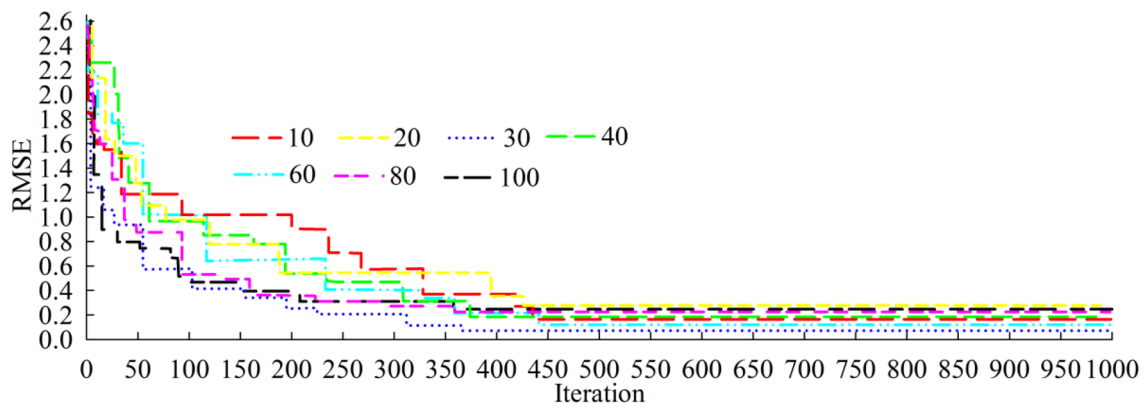
**Fig. 6** K -fold cross-validation using split randomized dataset**Fig. 5** Convergence history subjected to different firefly populations

Table 7 Determining the optimal parameters of MSVR in tenfold cross-validation procedure

Fold number	Train		Test		$C (\times 10^4)$	$\sigma (\times 10^{-1})$
	R^2	RMSE	R^2	RMSE		
1	0.92	0.795	0.96	0.528	2.181	120.28
2	0.94	0.517	0.95	0.465	1.742	66.822
3	0.93	0.628	0.92	0.586	0.098	524.754
4	0.95	0.543	0.90	0.767	1.014	74.297
5	0.92	0.548	0.94	0.475	9.335	978.906
6	0.93	0.634	0.96	0.579	1.756	55.156
7	0.96	0.425	0.95	0.402	0.423	44.278
8	0.93	0.701	0.94	0.604	8.077	267.428
9	0.95	0.696	0.92	0.675	10.152	83.203
10	0.91	0.716	0.93	0.695	1.948	211.572
Average	0.93	0.620	0.94	0.577	3.672	242.669

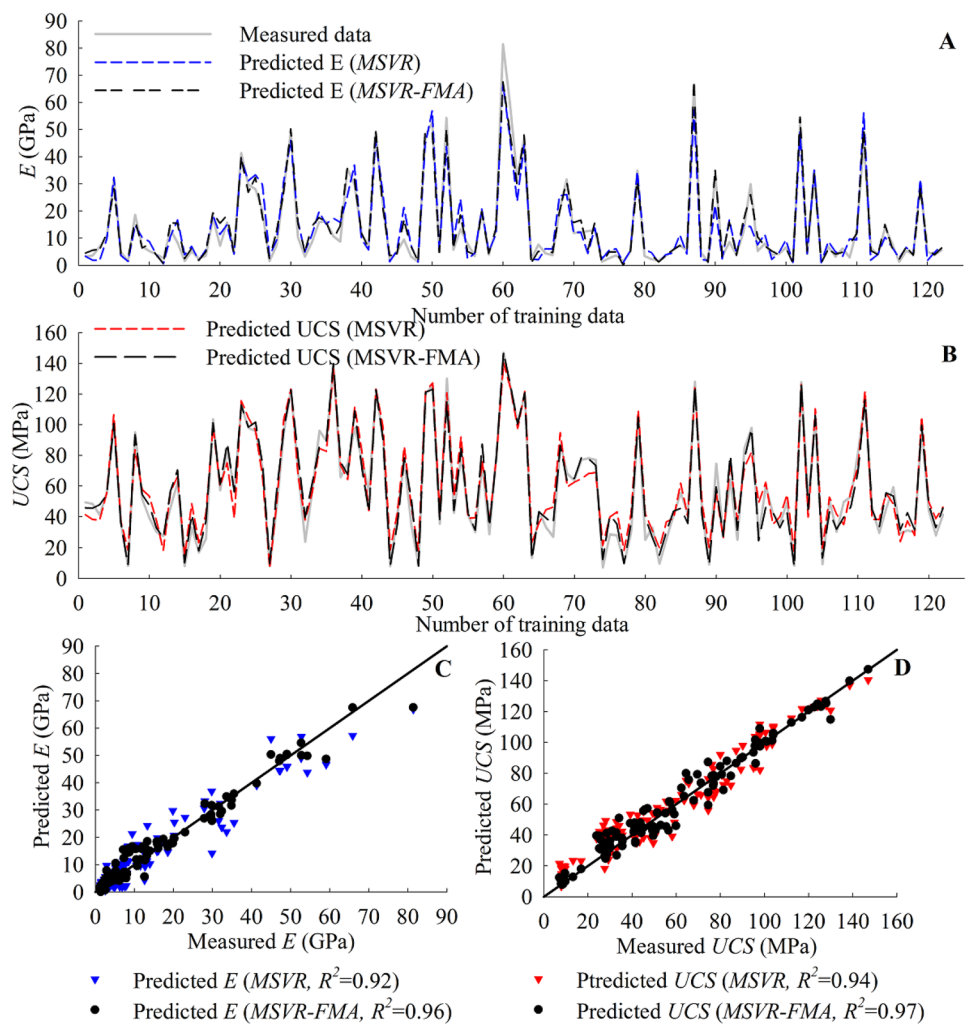
Fig. 7 Predictability level of optimum and hybridized MSVR in training stage for E (a, c) and UCS (b, d)

table layout method to present the predictability of machine learning classifier. Referring to conducted confusion matrix (Table 8), the calculated correct classification rate (CCR) showed 10.27% and 5.47% improvement in predictability

of UCS and E using hybrid MSVR-FMA (Table 9). These results reflect the significant influence of incorporated FMA on the accuracy progress of prediction process.

Table 8 Confusion matrix of applied models using validation datasets

Target output	Network output MSVR (<i>E</i>)												Results		
	<0.91	0.91–9.09	9.09–17.27	17.27–25.45	25.45–33.63	33.63–41.81	41.81–49.99	49.99–58.17	58.17–66.35	66.35–74.53	74.53–82.72	>82.72	Total	True	False
0.91–9.09	0	22	2	0	0	0	0	0	0	0	0	0	24	22	2
9.09–17.27	0	1	7	1	0	0	0	0	0	0	0	0	9	7	2
17.27–25.45	0	0	0	1	1	0	0	0	0	0	0	0	2	1	1
25.45–33.63	0	0	0	1	1	1	0	0	0	0	0	0	3	1	2
33.63–41.81	0	0	0	0	1	1	0	0	0	0	0	0	2	1	1
41.81–49.99	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0
49.99–58.17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58.17–66.35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
66.35–74.53	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0
74.53–82.72	0	0	0	0	0	0	0	1	0	1	0	0	2	1	1
Note	0	23	9	3	3	2	1	0	1	1	1	0	44	35	9
Network output hybrid MSVR-FMA (<i>E</i>)															
0.91–9.09	1	21	2	0	0	0	0	0	0	0	0	0	24	21	3
9.09–17.27	0	0	7	2	0	0	0	0	0	0	0	0	9	7	2
17.27–25.45	0	0	0	2	0	0	0	0	0	0	0	0	2	2	0
25.45–33.63	0	0	0	1	1	1	0	0	0	0	0	0	3	1	2
33.63–41.81	0	0	0	0	2	0	0	0	0	0	0	0	2	2	0
41.81–49.99	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0
49.99–58.17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58.17–66.35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
66.35–74.53	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0
74.53–82.72	0	0	0	0	0	0	0	0	0	2	0	0	2	2	0
Note	1	21	10	4	3	2	0	1	1	2	0	0	44	37	7
Target output	Network output MSVR (UCS)												Results		
	<6.08	6.08–23.14	23.14–40.2	40.2–57.26	57.26–74.32	74.32–91.38	91.38–108.44	108.44–125.5	125.5–142.56	142.56–159.62	159.62–176.70	>176.70	Total	True	False
6.08–23.14	0	7	0	0	0	0	0	0	0	0	0	0	7	7	0
23.14–40.2	0	0	2	2	0	0	0	0	0	0	0	0	4	2	2
40.2–57.26	0	0	3	10	0	0	0	0	0	0	0	0	13	10	3
57.26–74.32	0	0	0	0	4	0	0	0	0	0	0	0	4	4	0
74.32–91.38	0	0	0	0	1	2	0	0	0	0	0	0	3	2	1
91.38–108.44	0	0	0	0	0	6	0	0	0	0	0	0	6	6	0
108.44–125.5	0	0	0	0	0	1	2	0	0	0	0	0	3	2	1

Table 8 (continued)

Target output	Network output MSVR (UCS)												Results		
	< 6.08	6.08–23.14	23.14–40.2	40.2–57.26	57.26–74.32	74.32–91.38	91.38–108.44	108.44–125.5	125.5–142.56	142.56–159.62	159.62–176.70	> 176.70	Total	True	False
125.5–142.56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
142.56–159.62	0	0	0	0	0	0	0	0	1	2	0	0	3	2	1
159.62–176.70	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0
Note	0	7	5	12	5	2	7	2	1	2	1	0	44	36	8
Network output hybrid MSVR-FMA (UCS)															
6.08–23.14	0	7	0	0	0	0	0	0	0	0	0	0	7	7	0
23.14–40.2	0	0	3	1	0	0	0	0	0	0	0	0	4	3	1
40.2–57.26	0	0	2	11	0	0	0	0	0	0	0	0	13	11	2
57.26–74.32	0	0	0	0	4	0	0	0	0	0	0	0	4	4	0
74.32–91.38	0	0	0	0	0	3	0	0	0	0	0	0	3	3	0
91.38–108.44	0	0	0	0	0	0	6	0	0	0	0	0	6	6	0
108.44–125.5	0	0	0	0	0	0	1	2	0	0	0	0	3	2	1
125.5–142.56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
142.56–159.62	0	0	0	0	0	0	0	0	1	2	0	0	3	2	1
159.62–176.70	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0
Note	0	7	5	12	4	3	7	2	1	2	1	0	44	39	5

The area under the curve (AUC) of receiver operating characteristic (ROC) is one of the most important graphical metrics in performance and diagnostic ability of a classifier system. ROC is a probability curve that summarizes the trade-off between the true and false positive rates (TPR and FPR) for a predictive model at various threshold settings while AUC represents the degree of separability.

Therefore, AUC_{ROC} expresses the capability and strength of the model in distinguishing classes (Fig. 8a). In machine learning, precision shows the capability of a classification model in identifying only the relevant data points, while recall monitors all the related cases within a dataset. An optimal combination of precision and recall can be interpreted using F_1 -score as:

Table 9 CCR of optimized models for validation and test datasets

Model	CCR (%)				Progress (%)			
	Test		Validate		MSVR		MSVR-FMA	
	UCS	<i>E</i>	UCS	<i>E</i>	UCS	<i>E</i>	UCS	<i>E</i>
MSVR	80.3	76.8	81.2	79.5	1.11	3.4	–	–
MSVR-FMA	85.7	82.1	88.6	84.1	–	–	3.27	2.38

	MSVR		MSVR-FMA	
	UCS	<i>E</i>	UCS	<i>E</i>
MSVR				
UCS	–	2.1%	–8.35%	–3.45%
<i>E</i>	–2.1%	–	–10.27%	–5.47%
MSVR-FMA				
UCS	8.35%	10.27%	–	5.08%
<i>E</i>	3.45%	5.47%	–5.08%	–

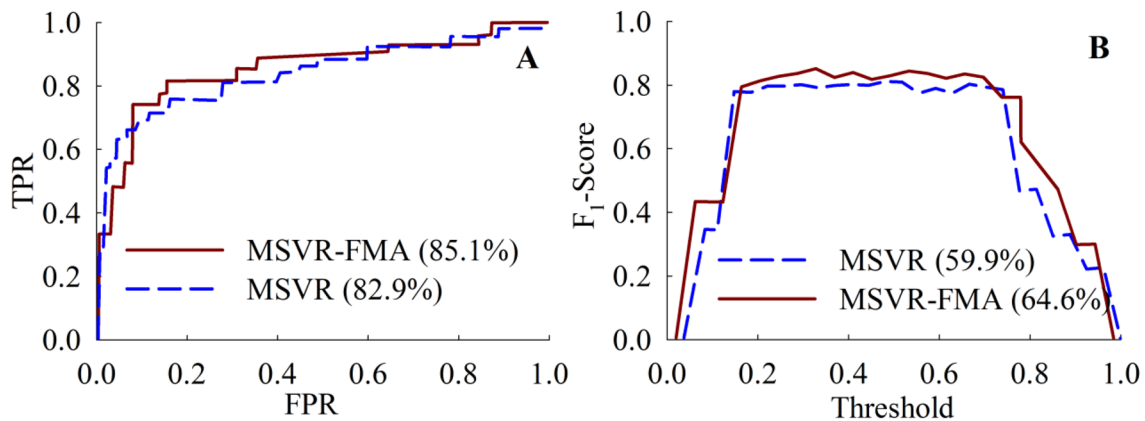


Fig. 8 Analyzed performance of MSVR and hybrid MSVR-FMA using AUC_{ROC} (a) and F_1 -score criteria (b)

Table 10 Comparison of statistical criteria to evaluate the MSVR and MVR models for all used datasets

Criteria	Abbreviation	MSVR-FMA		MSVR	
		UCS	<i>E</i>	UCS	<i>E</i>
Mean absolute percentage error	MAPE	8.21	7.46	9.51	10.32
Root mean squared error	RMSE	0.302	0.267	0.384	0.330
Mean absolute deviation	MAD	2.18	1.65	4.01	3.04
Variance account for	VAF	95.80	93.15	92.36	89.77
Index of agreement	IA	0.88	0.88	0.84	0.82
Coefficient of determination	R^2	0.98	0.97	0.95	0.93

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (19)$$

This criterion expresses a harmonic mean that can be used instead of a simple average. It avoids extreme values and thus is used in imbalanced classes when the false negatives and false positives are crucial [28]. High precision and low recall express extremely accurate model, but it misses a significant number of instances that are difficult to classify. In optimal recall and precision values, the F_1 -score of a balanced classification model tends to be maximize. This situation reflects veracity (correctly classified data) and robustness (not miss significant instances) of the classifier. In optimizing of the classifier to increase one and disfavor the other, the harmonic mean shows quick decreasing when both precision and recall are equal (Fig. 8b).

The performance of the presented models and forecasted outputs were also pursued using statistical error indices as reflected in Table 10. The formulation of these indices can widely be found in statistical textbooks. The MAPE is one of the most popular indexes for description of accuracy and size of the forecasting error. The MAD reflects the size of error in the same units as the data, and reveals that high predicted values cause higher error rates. The generic IA [60] indicates the compatibility of modeled and observations. The VAF as an intrinsically connected index between predicted and actual values is a representative of model performance. Therefore, higher values of VAF, IA and R^2 as well as smaller values of MAPE, MAD and RMSE are interoperated as better model performance (Table 10).

Sensitivity analyses cab express the influence of input parameters on predictability level and provide robust

calibrated models in the presence of uncertainty [12]. This implies that removing the least effective inputs may lead to better results. The importance of input variables using the cosine amplitude (CAM) and partial derivative (PaD) [26] are calculated as:

$$\text{CAM}_{ij} = \frac{\sum_{k=1}^m (x_{ik} \times x_{jk})}{\sqrt{\sum_{k=1}^m x_{ik}^2 \sum_{k=1}^m x_{jk}^2}}; \quad \text{PaD}_i = \frac{\sum_p \left(\frac{\partial O_p^k}{\partial x_i^p} \right)^2}{\sum_i \sum_p \left(\frac{\partial O_p^k}{\partial x_i^p} \right)^2} \quad (20)$$

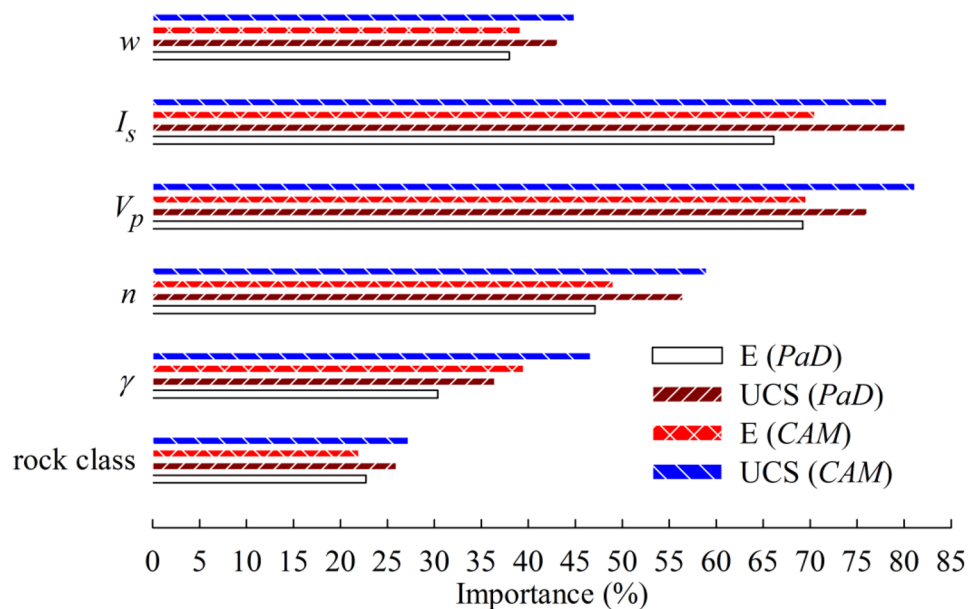
where CAM_{ij} and PaD_i express the importance (contribution) of i th variable. x_i and x_j denote elements of data pairs. O_p^k and x_i^p are output and input values for pattern P , and SSD_i is the sum of the squares of the partial derivatives respectively.

The results of CAM and PaD (Fig. 9) showed that I_s , V_p and n are the main effective factors on predicted UCS and E , while the rock class and γ expressed the least influences.

7 Conclusion

Due to the heterogeneity of rocks and dependency of strength parameters (UCS and E) to different physical and mechanical properties, conducting reliable and accurate predictive models is great of interest. In this paper, a new MSVR using 222 datasets of rock class, density (γ), porosity (n), P -wave velocity (V_p), water absorption (w) and point load index (I_s) for a wide variety of quarried rocks in Iran was developed. To enhance the progress and improve the efficiency, the MSVR successfully was hybridized with FMA. The results showed that by applying the FMA, the characterized internal

Fig. 9 Influence of input parameters on predicted UCS and E using different sensitivity analyses



properties of MSVR were optimized. Hybridizing procedure revealed that the CCR for UCS from 81.2% was promoted to 88.6% in MSVR-FMA. Similarly, for E this criterion was updated from 79.5 to 84.1%. These values indicate for 8.35% and 5.47% improvement in predictability level of UCS and E in MSVR-FMA. Investigating the robustness of models using AUC_{ROC} , F_1 -score exhibited superior performance in MSVR-FMA (85.1%) than MSVR (82.9%). The figured out accuracy performance of both classifiers using statistical error indices represented higher reliability in MSVR-FMA. According to evaluated criteria, 21.35% and 16.36% improvements in RMSE for UCS and E subjected to hybrid model was observed. Correspondingly, progress of 3.1% (UCS) and 4.16% (E) in MSVR-FMA was raised. Calculated IA showed that the MSVR-FMA with 5.54% and 6.82% progress for UCS and E is more compatible than MSVR. The implemented sensitivity analyses showed that I_s , V_p and n are the most effective factors on both UCS and E . This ranking can be interpreted with previous empirical correlations which mostly have been established by these three factors. The accuracy level of predicted outputs approved that the hybrid MSVR-FMA can efficiently utilizes a promising and superior alternative for the purpose of rock strength predictions in designing of construction projects.

References

1. Abbaszadeh Shahri A, Gheirati A, Espersson M (2014) Prediction of rock mechanical parameters as a function of P-wave velocity. *Int Res J Earth Sci* 2(9):7–14
2. Abbaszadeh Shahri A, Larsson S, Johansson F (2016) Updated relations for the uniaxial compressive strength of marlstones based on P-wave velocity and point load index test. *Innov Infrastruct Solut* 1:17. <https://doi.org/10.1007/s41062-016-0016-9>
3. Abbaszadeh Shahri A (2016) Assessment and prediction of liquefaction potential using different artificial neural network models—a case study. *Geotech Geol Eng* 34(3):807–815. <https://doi.org/10.1007/s10706-016-0004-z>
4. Abbaszadeh Shahri A, Larsson S, Renkel C (2020) Artificial intelligence models to generate visualized bedrock level: a case study in Sweden. *Model Earth Syst Environ*. <https://doi.org/10.1007/s40808-020-00767-0>
5. Abdi Y, Garavand AT, Sahamieh RZ (2018) Prediction of strength parameters of sedimentary rocks using artificial neural networks and regression analysis. *Arab J Geosci* 11:587. <https://doi.org/10.1007/s12517-018-3929-0>
6. Aboutaleb Sh, Behnia M, Bagherpour R, Bluekian B (2018) Using non-destructive tests for estimating uniaxial compressive strength and static Young's modulus of carbonate rocks via some modeling techniques. *Bull Eng Geol Environ* 77(4):1717–1728. <https://doi.org/10.1007/s10064-017-1043-2>
7. Alemdag S, Gurocak Z, Cevik A, Cabalar AF, Gokceoglu C (2016) Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming. *Eng Geol* 203:70–82. <https://doi.org/10.1016/j.enggeo.2015.12.002>
8. Amiri B, Hossain L, Crawford JW, Wigand RT (2013) Community detection in complex networks: multi-objective enhanced firefly algorithm. *Knowl Based Syst* 46(1):1–11. <https://doi.org/10.1016/j.knsys.2013.01.004>
9. Apostolopoulos T, Vlachos A (2011) Application of the firefly algorithm for solving the economic emissions load dispatch problem. *Int J Combin*. <https://doi.org/10.1155/2011/523806>
10. Arsuaga-Rios M, Vega-Rodriguez MA (2012) Multi-objective firefly algorithm for energy optimization in grid environments. In: Dorigo M et al (eds) *Swarm intelligence*. ANTS 2012. Lecture notes in computer science, vol 7461. Springer, Berlin. https://doi.org/10.1007/978-3-642-32650-9_41
11. Asheghi R, Abbaszadeh Shahri A, Khorsand Zak M (2019) Prediction of strength index parameters of different rock types using hybrid multi output intelligence model. *Arab J Sci Eng* 44(10):8645–8659. <https://doi.org/10.1007/s1336-019-04046-8>
12. Asheghi R, Hosseini SA, Sanei M, Abbaszadeh Shahri A (2020) Updating the neural network sediment load models using different sensitivity analysis methods—a regional application. *J Hydroinf*. <https://doi.org/10.2166/hydro.2020.098>
13. Awad M, Khanna R (2015) Support vector regression. In: *Efficient learning machines*. Apress, Berkeley, pp 67–80. https://doi.org/10.1007/978-1-4302-5990-9_4
14. Beiki M, Majdi A, Givshad AD (2013) Application of genetic programming to predict the uniaxial compressive strength and elastic modulus of carbonate rocks. *Int J Rock Mech Min Sci* 63:159–169. <https://doi.org/10.1016/j.ijrmm.2013.08.004>
15. Behnia D, Behnia M, Shahriar K, Goshtasbi K (2017) A new predictive model for rock strength parameters utilizing GEP method. *Procedia Eng* 191:591–599. <https://doi.org/10.1016/j.proeng.2017.05.222>
16. Bennett KP, Mangasarian OL (1992) Robust linear programming discrimination of two linearly inseparable sets. *Optim Methods Softw* 1:23–34. <https://doi.org/10.1080/10556789208805504>
17. Chandrasekaran K, Simon SP (2012) Network and reliability constrained unit commitment problem using binary real coded firefly algorithm. *Int J Electr Power Energy Syst* 42(1):921–932. <https://doi.org/10.1016/j.ijepes.2012.06.004>
18. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*. <https://doi.org/10.1145/1961189.1961199>
19. Cherkassky V, Mulier F (1998) *Learning from data: concepts, theory and methods*. Wiley, New York
20. Coelho LDS, Mariani VC (2013) Improved firefly algorithm approach applied to chiller loading for energy conservation. *Energy Build* 59:273–278. <https://doi.org/10.1016/j.enbui.2012.11.030>
21. Coelho LS, de Andrade Bernert DL, Mariani VC (2011) A chaotic firefly algorithm applied to reliability-redundancy optimization. In: 2011 IEEE Congress of Evolutionary Computation (CEC'11), 12117561, New Orleans, USA, pp 517–521. <https://doi.org/10.1109/CEC.2011.5949662>
22. Coelho LS, Mariani VC (2012) Firefly algorithm approach based on chaotic Tinkerbell map applied to multivariable PID controller tuning. *Comput Math Appl* 64(8):2371–2382. <https://doi.org/10.1016/j.camwa.2012.05.007>
23. Courant R, Hilbert D (1953) *Methods of mathematical physics*. Interscience, New York
24. Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik VN (1997) Support vector regression machines. In: *proceedings of NIPS'96, the 9th international conference on neural information processing systems*, pp 155–161
25. Gao ML, Li LL, Sun XM, Yin LJ, Li HT, Luo DS (2015) Firefly algorithm based particle filter method for visual tracking. *Optik* 126:1705–1711
26. Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial

- neural network models. *Ecol Model* 160(3):249–264. [https://doi.org/10.1016/S0304-3800\(02\)00257-0](https://doi.org/10.1016/S0304-3800(02)00257-0)
27. Ghaderi A, Abbaszadeh Shahri A, Larsson S (2019) An artificial neural network based model to predict spatial soil type distribution using piezocone penetration test data (CPTu). *Bull Eng Geol Environ* 78:4579–4588. <https://doi.org/10.1007/s10064-018-1400-9>
 28. Hand D, Christen P (2018) A note on using the F-measure for evaluating record linkage algorithms. *Stat Comput* 28(3):539–547. <https://doi.org/10.1007/s11222-017-9746-6>
 29. Imbens GW, Lemieux T (2008) Regression discontinuity designs: a guide to practice. *J Econ* 142(2):615–635
 30. ISRM (1981) Rock characterization testing and monitoring. In: Brown ET (ed) ISRM suggested methods, International Society of Rock Mechanics. Pergamon Press, Oxford
 31. Jahed Armaghani D, Mohamad ET, Momeni E, Monjezi M, Narayanasamy MS (2016) Prediction of the strength and elasticity modulus of granite through an expert artificial neural network. *Arab J Geosci* 9:48. <https://doi.org/10.1007/s12517-015-2057-3>
 32. Kavousi-Fard A, Samet H, Marzbani F (2014) A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst Appl* 41(13):6047–6056
 33. Li X, Chen X, Yan Y, Wei W, Wang J (2014) Classification of EEG signals using a multiple kernel learning support vector machine. *Sensors* 14(7):12784–12802. <https://doi.org/10.3390/s140712784>
 34. Li HM, Ye CM (2012) Firefly algorithm on multi-objective optimization of production scheduling system. *Adv Mech Eng Appl* 3(1):258–262
 35. Madhubabu N, Singh PK, Kainthola A, Mahanta B, Tripathy A, Singh TN (2016) Prediction of compressive strength and elastic modulus of carbonate rocks. *Measurement* 88:202–213. <https://doi.org/10.1016/j.measurement.2016.03.050>
 36. Mahdiabadi N, Khanlari G (2019) Prediction of uniaxial compressive strength and modulus of elasticity in calcareous mudstones using neural networks, fuzzy systems, and regression analysis. *Period Polytech Civil Eng* 63(1):104–114. <https://doi.org/10.3311/PPci.13035>
 37. Marichelvam MK, Prabakaran T, Yang XS (2013) A discrete firefly algorithm for the multi-objective hybrid flowshop scheduling problems. *IEEE Trans Evol Comput* 18(2):301–305. <https://doi.org/10.1109/TEVC.2013.2240304>
 38. Matin SS, Farahzadi L, Makaremi S, Chelgani SC, Sattari GH (2018) Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. *Appl Soft ing* 70:980–987. <https://doi.org/10.1016/j.asoc.2017.06.030>
 39. Min J, Lee Y (2005) Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Syst Appl* 28(4):603–614. <https://doi.org/10.1016/j.eswa.2004.12.008>
 40. Jiang M, Hu W, Qiu L, Shi M, Tan KC (2018) Solving dynamic multi-objective optimization problems via support vector machine. In: 2018 tenth international conference on advanced computational intelligence (ICACI). <https://doi.org/10.1109/ICACI.2018.8377567>
 41. Momeni E, Jahed Armaghani D, Hajihassani M, For M, Amin M (2015) Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement* 60:50–63. <https://doi.org/10.1016/j.measurement.2014.09.075>
 42. Moradian ZA, Behnia M (2009) Predicting the uniaxial compressive strength and static Young's modulus of intact sedimentary rocks using the ultrasonic test. *Int J Geomech* 9(1):14–19. [https://doi.org/10.1061/\(ASCE\)1532-3641\(2009\)9:1\(14\)](https://doi.org/10.1061/(ASCE)1532-3641(2009)9:1(14))
 43. Najibi AR, Ghafoori M, Lashkaripour GR, Asef MR (2015) Empirical relations between strength and static and dynamic elastic properties of Asmari and Sarvak limestones, two main oil reservoirs of Iran. *J Petrol Sci Eng* 126:78–82. <https://doi.org/10.1016/j.petrol.2014.12.010>
 44. Nocedal J, Wright SJ (1999) Numerical optimization. Springer-Verlag, New York
 45. Palchik V (2011) On the ratios between elastic modulus and uniaxial compressive strength of heterogeneous carbonate rocks. *Rock Mech Rock Eng* 44(1):121–128. <https://doi.org/10.1007/s00603-010-0112-7>
 46. Pedhazur EJ (1997) Multiple regression in behavioral research, 3rd edn. Harcourt Brace College Publishers, Fort Worth
 47. Pérez-Cruz F, Navia-Vázquez A, Alarcón-Diana PL, Artés-Rodríguez A (2000a) An IRWLS procedure for SVR. In: Proc. EUSIPCO, Tampere, Finland
 48. Pérez-Cruz F, Alarcón-Diana PF, Navia-Vázquez A, Artés-Rodríguez A (2000) Fast training of support vector classifiers. In: Leen T, Dietterich T, Tresp V (eds) Neural information processing systems, vol 13. MIT Press, Cambridge, pp 734–740
 49. Platt J (1999) Fast training of support vector machines using sequential minimal optimization, advances in Kernel methods support vector learning. MIT Press, Cambridge
 50. Powers DA, Xie Y (2008) Statistical methods for categorical data analysis, 2nd edn. Emerald, Howard House
 51. Rampriya B, Mahadevan K, Kannan S (2010) Unit commitment in deregulated power system using Lagrangian firefly algorithm. *IEEE Int Conf Commun Control Comput Technol (ICCCCT2010)* 11745352 Ramanathapuram India. <https://doi.org/10.1109/ICCCCT.2010.5670583>
 52. Sayadi MK, Ramezani R, Ghaffari-Nasab N (2010) A discrete firefly meta-heuristic with local search for makespan minimization in permutation flow shop scheduling problems. *Int J Ind Eng Comput* 1(1):1–10. <https://doi.org/10.5267/j.ijiec.2010.01.001>
 53. Singh R, Rk U, Ahmad M, Ansari MK, Sharma LK, Singh TN (2017) Prediction of geomechanical parameters using soft computing and multiple regression approach. *Measurement* 99:108–119. <https://doi.org/10.1016/j.measurement.2016.12.023>
 54. Stehman SV (1997) Selecting and interpreting measures of thematic classification accuracy. *Remote Sens Environ* 62(1):77–89. [https://doi.org/10.1016/S0034-4257\(97\)00083-7](https://doi.org/10.1016/S0034-4257(97)00083-7)
 55. Suttrop T, Igel C (2006) Multi-objective optimization of support vector machines. In: Jin Y (eds) Multi-objective machine learning. Studies in computational intelligence, vol 16. Springer, Berlin, pp 199–220. https://doi.org/10.1007/3-540-33019-4_9
 56. Tuia D, Verrelst J, Alonso L, Pérez-Cruz F, Camps-Valls G (2011) Multi output support vector regression for remote sensing biophysical parameter estimation. *IEEE Geosci Remote Sens Lett* 8(4):804–808. <https://doi.org/10.1109/LGRS.2011.2109934>
 57. Umrao RK, Sharma LK, Singh R, Singh TN (2018) Determination of strength and modulus of elasticity of heterogeneous sedimentary rocks: an ANFIS predictive technique. *Measurement* 126:194–201. <https://doi.org/10.1016/j.measurement.2018.05.064>
 58. Vapnik VN (1995) The nature of statistical learning theory, 2nd edn. Springer-Verlag, New York
 59. Wang H, Hu D (2005) Comparison of SVM and LS-SVM for regression. In: Proc. Int. Conf. on neural networks and brain. IEEE, New York, pp 279–283. <https://doi.org/10.1109/ICNNB.2005.1614615>
 60. Willmott CJ (1984) On the evaluation of model performance in physical geography. *Spatial Stat Models*. https://doi.org/10.1007/978-94-017-3048-8_23
 61. Yang XS (2008) Firefly algorithm. Luniver Press, Bristol
 62. Yang XS (2012) Chaos-enhanced firefly algorithm with automatic parameter tuning. *Int J Swarm Intell Res* 2(4):125–136. <https://doi.org/10.4018/jsir.2011100101>

63. Yang XS (2013) Multi-objective firefly algorithm for continuous optimization. *Eng Comput* 29(2):175–184
64. Yang XS (2014) Nature-inspired optimization algorithms. Elsevier. <https://doi.org/10.1016/C2013-0-01368-0>
65. Yun Y, Yoon M, Nakayama H (2009) Multi-objective optimization based on meta-modeling by using support vector regression. *Optim Eng* 10:167–181. <https://doi.org/10.1007/s11081-008-9063-1>
66. Yun Y, Nakayama H, Arakawa M (2004) Using support vector machines in multi-objective optimization. In: 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541). <https://doi.org/10.1109/IJCNN.2004.1379903>