RESEARCH

Estimation of static Young's modulus of sandstone types: effective machine learning and statistical models

Na Liu[1] · Yan Sun[2] · Jiabao Wang[3] · Zhe Wang[1] · Ahmad Rastegarnia[4] · Jafar Qajar[5,6]

Abstract

The elastic modulus is one of the important parameters for analyzing the stability of engineering projects, especially dam sites. In the current study, the effect of physical properties, quartz, fragment, and feldspar percentages, and dynamic Young's modulus (DYM) on the static Young's modulus (SYM) of the various types of sandstones was assessed. These investigations were conducted through simple and multivariate regression, support vector regression, adaptive neuro-fuzzy inference system, and backpropagation multilayer perceptron. The XRD and thin section results showed that the studied samples were classified as arenite, litharenite, and feldspathic litharenite. The low resistance of the arenite type is mainly due to the presence of sulfate cement, clay minerals, high porosity, and carbonate fragments in this type. Examining the fracture patterns of these sandstones in different resistance ranges showed that at low values of resistance, the fracture pattern is mainly of simple shear type, which changes to multiple extension types with increasing compressive strength. Among the influencing factors, the percentage of quartz has the greatest effect on SYM. A comparison of the methods' performance based on CPM and error values in estimating SYM revealed that SVR ($R^2 = 0.98$, RMSE $= 0.11$GPa, CPM $= +1.84$) outperformed other methods in terms of accuracy. The average difference between predicted SYM using intelligent methods and measured SYM value was less than 0.05% which indicates the efficiency of the used methods in estimating SYM.

Abbreviations

DYM	Dynamic Young's modulus
ML	Machine learning
SYM	Static Young's modulus
MLR	Multivariate linear regression
XRD	X-Ray diffraction
ANNs	Artificial neural networks
CPM	Computed performance metric
RFR	Random forest regression
R^2	Determination coefficient
ANFIS	Adaptive neuro-fuzzy inference system
RMSE	Root mean squared error
BPMLP	Backpropagation multilayer perceptron
MAPE	Mean absolute percentage error
CART	Classification and regression tree
VAF	Variance accounted for
RBF	Radial basis function
UCS	Uniaxial compressive strength
ANOVA	Analysis of variance
D	Density
SR	Simple regression
Vp	Compressional wave velocity

Communicated by Hassan Babaie.

✉ Zhe Wang
hahawojiaowangzhe@126.com

✉ Jafar Qajar
j.qajar@uu.nl

1 Beijing GrandTrend International Economic and Technical Consulting Co., Ltd, Beijing 100012, China

2 Guangzhou Changdi Spatial Information Technology Co., Ltd, Guangzhou, Guangdong 510663, China

3 Geophysical Survey Center, China Geological Survey. Langfang, Hebei 065000, China

4 Department of Geology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran

5 Department of Earth Sciences, Utrecht University, Utrecht 3584 CB, the Netherlands

6 Department of Petroleum Engineering, School of Chemical and Petroleum Engineering, Shiraz University, Shiraz 7134851154, Iran

🖄 Springer

SVR	Support vector regression
Eqs.	Equations
SVM	Support vector machine
XLD	Xiaolangdi Dam
GMF	Gaussian membership function
TGD	Three Gorges dam
MF	Membership functions
Qzr	Quartz ratio
U	Water uptake
Qz or Q	Quartz
n	Porosity
Ft or FM	Fragment
SHN	Schmidt hardness number
Ch	Chert
ISRM	International society for rock mechanics
Fr or FP	Feldspar
Vs	Shear wave velocity
Fm-m	Metamorphic fragments
ASTM	American society for testing and materials
MFIS	Mamdani fuzzy inference system
FFNN	Feedforward neural networks

Introduction

The assessment of rock geo-mechanical properties plays a pivotal role in understanding the behavior of subsurface formations and optimizing various engineering applications (Abdelhedi et al. 2023; Liu et al. 2023; Ameen et al. 2009). The elastic modulus represents the rigidity and resistance of the rocks to failure. This parameter is one of the important parameters for analyzing the stability of engineering structures, especially dam sites. Recent advancements in the field have witnessed a diverse array of studies employing innovative approaches to characterize rock dynamic properties (Li and Dias 2023; Bouchaala et al. 2024; Motahari et al. 2022; Pappalardo and Mineo 2022). This introduction explores key findings from prominent research articles that delve into the intricacies of rock mechanics and elastic modulus estimation. The focus is on factors affecting dynamic properties and various methodologies, including traditional experimental approaches and innovative machine learning (ML) techniques.

Zhang et al. (2023) summarized factors affecting static and dynamic properties and proposed a method for extracting static Young's moduli (SYM) by analyzing elastic wave velocities of the various rocks. Rahman and Sarkar (2023a) explored correlations between uniaxial compressive strength and dynamic elastic properties, emphasizing diverse rock types. Shahani et al. (2022) applied machine learning (ML) models for intelligent predictions at Thar Coalfield, while Manda et al. (2023) used a machine learning approach to predict geo-mechanical properties from well logs. Additionally, Li and Dias (2023) assessed rock elasticity modulus using hybrid random forest regression (RFR) models, combining data-driven and soft techniques. Daraei and Zare (2019) predicted the static elastic modulus of limestone using downhole seismic tests, demonstrating the applicability of ML to diverse geological formations. Mahmoud et al. (2020) applied backpropagation multilayer perceptron (BPMLP), support vector regression (SVR), and Mamdani fuzzy inference system (MFIS) for SYM predictive models of sandstone formations. Khan et al. (2022) utilized multivariate statistics alongside ML models to forecast uniaxial compressive strength (UCS) and SYM (static Young's modulus). They stated that the random forest regression (RFR) has the best performance for predicting. Researchers have increasingly turned to artificial neural networks (ANNs), multivariate linear regression (MLR), and support vector regression (SVR) for their efficacy in estimating the dynamic and mechanical characteristics of rocks (Motahari et al. 2022; Mahmoud et al. 2020; Rastegarnia et al. 2021). Pappalardo and Mineo (2022) predicted static elastic modulus through regression models and BPMLP, showcasing the versatility of predictive modeling in rock mechanics. Rastegarnia et al. (2021) focused on the engineering characteristics and static properties of clay-bearing rocks, contributing valuable insights into this specific rock type. They stated that the BPMLP method has a conservative trend for forecasting SYM. Fang et al. (2023) assessed the applicability of BPMLP, SVR, ANFIS, and multiple linear regression (MLR) methods to estimate the SYM and UCS of the rock samples. Khosravi et al. (2022) contributed to the evaluation and prediction of both static and dynamic parameters in rocks, offering a holistic perspective on rock behavior. Additionally, Onaloa et al. (2018) provided a SYM model for formation evaluation, enriching the understanding of rock properties for practical applications in the petroleum industry. Guo et al. (2023) conducted a comprehensive assessment of rock geo-mechanical properties and wave velocities, providing valuable insights into the complex interplay between various rock parameters. They stated that the SVR method has a higher accuracy than ANFIS, BPMLP, and MLR models. Soustelle et al. (2023) contributed to the understanding of carbonates by investigating the relationship between static and dynamic elastic moduli, shedding light on the nuanced behavior of these formations. In the realm of carbonate rocks, Hadi and Nygaard (2023) focused on estimating UCS and SYM using petrophysical properties, offering practical applications for the petroleum industry. Abdi et al. (2023) employed the RFR for forecasting the SYM of weak rock samples, addressing challenges associated with rocks of lower strength. Ebrahimi et al. (2023) verified the use of daily drilling reports in conjunction with ML methods for the estimation

of Young's modulus, showcasing the potential for real-time applications in the oil, gas, and petrochemical industries. Kheirollahi et al. (2023) integrated traditional well data with computational techniques to enhance velocity estimations. Ghafoori et al. (2018) used statistical methods to forecast SYM in the Asmari formation, promising improved accuracy. Their results showed that the dynamic Young's modulus was 5 times larger than the static Young's modulus. Motahari et al. (2022) presented a new relationship between SYM and DYM of the feldspathic litharenite samples. They stated that by increasing the amount of quartz in the samples, the compressional and shear wave speeds increased. Also, BPMLP, SVR, and MLR methods are conservative in estimating Vs. Comparing the performance of the methods in estimating Vs showed that SVR has higher accuracy than other models. Xie et al. (2024)

investigated the factors affecting shear wave velocity (Vs) of the limestone samples using Gaussian process regression (GPR), BPMLP, and MLR models. They stated that the GPR showed higher accuracy compared to the BPMLP and MLR. Table 1 summarizes some of the previous relationships to estimate the SYM.

Considering the challenges in measuring the mechanical properties of rocks in civil and mining projects, this research prioritizes SYM estimation using simple and cost-effective methods. After assessing petrography, physical and mechanical properties of the sandstone samples types, simple regression (SR), MLR, and artificial intelligence methods including SVR, ANFIS, and backpropagation multilayer perceptron (BPMLP) were employed to build SYM prediction models. The models utilized mineralogy, physical properties, and dynamic Young's modulus

Table 1 Some of the previous relationships to estimate the SYM

Provided relationship	Units and variables	Authors	Rock type	Equations no. (Eqs.)
$SYM = 0.0015Vp - 2.516$	SYM (GPa) and Vp (m/s)	Kurtulus et al. (2012)	Travertine	(1)
$SYM = 23.643 - 1.786*U$	SYM (GPa) and U (%)	Fang et al. (2023) using U	Sandstones	(2)
$SYM = 23.52 - 1.1*n$	SYM (GPa) and n (%)	Fang et al. (2023)-using n	Sandstones	(3)
$SYM = 11.237Vp - 6.894$	SYM (GPa) and Vp (km/s)	Bejarbaneh et al. (2018)	Sandstones	(4)
$SYM = 2.06Vp^{2.78}$	SYM (GPa) and Vp (m/s)	Moradian and Behnia (2009)	Various stones	(5)
$SYM = 0.005Vp + 0.621$	SYM (GPa) and Vp (m/s)	Abdi et al. (2023)	Sandstones	(6)
$SYM = 0.091*DYM^{1.552}$	SYM (GPa) and DYM (GPa)	Davarpanah et al. (2020)	Different stones	(7)
$SYM = 0.867*DYM - 2.085$	SYM (GPa) and DYM (GPa)	Brotons et al., (2014)	Various stones	(8)
$SYM = 0.932*DYM - 3.421$	SYM (GPa) and DYM (GPa)	Brotons et al., (2016)	Various stones	(9)
$SYM = 0.564*DYM - 3.494$	SYM (GPa) and DYM (GPa)	Fei et al. (2016)	Sandstones	(10)
$SYM = 0.022*DYM^{1.774}$	SYM (GPa) and DYM (GPa)	Ghafoori et al. (2018)	Limestone rocks	(11)
$SYM = 0.581*DYM - 4.71$	SYM (GPa) and DYM (GPa)	Salehi et al. (2011)	Limestone rocks	(12)
$SYM = 0.352*DYM^{1.149}$	SYM (GPa) and DYM (GPa)	Najibi et al. (2011)	Limestone rocks	(13)
$SYM = 0.541*DYM + 12.582$	SYM (GPa) and DYM (GPa)	Ameen et al. (2009)	Carbonate rocks	(14)
$SYM = 0.008Vp - 5.619$	SYM (GPa) and Vp (m/s)	Azimian and Ajalloeian (2015)	Marl	(15)
$SYM = 10.67Vp - 18.71$	SYM (GPa) and Vp (km/s)	Yasar and Erdogan (2004)	Carbonate rocks	(16)
$SYM = 0.919Vp^{1.9122}$	SYM (GPa) and Vp (km/s)	Altindag (2012)	Sedimentary rocks	(17)
$SYM = 0.4029*DYM$	SYM (GPa) and DYM (GPa)	Pereira et al. (2021)	Basalt rocks	(18)
$SYM = 0.0811*DYM^{1.491}$	SYM (GPa) and DYM (GPa)	Daraei and Zare (2019)	Limestone	(19)
$SYM = 0.7134*DYM + (1.9584)$	SYM (GPa) and DYM (GPa)	Onaloa et al. (2018)	Various rocks	(20)
$SYM = 0.03*DYM^{1.76}$	SYM (GPa) and DYM (GPa)	Khosravi et al. (2022)	Limestone	(21)
$SYM = 2.08*DYM^{0.54}$	SYM (GPa) and DYM (GPa)	Guo et al. (2023)	Sedimentary rocks	(22)
$SYM = 0.69*DYM - 16.58$	SYM (GPa) and DYM (GPa)	Sharifi et al. (2021)	Sedimentary rocks	(23)
$SYM = 0.49*DYM - 7.651$	SYM (GPa) and DYM (GPa)	Sharifi et al. (2023)	Carbonate rocks	(24)
$SYM = 1.09*DYM - 17.99$	SYM (GPa) and DYM (GPa)	Kotsanis et al. (2021)	Prasinites	(25)
$SYM = 27.161n^{-0.37}$	SYM (GPa) and n (%)	Köken (2021)	Sandstones	(26)
$SYM = 35.87 - 1.741n - 0.29ft$	SYM (GPa), Ft (%), and n (%)	Köken (2021)	Sandstones	(27)
$SYM = 14.20 - 1.68n + 0.22Qz$	SYM (GPa), Qz (%), and n (%)	Köken (2021)	Sandstones	(28)
$SYM = -11.11 + 9.05Vp$	SYM (GPa) and Vp (km/s)	Köken (2021)	Sandstones	(29)

Abbreviations: *D* Density, *U* water uptake, *n* porosity, *DYM* Dynamic Young's modulus, *SYM* Static Young's modulus, *Vp* Compressional wave velocity

(DYM) from obtained sandstone samples at two large dam sites in China to enhance precision and overcome limitations in predicting SYM.

Case studies

This study investigates the engineering geological characteristics of the two large dam sites in China. Situated on the Yellow River in China, the Xiaolangdi Dam (XLD) was completed in 1995, boasting dimensions of 154 m in height and 1650 m in length. Nestled within Henan province, approximately 20 km northwest of Luoyang, the XLD serves a diverse array of functions, including electricity generation and flood control. The geological makeup of the dam site comprises sedimentary rocks, primarily sandstone, and mudstone, forming its foundational base. These rock formations significantly influence both the stability and design of the dam. Beneath the surface lie Permian and Triassic bedrock, overlaid by Quaternary deposits. The Permian formation features reddish and brown silty claystone and argillaceous siltstone, interspersed with layers of fine calcareous sandstone and medium-coarse siliceous sandstone. Similarly, the Triassic formation consists of red-brown limestone sandstone, alongside fine siliceous or calcareous sandstones and argillaceous siltstone, occurring in stratified layers. Understanding these geological characteristics is paramount for ensuring the durability and functionality of the Xiaolangdi Dam. Engineers and geologists meticulously consider these factors during the design, construction, and ongoing maintenance processes to safeguard its integrity and effectiveness for its intended purposes.

Constructed in 2006 on the Yangtze River, the Three Gorges dam (TGD) serves a multifaceted role, encompassing electricity generation, augmentation of the river's shipping capacity, and mitigation of downstream flooding risks. The dam site presents a rich geological tapestry, characterized by sandstone, shale, and limestone formations. These sedimentary rocks exhibit variances in strength, permeability, and stability. The geological profile of the TGD site encompasses a blend of sedimentary rock formations, along with considerations for seismic activity risks, challenges associated with slope stability, and strategies for sedimentation management. This study focused on the analysis of sandstone types sourced from the dam site. Figure 1 depicts the specific locations of the studied dam sites, with a predominant emphasis on sampling from the TGD. Further details regarding the sampling locations, gathered from both the dam axis and reservoir, are illustrated in Fig. 1 and Table 2. The sampling locations along the reservoir and TGD axis cover depths ranging from 0.50 to 38 m. These areas primarily comprise sandstone formations, including arenite, litharenite, and feldspathic litharenite types, predominantly

sourced from the Badong formation dating back to the Jurassic period (Table 2).

To ensure precise mechanical property measurements, samples exhibiting cracks and joints were intentionally excluded. This precaution was taken due to the anisotropic behavior of these features, which can lead to stress concentration and potential measurement errors (Yu et al. 2021). Consideration of anisotropy is particularly critical in the design of rock structures, especially when dealing with metamorphic rocks like schist, which are characterized by their sheet-like structure and mechanical variability (Fang et al. 2023).

Materials and methods

Laboratory tests

In the laboratory studies, we examined 122 rock samples with a length to diameter ratio ranging from 2.5 to 3 under natural conditions. To mitigate stress concentration concerns, specimens exhibiting joints and cracks were omitted. Prior to testing, the ends of the specimens were parallelized and polished. For petrography analysis, 122 thin sections were meticulously prepared and scrutinized in the microscopic laboratory using transmitted light. Thin section identifications of the samples were conducted based on Folk (1980) instructions. X-ray diffraction (XRD) analysis employed a D8-Advance, Bruker AXS equipment via copper Kα beam (λ = 1.5406Å) with a 2θ angle range of 2 to 60 degrees, using powder of samples with 50-micron size.

The determination of the quartz ratio (Qzr) was derived from XRD results using Eq. 30, incorporating the percentages of quartz (Qz), fragment (Ft), and feldspar (Fr).

$$Qzr = \frac{Qz\%}{(Qz\% + Ft\% + Fr\%)} \qquad (30)$$

Density (D), porosity (n), Schmidt hardness number (SHN), and water uptake by weight (U) were determined using methods specified in ISRM standards (ISRM 1981). The porosity test was determined based on the saturation and immersion method. This method is used for rock samples with regular, irregular geometric shapes or rock pieces. Also, the sample should not be crisp and fragile and should not be swollen and should not disintegrate or disintegrate due to exposure to water or a greenhouse. In this test, the sample is washed with water to remove the dust on the surface of the water. The sample is placed in water and in a vacuum of less than 800 pascals for at least one hour to be saturated. During this period, the sample should be moved alternately to remove the air bubbles. Then the sample is placed in the basket and enters the water tub. In this case, the

Fig. 1 Studied dams: Xiaol-angdi dam cross-sectional view (**a** part; abbreviations: SIS: sandstone interbedded with silt-stone, FS: fine sandstone, SSCM: siltstone interbedded with sandstone, conglomerate and silty mudstone, FSCS: fine sandstone interbedded with silt-stone and conglomerate, S: silt-stone, SC: silty claystone, SSC: fine sandstone interbedded with siltstone and silty claystone, al Q3: alluvium deposit of sand and gravel of the Pleistocene, F: fault), and Three Gorges dam site and reservoir geological map (**b** part)
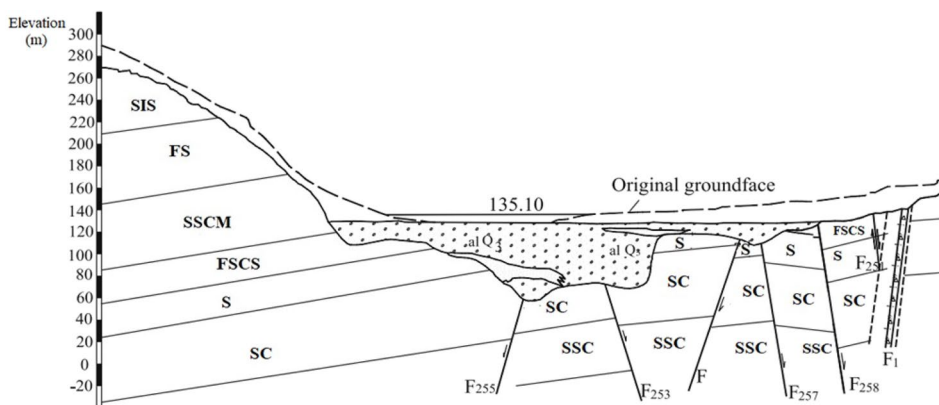
SIS: sandstone interbedded with siltstone, FS: fine sandstone, SSCM: siltstone interbedded with sandstone, conglomerate and silty mudstone, FSCS: fine sandstone interbedded with siltstone and conglomerate, S: siltstone, SC: silty claystone, SSC: fine sandstone interbedded with siltstone and silty claystone, al Q3: alluvium deposit of sand and gravel of the Pleistocene, F: fault

(a)
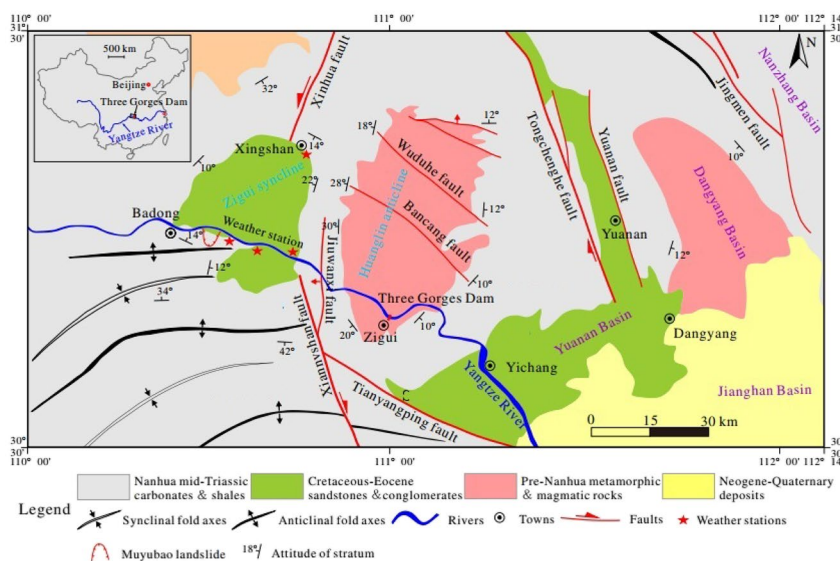
(b)

Table 2 Geological properties of the collected samples

Site	Geological Age	Sandstone types based on the Folk (1980) classification	Number of used samples	Depth (m)
XLD	Triassic and Permian	Feldspathic litharenite and litharenite	40	0.20–86
TGD	Jurassic	Arenite, feldspathic litharenite, and litharenite	82	0.50 to 38

basket is hung on the scale with a wire. The saturated mass of the basket and the sample is measured with an accuracy of 0.1 g. Finally, the result of dividing the volume of pores by the total volume of the sample, which is expressed as a percentage, is called porosity.

$$n(\%) = \frac{V_v}{V_t} * 100 \tag{31}$$

In this Eq. n is porosity, Vv is volume of pores in the sample, and Vt is total volume of rock sample.

In the present investigation, we employed the N-type Schmidt hammer, featuring a tensioned spring exerting force near the specimens. This device was consistently used in a vertical orientation for all samples in this study. The obtained hardness is affected by the hammer's orientation, using upright, horizontal, or upright positions within a ± 5 degrees deviation (ISRM 1981).

Compressional (Vp) and shear (Vs) wave tests followed the ASTM D2845 standard at a frequency of 0.50MHz (ASTM 2008). The device used to measure the Vp and Vs is the Pundit Plus device made by Proceq, Switzerland. This device has a transducer with pressure and shear wave speed, frequency 1 MHz, diameter and length 50 mm, maximum force tolerance 220 kilonewtons and temperature range 0 to 70 degrees. Prior to mechanical tests, sample ends were parallelized. Uniaxial compressive strength (UCS) tests were conducted per ASTM's recommended method with a constant loading rate of 0.70MPa/S (ASTM 2002). In the UCS test, when a static load was gradually applied to the sample, the axial strain was measured using a strain gauge. Then the stress–strain curve was drawn. Various definitions of Young's modulus are provided in Fig. 2. The secant static Young's modulus (SYM) was determined using the stress–strain curve. The DYM was determined using Eq. 32 (Pereira et al. 2021).

$$DYM = \frac{D*Vs^2(3*Vp^2 - 4*Vs^2)}{Vp^2 - Vs^2} \qquad (32)$$

where, Vp, Vs, and D are in km/s and g/cm^3, respectively.

The SVR approach

The SVR algorithm aims to identify the optimal regression function, minimizing prediction error while adhering to specified ε-tube bounds (Hussan et al. 2023; Fang et al. 2023). This method looks for a border that, in addition to separating the data of two groups, has the greatest distance to the closest data of the two groups (Kookalani and Cheng 2021; Maleki and Emami 2019). The closest data of each group to the decision boundary are called support vectors. In some problems, the model input data are not linearly separable for classification. In such cases, the support vector machine (SVM) uses a non-linear imager to transfer the data to a new and larger space (Fang et al. 2023; Dutta et al. 2024; Guo et al. 2023). Upon reviewing published articles on estimating using SVR, it was evident that among various kernel functions, the radial basis function (RBF) outperforms others in terms of efficiency (Fang et al. 2023;

Khosravi et al. 2022). Consequently, we adopted this specific kernel function to estimate SYM in our research.

The ANFIS

The ANFIS, a hybrid machine learning (ML) approach merging ANNs and fuzzy logic, facilitates data-driven inference and predictions (Khajevand 2022). In this study, ANFIS based on the Gaussian membership function (GMF) and Sugeno fuzzy system was employed to estimate SYM. ANFIS-GMF incorporates fuzzy logic, managing uncertain information through linguistic variables and membership functions (Khajevand 2023a). Utilizing feedforward neural networks (FFNN), often with a hidden layer, this method maps input variables to output values, harnessing the strengths of fuzzy logic and NNs (Tashayo et al. 2020). Fuzzy logic handles linguistic rules and input-to-fuzzy set mapping, while the neural network learns fuzzy set parameters from data (Guo et al. 2023; Hasheminezhad and Sadeghi 2023; Khajevand 2023b).
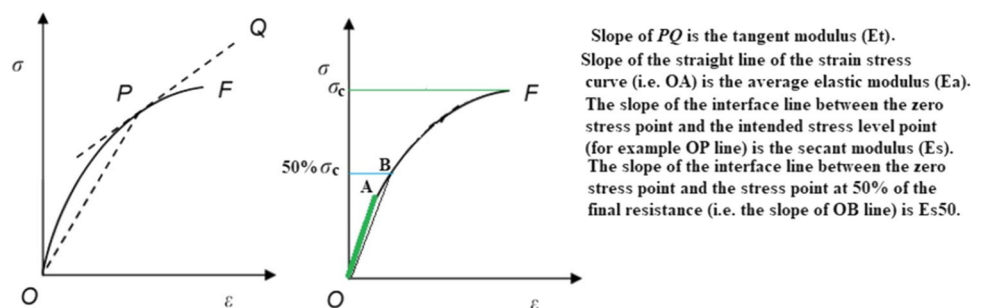
The BPMLP method

The BPMLP refers to a type of ANN that uses the backpropagation algorithm for training. In a multilayer perceptron, information moves through an input layer, one or more hidden layers, and an output layer (Shamsashtiany and Ameri 2018). The BPMLP method was completely described in previous studies (Guo et al. 2023; Khajevand 2023b, 2023a). In the current research, a feed-forward back-propagation algorithm with Levenberg Marquart (LM) training algorithm was used for the BPMLP training. We utilized collected equations from previous studies to predict the number of neurons using the BPMLP method (Fang et al. 2024). Accordingly, neurons 1 to 4 were investigated. MATLAB software was employed for BPMLP implementation to estimate the SYM.

Normalizing data and evaluation metrics

In this research, for intelligent modeling, normalization of data between -1 and 1 was done using Eq. 33.

Fig. 2 Various definitions of Young's modulus

Slope of *PQ* is the tangent modulus (Et).
Slope of the straight line of the strain stress curve (i.e. OA) is the average elastic modulus (Ea).
The slope of the interface line between the zero stress point and the intended stress level point (for example OP line) is the secant modulus (Es).
The slope of the interface line between the zero stress point and the stress point at 50% of the final resistance (i.e. the slope of the OB line) is Es50.

$$A_n = 2\left(\frac{A_a - A_{min.}}{A_{max.} - A_{min.}}\right) - 1 \qquad (33)$$

Here, A_n, A_a, A_{min}, and A_{max} represent the normalized, actual, minimum, and maximum values for the A parameter, respectively. To evaluate the effectiveness of the methods, calculated performance metric (CPM), mean absolute percentage error (MAPE), variance accounted for (VAF), root mean squared error (RMSE), and determination coefficient were employed (Eqs. 34–37). These metrics were extensively used for evaluating models (Shakir 2023; Xiao et al. 2023; Song et al. 2022).

$$CPM = R.^2 - RMSE + VAF \qquad (34)$$

$$\text{RMS}\left(\sqrt{\frac{1}{N_{SYM}}\sum_{i=1}^{N_{SYM}}\left(A_{SYM} - P_{SYM}\right)^2}\right)E = \qquad (35)$$

$$VAF = \left[1 - \frac{var(A_{SYM} - P_{SYM})}{var(A_{SYM})}\right] \qquad (36)$$

$$MAPE = \frac{\sum_{i=1}^{N_{SYM}}\left|\left(\frac{A_{SYM} - P_{SYM}}{A_{SYM}}\right)\right|}{N_{SYM}} \qquad (37)$$

where, N_{SYM}, A_{SYM}, and P_{SYM} represent the total experiments of SYM, actual and predicted SYM values, respectively.

Results and discussion

In this section, we present and compare laboratory test results of intact rock at the investigation sites, followed by the introduction of various experimental relationships and models for estimating the SYM.

One of the indirect methods of estimating rock properties at the construction site of construction projects is the use of indirect methods such as empirical relationships and intelligent models (Teshnizi et al. 2021). For this reason, in recent years, the presentation of models and relationships to estimate mechanical and dynamic properties, and destructive tests have been of interest to researchers. These methods are more important in the stages of identifying constructions and especially when sampling weak rocks (Pappalardo et al. 2022).

Analysis of laboratory results

The petrographic and geological features of the samples under examination are displayed in Table 2. Figures 3 and 4

present examples of laboratory activities, including an XRD pattern of a Feldspathic litharenite sample and thin sections from the studied samples.

Table 3 presents the geo-mechanical properties of 122 samples. Through XRD analysis (evidenced by a sample in Fig. 4) and thin section examination (some images in Fig. 3), the 122 sandstone samples were classified into arenite, litharenite, and feldspathic litharenite per Folk's (1980) categorization, encompassing components like feldspar, murky minerals, quartz, albite, carbonate and metamorphic fragments, and chert. Gypsum and calcite, as types of cement, were characterized by semi-rounded to angular grains and poor to medium sorting.

The examined samples (refer to Table 3) exhibit a broad spectrum of properties. SYM values range from 8.62 GPa to 35.02 GPa, while DYM values span from 29.81 GPa to 85.26 GPa. Porosity levels vary between 0.001% and 13.08%, and Schmidt hardness number (SHN) varies from 29 to 55. Density values fall in the ranges of 2.35 g/cm^3 to 2.77 g/cm^3. Based on the results, arenite and litharenite samples, incorporating Gypsum and clay elements, exhibit the lowest P-wave velocity, density, SHN, DYM, UCS, and SYM values. Conversely, specimens with high Qtr demonstrate the highest dynamic and static values.

The mineralogy of sandstone significantly influences its mechanical properties. For instance, quartz-rich sandstones tend to be harder and more durable, while clay-rich varieties may exhibit lower strength. The arrangement and interlocking of minerals impact factors like compressive strength, porosity, and permeability in sandstone rocks, thereby affecting their overall mechanical behavior (Diaz-Acosta et al. 2023; Liang et al. 2024). Sandstones with a higher quartz content generally have higher hardness and abrasion resistance (Motahari et al. 2022; Rastegarnia et al. 2022). Quartz is a hard and durable mineral, contributing to the overall strength of the rock. The minerals binding the sand grains together (cement) influence the rock's strength. Silica, calcite, and iron oxides are common cementing materials (Etemadi et al. 2020). Silica cements often result in harder and more competent sandstones. The presence of clay minerals can affect the cohesion and porosity of sandstone (Tofighkhah et al. 2023). Higher clay content may reduce the overall strength of the rock and increase its susceptibility to weathering (Bagherzadeh Khalkhali et al. 2019).

The arrangement of minerals influences the porosity of sandstone. Well-sorted, tightly packed grains generally lead to lower porosity and higher strength, while poorly sorted or loosely packed grains can result in higher porosity and lower strength. The size and sorting of minerals within the rock impact its permeability and strength (He et al. 2021; Li et al. 2023). Well-sorted and fine-grained sandstones often exhibit higher shear strength. The mechanical properties of sandstone are intricately linked to its mineral composition,

Fig. 3 An example of litharenite thin section from the TGD (**a**), feldspathic litharenite thin section from the XLD (**b**), litharenite thin section from the TGD (**c**), and litharenite thin section from the TGD (**d**): (Q, FM, FP, Ch, and Fm-m are quartz, rock fragments, feldspar, chert, and metamorphic fragments, respectively)
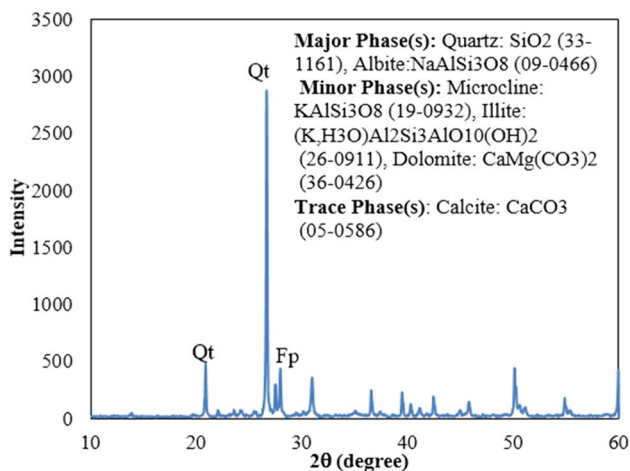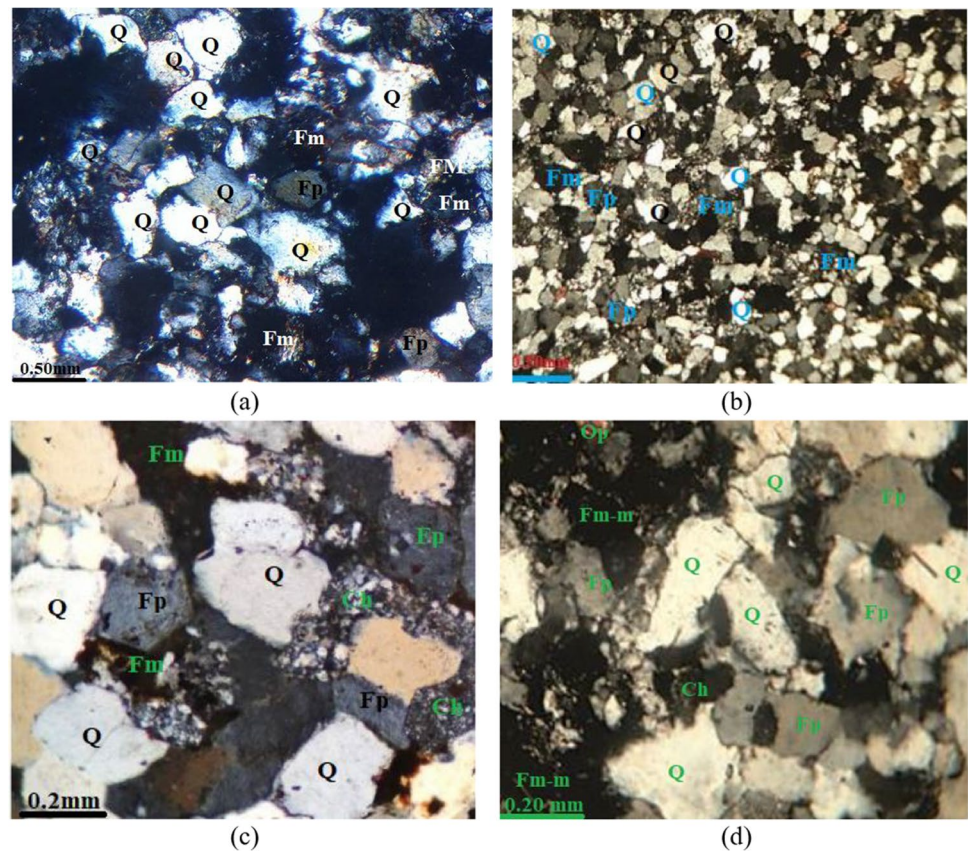

Fig. 4 An example of a feldspathic litharenite XRD pattern

with minerals like gypsum and clay playing pivotal roles in shaping the rock's behavior. The inclusion of gypsum may render the rock less robust, and its solubility in water can contribute to increased porosity over time, affecting both strength and durability. In contrast, the presence of clay minerals introduces a different set of influences on the mechanical properties of sandstone. Clay minerals, such as kaolinite, illite, or smectite, bring cohesion to the rock, impacting its

strength (Kafash Bazari 2023). The plasticity inherent in clay minerals also plays a role in shaping the deformation behavior of the sandstone. Moreover, these minerals contribute to porosity and permeability, affecting fluid flow and storage properties within the rock (Sharifi et al. 2023; Motahari et al. 2022; Khajevand 2023c).

The interaction between gypsum, clay minerals, and the cementing materials in sandstone is a dynamic process that further modulates the mechanical characteristics of the rock. Gypsum's susceptibility to dissolution in water and the potential for certain clay minerals to expand or contract with changes in moisture can lead to long-term changes in the sandstone's strength and stability (Ghavami and Rajabi 2021; Chen et al. 2023; Taheri and Ziad 2021; Rahman and Sarkar 2023b). Figure 5 presents box plots and normal curves illustrating the variables. The data exhibit normal distributions, supporting statistical analysis. Moreover, with a dataset exceeding 25 data points (122 samples), statistical analysis and Pearson correlation can be applied across the entire dataset, assuming a normal distribution.

Effects of failure modes and sandstone types on mechanical properties

Efforts have been undertaken to categorize various failure patterns observed in cylindrical samples. Figure 6 illustrates

Table 3 Statistical statistics of the measured properties

	Vp (km/s)	D (g/cm3)	n (%)	UCS (MPa)	SYM (GPa)	SHN	Qz (%)	Ft (%)	Fr (%)	Qtr	DYM (GPa)
Mean	4.71	2.61	4.47	87.02	22.47	43.00	12.13	45.09	40.23	0.12	55.86
Standard error	0.07	0.01	0.29	2.52	0.63	0.59	0.20	0.53	0.30	0.00	1.60
Standard deviation	0.78	0.13	3.20	27.63	6.87	6.47	2.17	5.78	3.28	0.02	17.57
Sample variance	0.60	0.02	10.26	763.49	47.22	41.91	4.73	33.39	10.75	0.00	308.77
Kurtosis	(1.27)	(0.98)	(0.64)	-1.03	(1.09)	(0.86)	(0.99)	(0.83)	(0.95)	(0.92)	(1.35)
Skewness	0.22	(0.63)	0.39	0.06	0.04	0.02	0.02	(0.13)	0.19	0.04	0.11
Minimum	3.60	2.35	0.001	37	8.62	29.00	7.73	32.34	34.17	0.08	29.81
Maximum	6.13	2.77	13.08	143	35.02	55.00	16.62	57.00	47.14	0.17	85.26

Fig. 5 Box plots and normal curves of the variables
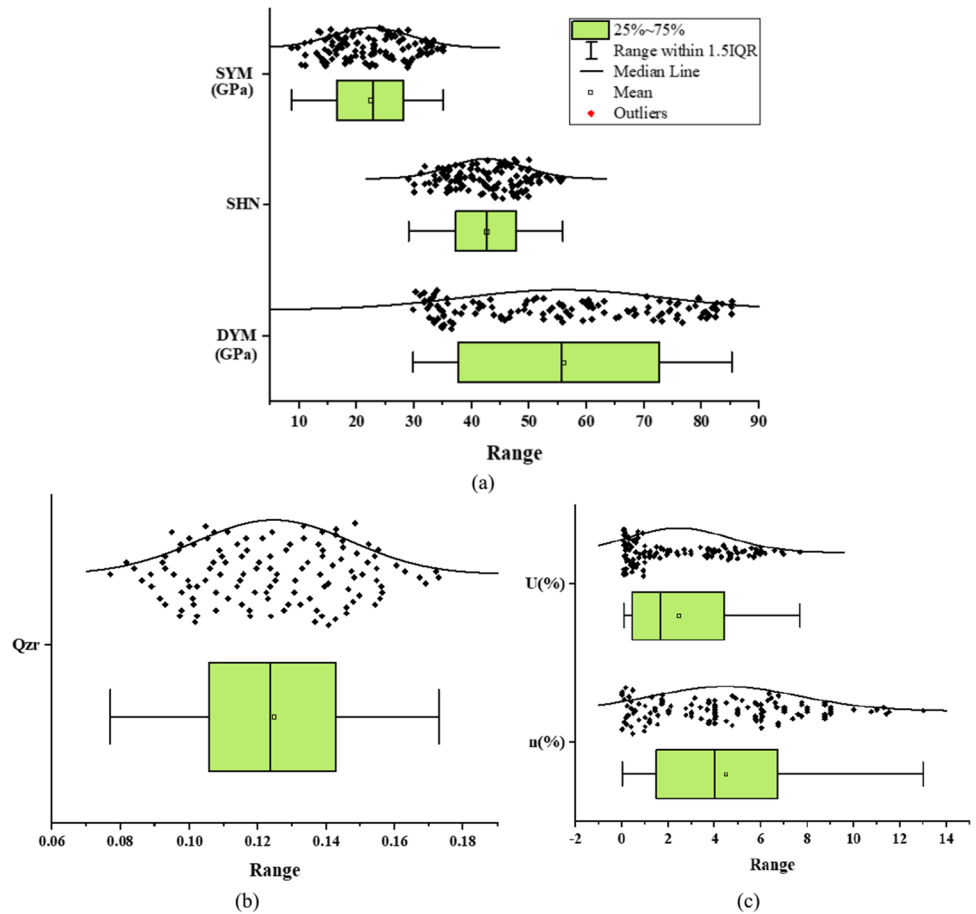
the different types of failure modes observed during UCS tests (Basu et al. 2013). Analyzing these failure patterns offers insights into the predominant stress states within the rock samples under examination. Tensile fractures typically form perpendicular to the direction of the minimum principal stress, while aligning with the direction of the average and maximum stresses (Basu et al. 2013; Bouchaala et al. 2023). However, the presence of microscopic discontinuities introduces a scale effect, wherein the strength decreases as sample dimensions increase. In larger samples, discontinuities have a more pronounced impact on strength, leading to predominantly shear-type fracture patterns (Basu et al. 2013).

In feldspathic litharenite samples, the prevalent fracture pattern observed in most cases is axial fracturing (AF) type, as depicted in Fig. 7. Conversely, the Y-shaped fracture (YSF) pattern is the least common among the samples examined. Similarly, the majority of fracturing patterns observed in both arenites and litharenites are of the AF type. The wide variability in compressive strength from axial torque contributes to the diverse failure patterns observed under axial pressure (Basu et al. 2013). Consequently, analyzing these
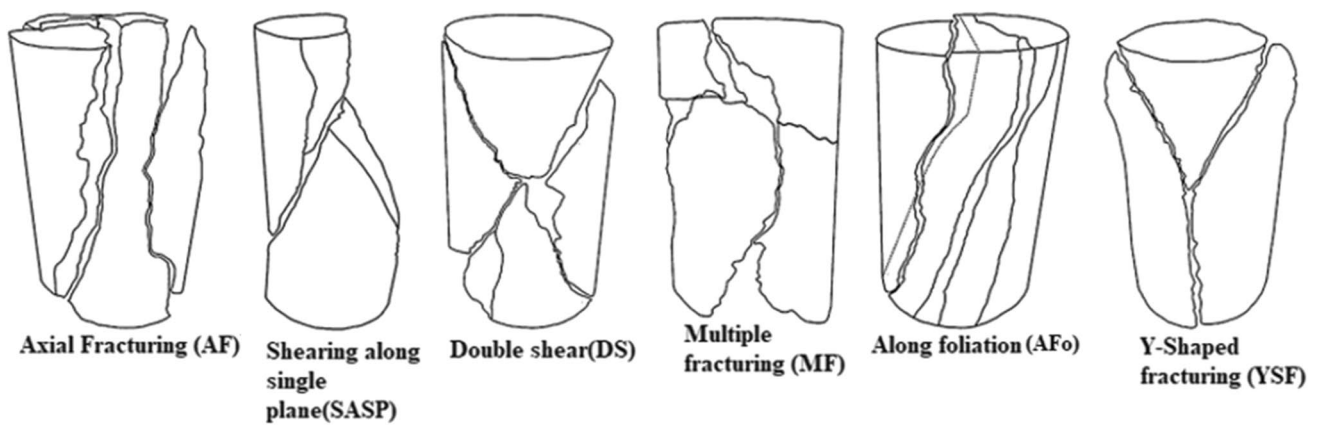
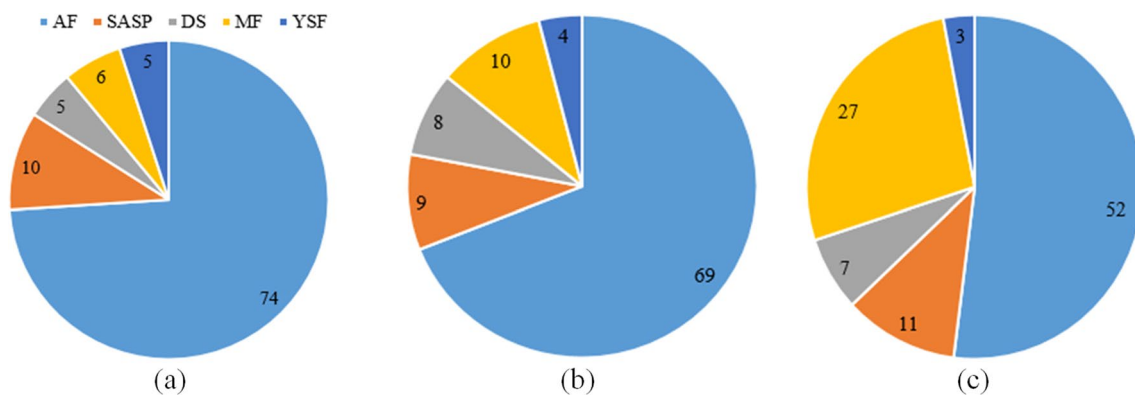**Fig. 6** Fracture pattern under UCS test (Basu et al. 2013)

Fig. 7 Fracture pattern observed for each sandstone types (values are in percent): Arenite (**a**), Litharenite (**b**), and Feldspathic litharenite (**c**)

failure patterns can elucidate the underlying reasons for such broad variations. As the resistance of the samples increases, transitioning from arenites to feldspathic litharenites, there is an escalation in multiple fracturing (MF) behaviors while the occurrence of axial fracture (AF) decreases. In essence, samples with lower strength exhibit a heightened tendency towards AF as the predominant failure mode.

The comparison of the characteristics of arenite and feldspathic litharenite and litharenite samples shows that the highest values of the average dynamic properties are related to feldspathic litharenite samples. Also, the highest value of density and compressive strength is related to litharenite feldspathic. The high specific weight of the feldspathic litharenite type is related to the high amounts of opaque heavy minerals in this type. Feldspathic litharenite type sandstones showed the highest average values of static and dynamic Young's modulus. The low resistance of arenite and litharenite types is mainly due to the presence of sulfate cement, clay minerals, high porosity, and carbonate fragments in this type. For example, the results related to the physical and mechanical parameters of samples 1 to 27 of each sandstone type are presented in Fig. 8.

The quartz content in sandstone is a key determinant of the rock's durability and its resistance to weathering processes (Wang et al. 2024). Moreover, the type and amount of cementing minerals, such as silica, calcite, or iron oxides, in sandstone play a crucial role in determining the rock's strength. A comprehensive understanding of the mineral composition of sandstone is vital for assessing their engineering geological characteristics. These insights into the role of calcite, clay minerals, and other components contribute to informed decision-making in construction and various engineering applications where these rocks are utilized (Rastegarnia et al. 2020).

Correlation matrix between inputs and SYM

To explore the influence of geo-mechanical parameters on the static Young's modulus (SYM), we analyzed the correlation matrix and the two-variable regression models (as shown in Tables 4, and Figs. 9 and 10). The results illustrated the potential for estimating SYM based on various factors including DYM, water uptake (U), porosity(n), quartz ratio (Qzr), and SHN as indicated in Fig. 9, due to

Fig. 8 Comparing physical and mechanical properties of the litharenite (L) and feldspathic litharenite (FL) for samples 1–27
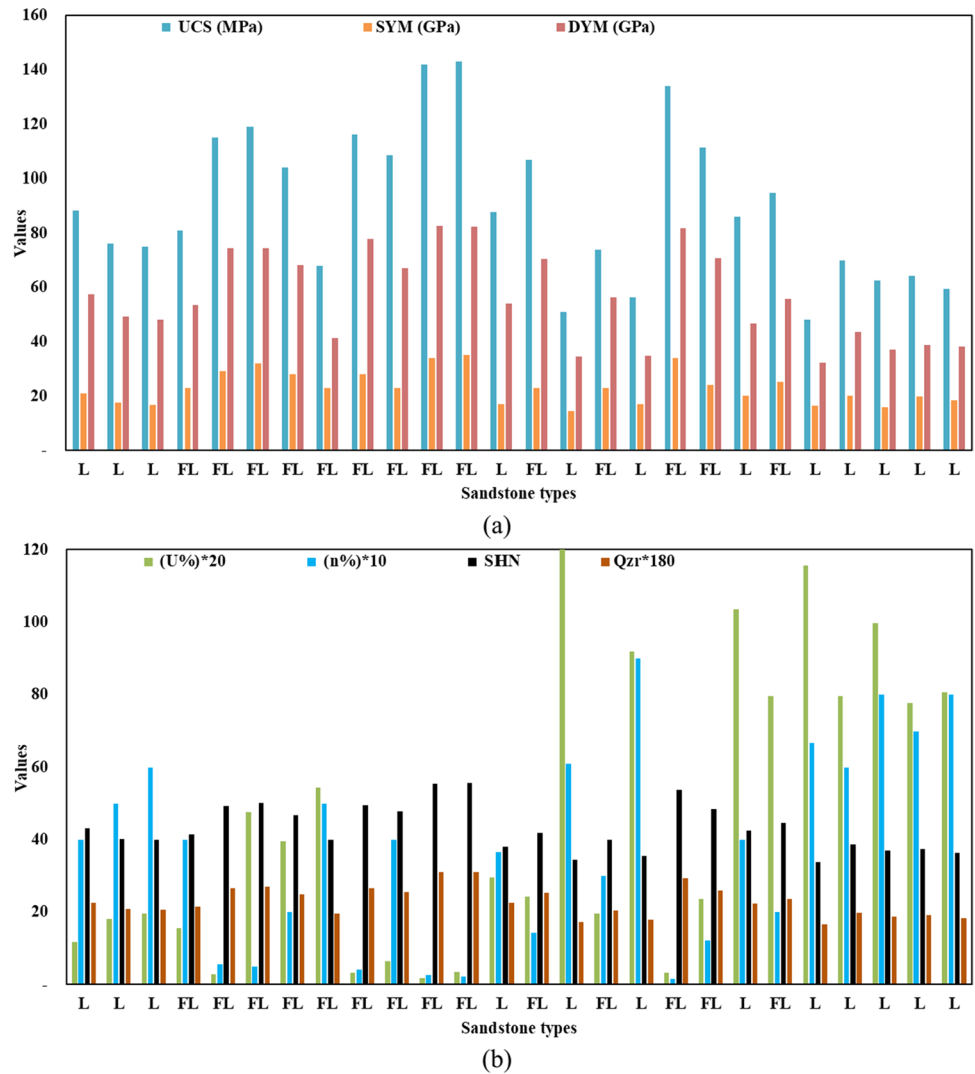

(a)

(b)

Table 4 Regression equations between static Young's modulus and inputs

Equations	%R-sq	Analysis of Variance		Eq. No
		F-Value	P-Value	
SYM = 1.821 + 0.3696DYM	89.31	985.94	0.00	(38)
SYM = 31.158 − 1.9435n	82.10	541	0.00	(39)
SYM = −21.21 + 1.0272SHN	93.63	1435.21	0.00	(40)
SYM = − 12.83 + 283.34Qzr	87.57	831	0.00	(41)
SYM = 28.995 − 2.680U	73.32	324	0.00	(42)

their correlation coefficients exceeding 0.84. Notably, water uptake exhibited the least impact on SYM. Following this, both simple and multivariate regression and machine learning (ML) techniques were employed to scrutinize parameter effects and determine the most suitable equation for SYM estimation.

Figure 10 shows a high correlation between the dynamic and static Young's modulus so that it is possible to estimate the static modulus based on the dynamic modulus. Based on the results of the current study, SYM is 2.5 times greater than DYM. The static and dynamic elastic modulus, which describe a material's stiffness, are influenced by various factors (Ghafoori et al. 2018; Li and Fan 2024; Zhang et al. 2023):

- Material composition: different rock types have different stiffness due to their atomic and molecular arrangements.
- Crystal structure: crystalline samples' elastic modulus depends on their crystal lattice arrangement.
- Temperature: elastic modulus decreases as temperature rises due to increased thermal vibrations.
- Strain rate: higher strain rates often lead to stiffer rock responses.
- Microstructure: features like grain size affect modulus; finer grains typically result in higher stiffness.

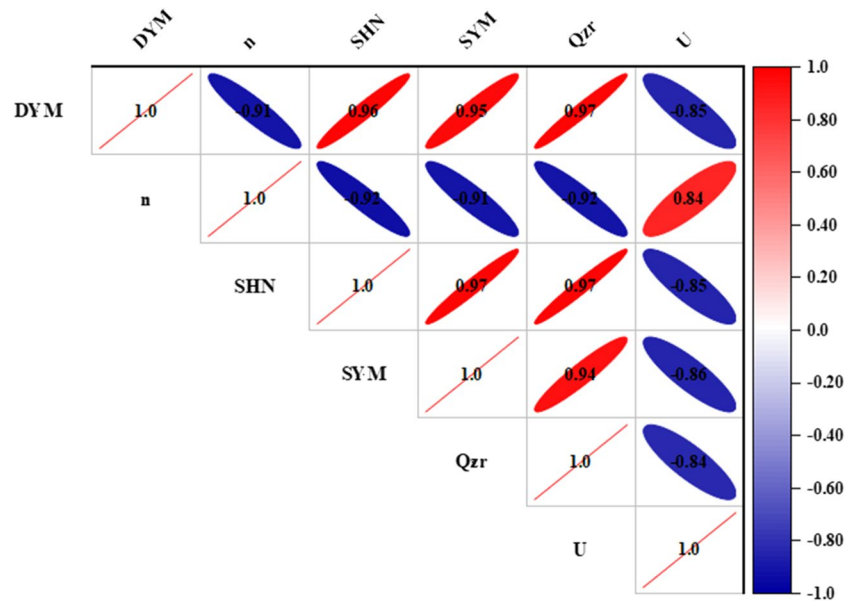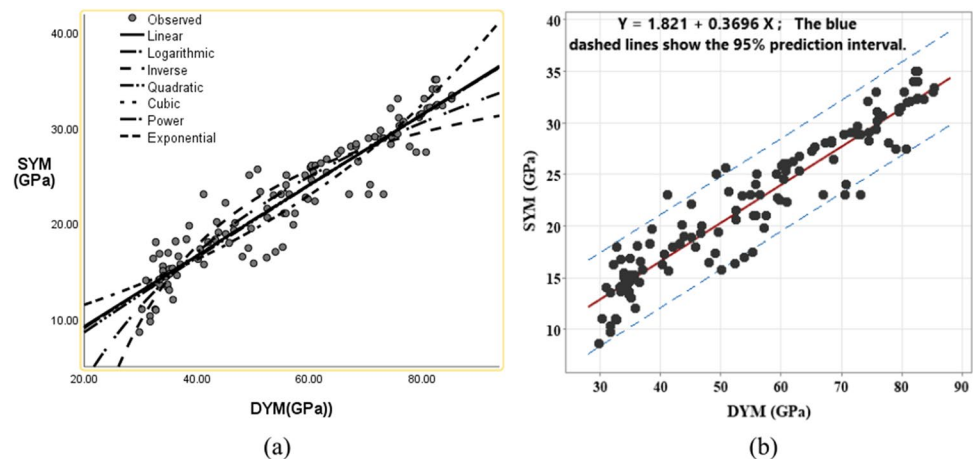**Fig. 9** Correlation matrix of the variables

Fig. 10 Schematic view of the relationships between dynamic and static Young's modulus: All fits (**a**) and the best fit (**b**)


(a)

(b)

- Presence of defects: cavities and defects like vacancies and dislocations can alter modulus by hindering or facilitating atomic movement.
- Moisture content: moisture absorption can change internal structure and stiffness.
- Applied stress/strain: material moduli can vary with the type and direction of applied stress or strain.
- Loading history: past mechanical loading affects modulus, with cyclic loading or creep causing microstructural changes.
- Frequency: In dynamic testing, frequency affects moduli due to viscoelastic effects.

Understanding these factors is essential for predicting and controlling rock behavior in various applications, using experimental and computational methods for characterization.

Evaluating previous empirical Eqs.

Table 1 outlines various relationships proposed by prior researchers for predicting SYM. Utilizing these relationships, SYM values were calculated for each sample, and the laboratory test outcomes were then compared with the estimated values derived from these empirical relationships. Evaluation results for the previous relationships are depicted in Figs. 11 and 12. Multiple metrics were employed to evaluate the predictive accuracy of the earlier relationships. In terms of the correlation coefficient, most relationships demonstrated satisfactory accuracy in SYM estimation. Notably, according to the CPM, among the examined background equations, the one proposed by Pereira et al (2021) based on Vp showcased superior precision in predicting SYM compared to other equations from the research background (Fig. 12). Despite having a high correlation coefficient, the

Fig. 11 Correlation between measured and predicted SYM
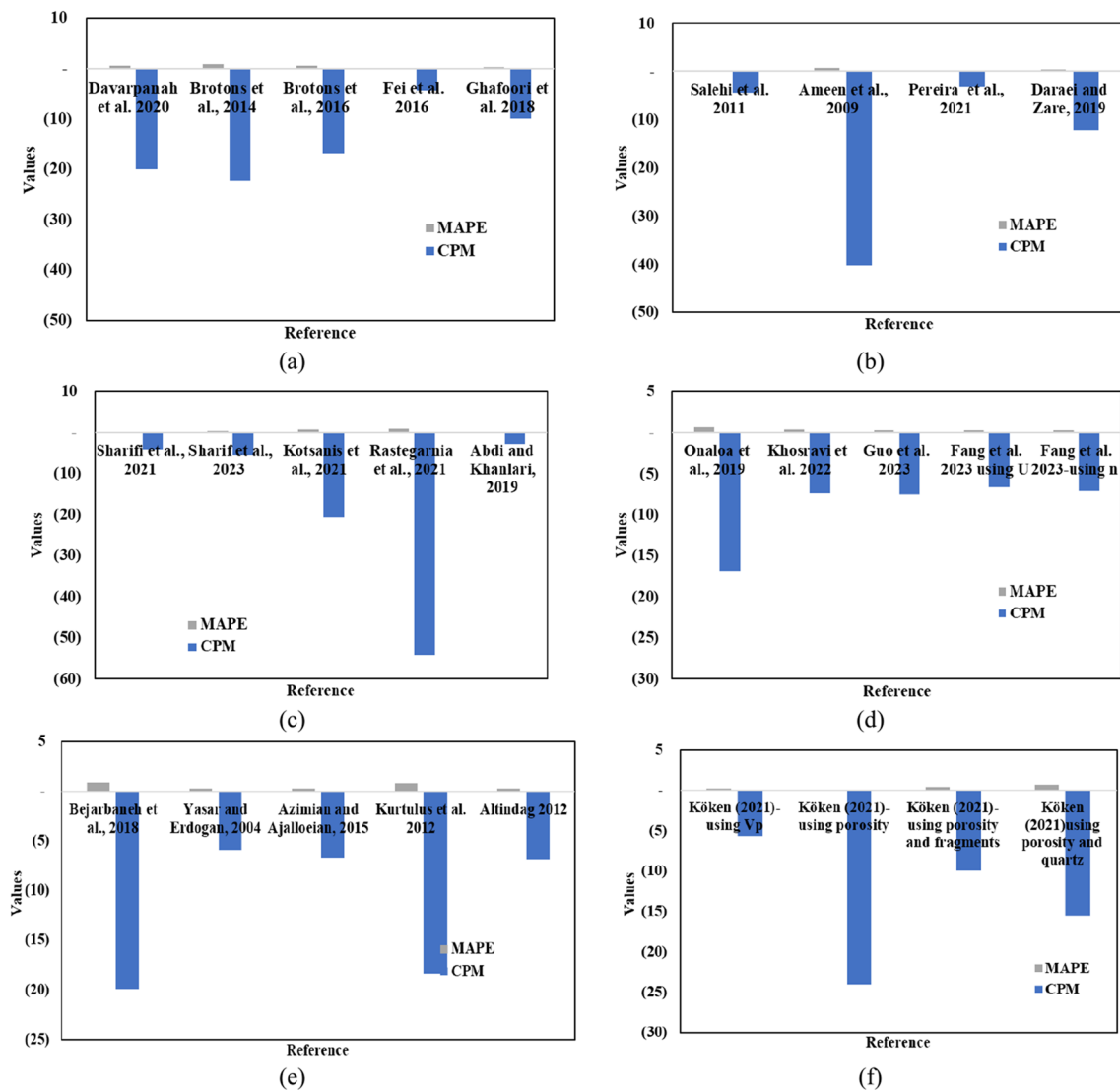

(a)

(b)

(c)

(d)

(e)

(f)

(g)
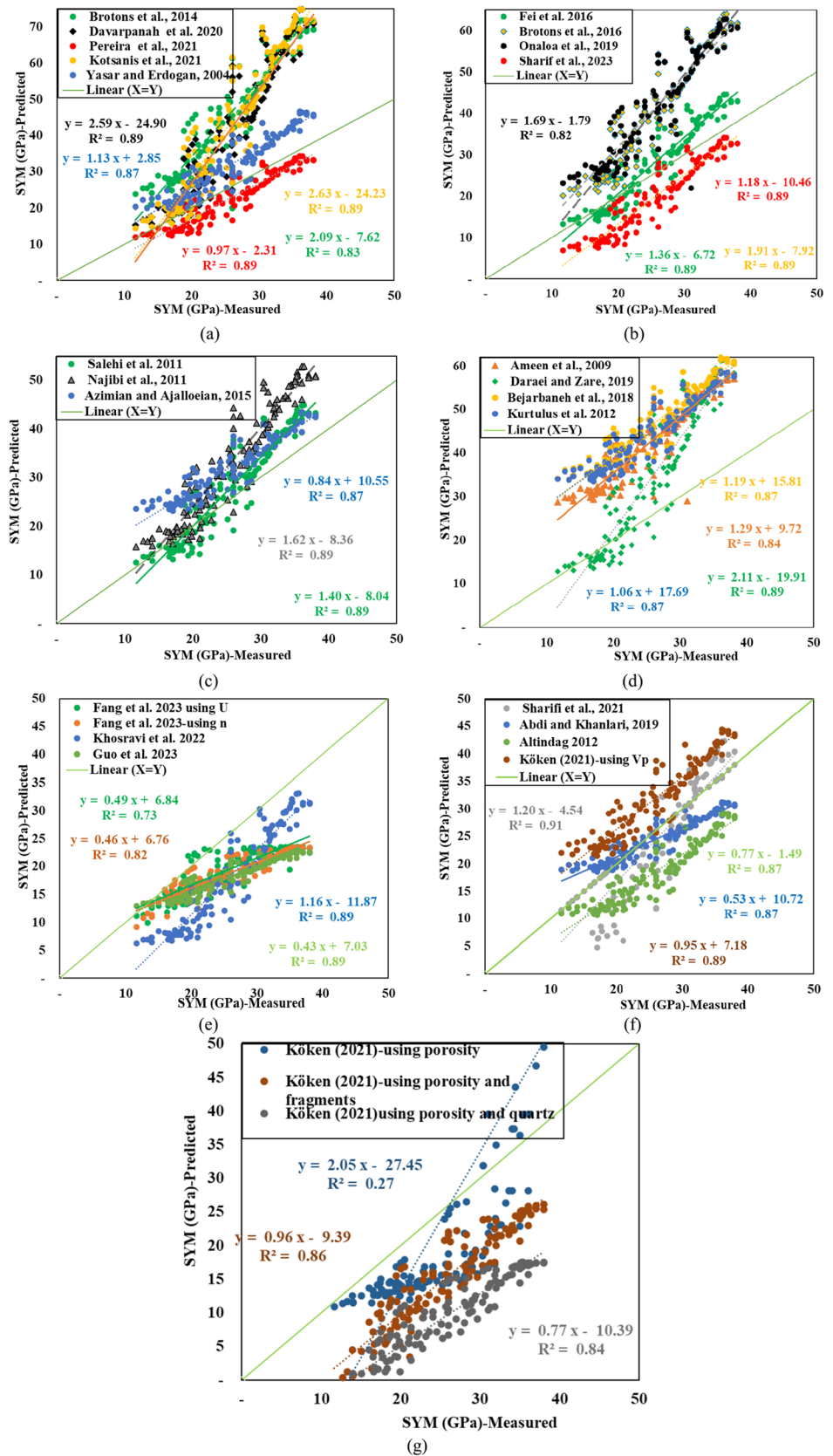
Fig. 12 Performance of previous proposed Eqs. to estimate SYM

CPM values of all previous relationships are negative, indicating low accuracy for SYM estimation (Fig. 12). Additionally, the values derived from most previous relations deviated significantly from the diametric line, resulting in their mean values differing substantially from the mean results of this study. Consequently, the development of local models for each area is deemed necessary (Li and Dias 2023). Thus, we explored the suitability of various modeling techniques for SYM estimation.

MLR Analysis

The MLR analysis was conducted utilizing Minitab software (Version 20) to construct predictive equations for forecasting SYM, as shown in Table 5. Model accuracy was assessed through analysis of variance (ANOVA), with

findings detailed in Tables 5 and 6, demonstrating significance below the 5% threshold and confirming the equation's efficacy in estimating SYM. The t-test was also employed to evaluate constant coefficients, with specific instances presented in Tables 5 and 6. The results show that, statistically, the variables of water uptake (U) and porosity (n) do not have much effect on the static modulus. Because the level of confidence in the t-test is more than 5%. These two variables were removed to check their effect on SYM. Removing these two variables has a very slight effect on the relationship determination coefficient (Table 5). Therefore, SYM estimation with three inputs of Schmidt number (SHN), dynamic modulus (DYM), and quartz ratio (Qzr) is more economical and preferred. In this way, these three variables were also considered as inputs in ML methods to estimate SYM. An example of coefficients and results of analysis of variance

Table 5 Results of MLR to estimate SYM

Input number	Inputs	Equation	R-square %	RMSE	Eq. No
5	DYM, n, SHN, Qzr, U	$SYM = -11.06 + 0.1162\,DYM - 0.083\,n + 0.924\,SHN - 90.7\,Qzr - 0.246\,U$	94.51	1.03	(43)
4	DYM, SHN, Qzr, U	$SYM = -12.56 + 0.1187\,DYM + 0.942\,SHN - 88.3\,Qzr - 0.266\,U$	94.49	1.05	(44)
4	DYM, n, SHN, Qzr	$SYM = -12.83 + 0.1291\,DYM - 0.135\,n + 0.956\,SHN - 96.0\,Qzr$	94.36	1.02	(45)
3	DYM, SHN, Qzr	$SYM = -15.65 + 0.1352\,DYM + 0.990\,SHN - 92.6\,Qzr$	94.31	1.01	(46)
3	n, U, DYM	$SYM = 11.55 - 0.441\,n - 0.435\,U + 0.2498\,DYM$	90.89	1.04	(47)
3	n, DYM, SHN	$SYM = -14.81 - 0.101\,n + 0.0725\,DYM + 0.7919\,SHN$	93.98	1.05	(48)
3	n, Qzr, SHN	$SYM = -17.04 - 0.203\,n - 33.8\,Qzr + 1.050\,SHN$	93.79	1.06	(49)
2	DYM, Qzr	$SYM = -3.52 + 0.2470\,DYM + 97.9\,Qzr$	89.93	1.06	(50)
2	SHN, Qzr	$SYM = -21.70 + 1.109\,SHN - 24.0\,Qzr$	93.67	1.04	(51)
2	SHN, DYM	$SYM = -16.89 + 0.8221\,SHN + 0.0786\,DYM$	93.95	1.04	(52)
2	n, U	$SYM = 31.016 - 1.362\,n - 1.008\,U$	85.14	1.06	(53)
2	n, Qzr	$SYM = 0.47 - 0.646\,n + 199.7\,Qzr$	89.12	1.07	(54)
2	n, DYM	$SYM = 9.44 - 0.553\,n + 0.2774\,DYM$	90.41	1.03	(55)
2	U, DYM	$SYM = 6.69 - 0.581\,U + 0.3077\,DYM$	90.25	1.05	(56)

Table 6 Features of the ideal MLR model (i.e., Eq. 46 presented in Table 5)

(a) Coefficients and t-test					(b) Analysis of variance					
Term	Coef	SE Coef	T-Value	*P*-Value	Source	DF	Adj SS	Adj MS	F-Value	*P*-Value
Constant	-15.65	2.04	-7.66	0.00	Regression	3	5299.17	1766.39	640.36	0.00
DYM	0.1352	0.0375	3.60	0.00	DYM	1	35.84	35.84	12.99	0.00
SHN	0.990	0.105	9.44	0.00	SHN	1	245.90	245.90	89.14	0.00
Qzr	-92.6	34.2	-2.71	0.008	Qzr	1	20.26	20.26	7.35	0.01

and t-test is presented in Table 6. These tests were extensively used in the literature to validate models and empirical relationships (Zhao et al. 2024; AlHamad et al. 2021; Shirnezhad et al. 2021).

The ANFIS model results

Utilizing MATLAB software, we developed an ANFIS-GMF model incorporating a Sugeno system. The ANFIS-GMF analysis utilized Gaussian membership functions distributed across rules, with linear membership functions employed for the SYM output layer. Through trial-and-error testing, appropriate membership degrees for input combinations were determined. To train the fuzzy inference system (FIS), we employed a hybrid learning algorithm that combines backpropagation with the least squares method. The developed ANFIS-GMF summarizes that the FIS type is GENFIS2 with input membership function type Gaussian and output membership function type linear. There are 4 membership functions (MFs) for the input, distributed across 4 rules of the Sugeno type (Fig. 13). The error goal is set at 0.00, with training continuing for 180 epochs. The influence radius is specified as 0.53. Figure 14 shows the developed ANFIS performance in SYM prediction.

The BPMLP results

In this study, a BPMLP was utilized with data normalized to the range of [-1,1] like other ML models. The input layer consisted of 4 neurons, corresponding to the independent variables, while the output layer featured a single SYM output. To address the complexity of the problem, we opted for a single hidden layer with minimal neurons, as suggested by previous research (Shahani et al. 2022; Alsalami et al. 2023). The dataset was split into training (93 cases) and testing (31 cases), with training determining weights and testing evaluating model performance. Relationships based on the number of dependent and independent variables have been presented by previous researchers to determine the number of neurons to be examined. Following equations proposed by various researchers (Fang et al. 2024), we examined neurons 1 to 4 to identify the best neuron for predicting SYM using BPMLP (Table 7). We applied a Sigmoid transfer function between the input and hidden layers and a linear transfer function between the hidden and output layers. The BPMLP training utilized the LM algorithm. Results revealed the highest correctness in approximating SYM for the third neuron with 3 inputs and LM training algorithm, as detailed in Table 7. The LM algorithm's adaptive learning rate adjustment capability
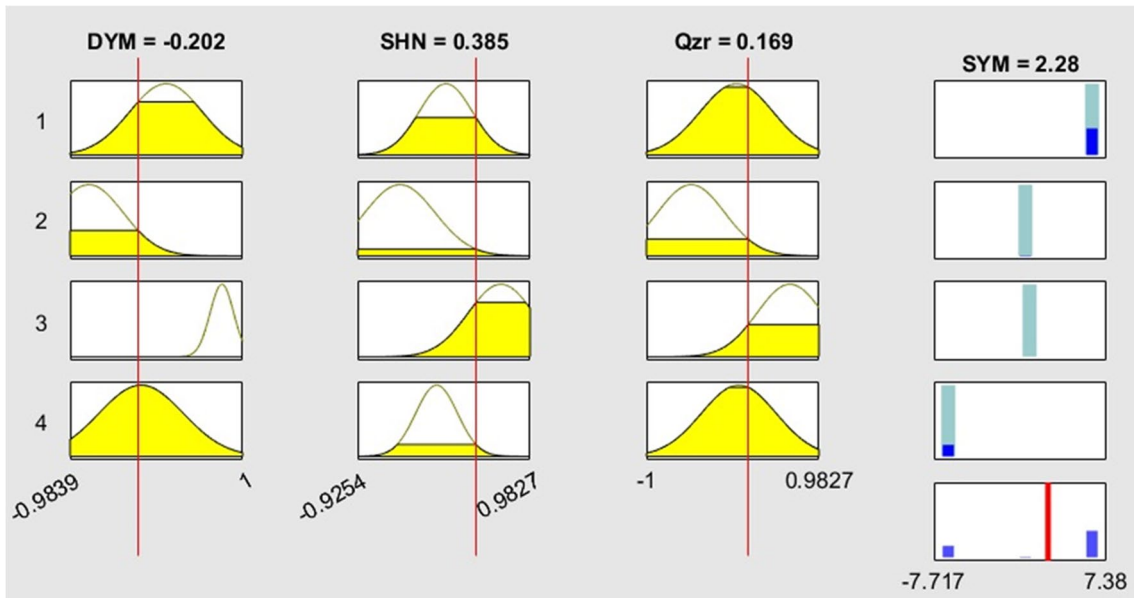
Fig. 13 Rule diagram using ANFIS method in SYM prediction

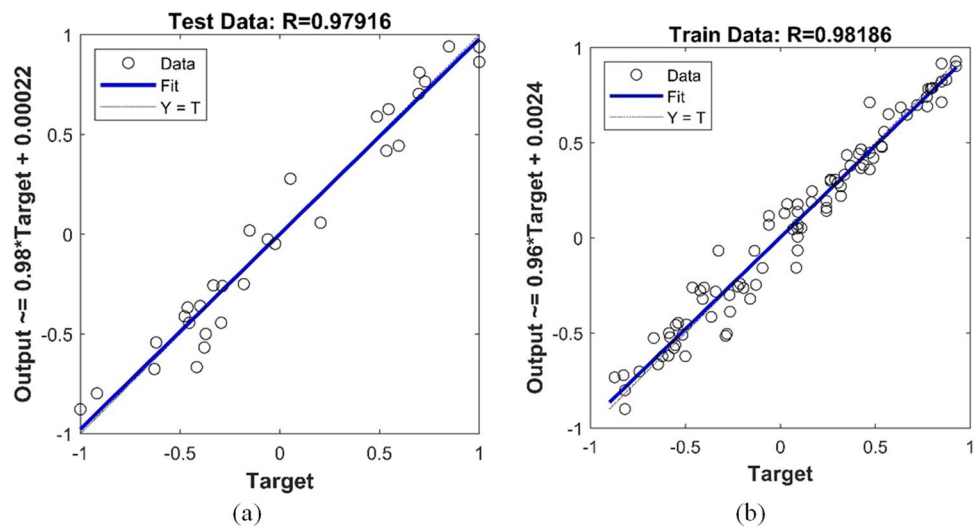**Fig. 14** The ANFIS performance in SYM prediction

Table 7 Results of neuron evaluation using the BPMLP model

Neuron number	R	MSE
1	0.97	0.03
2	0.97	0.04
3	**0.98**	**0.02**
4	0.96	0.02

proved advantageous in technical and engineering problems (Rastegarnia et al. 2021; AlHamad et al. 2022). Figure 15 shows the results of the ideal BPMLP model in SYM prediction.

The SVR results

SVR modeling results are summarized according to Table 8:

Factors affecting SVR are summarized below (Joseph and Swalih 2023; Fang et al. 2023).

- Curse of dimensionality: As the number of input variables increases, the feature space expands, leading to sparser data points in high-dimensional spaces. This makes it difficult for SVR to find an optimal hyperplane that effectively separates classes.
- Overfitting: introducing more input variables can result in overfitting, where the model becomes overly com-

**Fig. 15** The BPMLP performance in SYM prediction

(a)

(b)
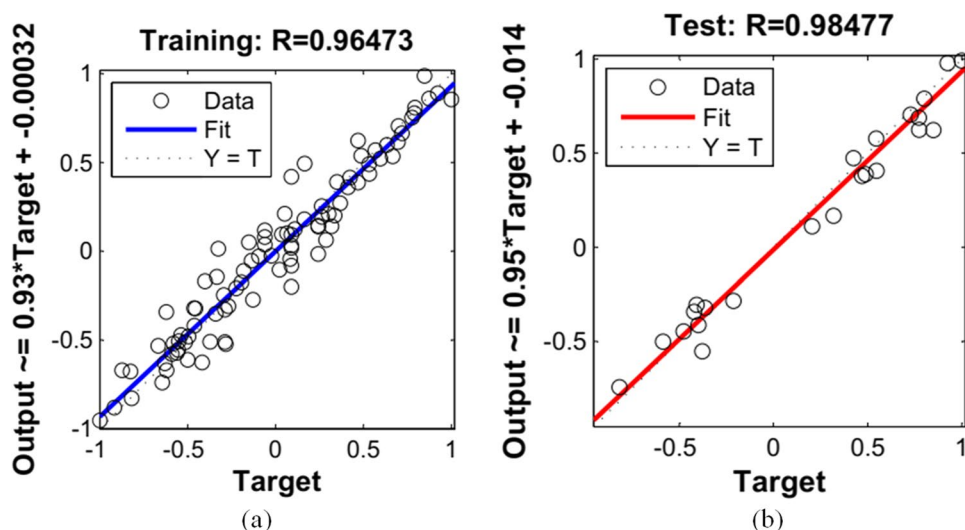
Table 8 Results of the SVR method to estimate SYM

Phase	Details
Dataset split	The dataset for SYM modeling using the SVR method was divided randomly into a training set consisting of 92 cases and a test set containing 30 cases
Kernel function parameter optimization	To optimize the achieved parameters of the SVR model, such as epsilon = 0.02, c = 51, and Gamma = 1, a trial-and-error approach was employed
Best SVR performance	The best performance in SVR was achieved with the implementation of the RBF kernel using four inputs, yielding the most favorable results
Visualization of results	Figure 16 illustrates the most accurate results obtained with SVR utilizing the RBF kernel for SYM estimation, specifically employing SHN, DYM, and Qzr inputs

Fig. 16 The SVR performance in SYM prediction

(a)

(b)
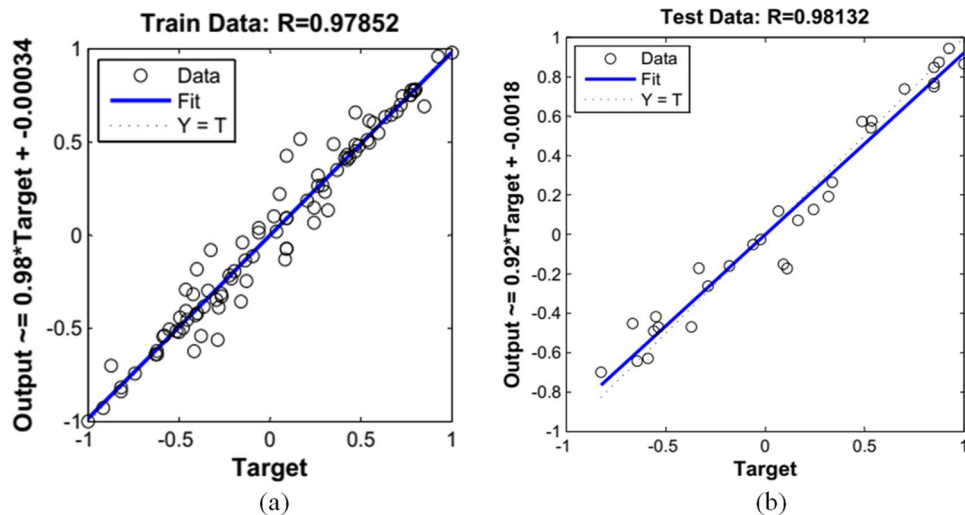
plex and captures noise from the training data instead of underlying patterns. Consequently, the model may struggle to generalize well to unseen data.

- Increased computational complexity: A higher number of input variables can escalate the computational require-

ments for SVR training and prediction. This can lead to longer processing times and resource-intensive operations, impeding practical application.

- Difficulty in kernel selection: SVR often relies on kernel functions to map data into higher-dimensional spaces.

Table 9 Results of statistical criteria to estimate SYM based on all data sets

Approaches	R^2	MAPE%	RMSE	VAF	CPM
BPMLP	0.96	4.87	0.14	0.95	1.77
MLR	0.95	6.24	1.01	0.94	0.88
SVR	0.98	4.68	0.11	0.97	1.84
ANFIS	0.97	6.26	0.12	0.96	1.81

With an increasing number of input variables, selecting an appropriate kernel and fine-tuning its parameters becomes more challenging, impacting the model's overall performance.

- To tackle these challenges, it is crucial to employ feature selection or dimensionality reduction techniques. These methods help identify and retain only the most relevant input variables, mitigating overfitting, alleviating the curse of dimensionality, and enhancing the accuracy of the SVR model.

Method comparisons and discussions

In our evaluation of various models, we examined different statistics, as illustrated in Table 9. Notably, SVR demonstrated the highest accuracy, characterized by CPM and %MAPE values in forecasting SYM. All methods displayed high R^2 values, typically around 0.95 or 0.98, indicating a robust ability to explain variance in SYM prediction. Based on the results, the performance of the SVR is better than the BPMLP and other models (Table 9). BPMLP uses the risk minimization method (ERM) but SVR uses the structural risk minimization method. Also, the precision of these approaches is influenced by the number of inputs, the number of samples, and the choice of training algorithm (Fang et al. 2024; Gautam et al. 2023). Khajevand (2023b) evaluated MLR, ANFIS, and BPMLP in predicting UCS of various rock samples. The outcomes showed that the ANFIS produced the most precise results. Khan et al. (2022) stated that the RFR outperforms KNN (K-nearest neighbors), MLR, and BPMLP in guessing rock UCS. Fang et al. (2024) demonstrated that the SVR provides the most correct outcomes for forecasting rock tensile strength among BPMLP, MLR, RFR, and CART.

Each of the methods successfully displayed the correlations between input parameters and their impacts on SYM, showcasing strong generalization abilities for new data. The average values of the static modulus using intelligent methods were found to be equal to 22.46GPa, which is about 0.04% different from the experimental value (22.47GPa). This discrepancy is less than 1% and indicates the efficiency of the methods used in estimating Young's modulus.

Conclusions

In this research, after evaluating petrography, physical and mechanical characteristics of the sandstone samples, intelligent and statistical analyses including SVR, ANFIS, MLR, and BPMLP were employed to estimate the SYM (static Young's modulus) of different types of sandstones found at the Three Gorges and Xiaolangdi dam sites. Petrography studies classified samples into arenite, feldspathic litharenite, and litharenite categories. The distinct sandstone types between the two dam sites resulted in variations in their geomechanical properties. Samples of feldspathic litharenite at the dam sites exhibited the highest strength, attributed to the influence of feldspar, while litharenite and arenite types showed variability in strength based on their mineral compositions. Analysis of fracture mode patterns revealed that samples with low strength values (arenite and litharenite) predominantly exhibited axial fractures, transitioning to multiple fracture types in feldspathic litharenite samples as UCS and SYM increased. Linear models demonstrated the highest correlation coefficients of inputs with SYM estimation. The ratio of quartz and water uptake showed the highest and lowest impacts on SYM, respectively.

The SVR model outperformed other methods in terms of accuracy ($R^2 = 0.98$, RMSE$= 0.11$GPa, CPM $= +1.84$) to estimate SYM. Notably, intelligent models demonstrated no overfitting, with test data correctness equal to or greater than training data. The average difference between the values of static modulus using ML methods and its experimental value is about -0.04%, which is less than 1% and indicates the efficiency of the methods used in estimating static Young's modulus.

In summary, the study highlights effective approaches for selecting the most suitable machine learning methods for accurately predicting rock SYM, particularly beneficial in the site investigation stages of civil engineering projects. Testing on larger datasets and data from diverse sites and lithologies is recommended for enhanced prediction accuracy in future models.

Authors contributions Na Liu, Yan Sun, Zhe Wang and Jiabao Wang: Sampling, Methodology, Performing rock mechanics tests, Writing—Original Draft, Writing—Review and Editing. Jafar Qajar, and Ahmad Rastegarnia: Compilation, Data analysis, Analysis of results, Writing—Review and Editing.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate Not applicable.

Informed consent The authors are fully aware and satisfied with the article's contents.

Competing interest The authors declare no competing interests.

References

Abdelhedi M, Jabbar R, Said AB, Fetais N, Abbes C (2023) Machine learning for prediction of the uniaxial compressive strength within carbonate rocks. Earth Sci Inform 16:1473–1487. https://doi.org/10.1007/s12145-023-00979-9

Abdi Y, Momeni E, Armaghani DJ (2023) Elastic modulus estimation of weak rock samples using random forest technique. Bull Eng Geol Environ 82(5):1–20. https://doi.org/10.1007/s10064-022-02651-9

AlHamad M, Akour I, Alshurideh M, Al-Hamad A, Kurdi BARN, Alzoubi H (2021) Predicting the intention to use google glass: A comparative approach using machine learning models and PLS-SEM. IJDNS 5(3):311–320

AlHamad AQ, Alomari KM, Alshurideh M, Al Kurdi B, Salloum S, Al-Hamad AQ (2022) The adoption of metaverse systems: a hybrid SEM-ML method. In 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) (pp. 1–5). IEEE

Alsalami Z (2023) Modeling of Optimal Fully Connected Deep Neural Network based Sentiment Analysis on Social Networking Data. JSIoT 2(1):114–132

Altindag R (2012) Correlation between P-wave velocity and some mechanical properties for sedimentary rocks. J S Afr Inst Min Metall 112:229–237

Ameen MS, Smart BG, Somerville JM, Hammilton S, Naji NA (2009) Predicting rock mechanical properties of carbonates from wireline logs (A case study: Arab-D reservoir, Ghawar field, Saudi Arabia). Mar Pet Geol 26(4):430–444. https://doi.org/10.1016/j.marpetgeo.2009.01.017

ASTM D2938-95 (2002) Standard Test Method for Unconfined Compressive Strength of Intact Rock Core Specimens. ASTM: West Conshohocken, PA, USA

ASTM D2845-08 (2008) Standard test method for laboratory determination of pulse velocities and ultrasonic elastic constants of rock. West Conshohocken, PA, USA

Azimian A, Ajalloeian R (2015) Empirical correlation of physical and mechanical properties of marly rocks with P wave velocity. Arab J Geosci 8(4):2069–2079. https://doi.org/10.1007/s12517-014-1274-2

Bagherzadeh Khalkhali A, Safarzadeh I, Rahimi Manbar H (2019) Investigating the effect of nanoclay additives on the geotechnical properties of clay and silt soil. J Civ Eng Mater Appl. 3(2):63–74. https://doi.org/10.22034/JCEMA.2019.92088

Basu A, Mishra DA, Roychowdhury K (2013) Rock failure modes under uniaxial compression, Brazilian, and point load tests. Bull Eng Geol Environ 72:457–475. https://doi.org/10.1007/s10064-013-0505-4

Bejarbaneh BY, Bejarbaneh EY, Fahimifar A, Armaghani DJ, Abd Majid MZ (2018) Intelligent modelling of sandstone deformation behavior using fuzzy logic and neural network systems. Bull Eng Geol Environ 77(1):345–361

Bouchaala F, Mohamed AA, Bouzidi JMS, Y, Ali MY, (2023) Azimuthal Investigation of a Fractured Carbonate Reservoir. SPE Reserv Eval Eng 26(03):813–826

Bouchaala F, Ali MY, Matsushima J, Jouini MS, Mohamed, AAI, Nizamudin S (2024) Experimental study of seismic wave attenuation in carbonate rocks. SPE Journal 29(04):1933–1947. Paper Number: SPE-218406-PA. https://doi.org/10.2118/218406-PA

Brotons V, Tomás R, Ivorra S, Grediaga A (2014) Relationship between static and dynamic elastic modulus of calcarenite heated at different temperatures: the San Julián's stone. Bull Eng Geol Environ 73(3):791–799. https://doi.org/10.1007/s10064-014-0583-y

Brotons V, Tomás R, Ivorra S, Grediaga A, Martínez-Martínez J, Benavente D, Gómez-Heras M (2016) Improved correlation between the static and dynamic elastic modulus of different types of rocks. Mater Struct 49(8):3021–3037. https://doi.org/10.1617/s11527-015-0702-7

Chen G, Zhang K, Wang S, Xia Y, Chao L (2023) iHydroSlide3D v10: an advanced hydrological–geotechnical model for hydrological simulation and three-dimensional landslide prediction. Geosci. Model Dev. 16(10):2915–2937. https://doi.org/10.5194/gmd-16-2915-2023

Daraei A, Zare SA (2019) Model between Dynamic and Static Moduli of Limestone in Asmari Geological Formation based on Laboratory and In-situ Tests. JE;G 12(4):617–634 (http://jeg.khu.ac.ir/article-1-2526-fa.html)

Davarpanah SM, Ván P, Vásárhelyi B (2020) Investigation of the relationship between dynamic and static deformation moduli of rocks. Geomech Geophys Geo-Energy Geo-Resour 6(1):1–14. https://doi.org/10.1007/s40948-020-00155-z

Diaz-Acosta A, Bouchaala F, Kishida T, Jouini MS, Ali MY (2023) Investigation of fractured carbonate reservoirs by applying shear-wave splitting concept. Adv Geo-Energy Res 7(2):99–110. https://doi.org/10.46690/ager.2023.02.04

Dutta A, Sarkar K, Tarun K (2024) Machine learning regression algorithms for predicting the susceptibility of jointed rock slopes to planar failure. Earth Sci Inform 17:2477–2493. https://doi.org/10.1007/s12145-024-01296-5

Ebrahimi P, Ranjbar A, Mohammadi Nia F, Ghimatgar H, Hashemizadeh A (2023) Young's Modulus Estimation Using Machine Learning Methods and Daily Drilling Reports. J Oil Gas Petrochem Technol 10(1):1–24

Etemadi M, Pouraghajan M, Gharavi H (2020) Investigating the effect of rubber powder and nano silica on the durability and strength characteristics of geopolymeric concretes. J Civ Eng Mater Appl 4(4):243–252

Fang Z, Qajar J, Safari K, Hosseini S, Nehdi ML (2023) Application of Non-Destructive Test Results to Estimate Rock Mechanical Characteristics—A Case Study. Minerals 13(4):472. https://doi.org/10.3390/min13040472

Fang Z, Cheng J, Xu C, Xu X, Qajar J, Rastegarnia A (2024) Comparison of machine learning and statistical approaches to estimate rock

tensile strength. Case Stud Constr Mater 20:e02890. https://doi.org/10.1016/j.cscm.2024.e02890

Fei W, Huiyuan HB, Jun Y, Yonghao Z (2016) Correlation of Dynamic and Static Elastic Parameters of Rock. Electron J Geotech Eng 21:1551–1560

Folk RL (1980) Petrology of Sedimentary Rocks. Hemphill, Austin, 600p., http://hdl.handle.net/2152/22930

Gautam R, Sinha A, Mahmood HR, Singh N, Ahmed S, Rathore N, Raza MS (2023) Enhancing Handwritten Alphabet Prediction with Real-time IoT Sensor Integration in Machine Learning for Image. JSIoT 2(1):53–64

Ghafoori M, Rastegarnia A, Lashkaripour GR (2018) Estimation of static parameters based on dynamical and physical properties in limestone rocks. J African Earth Sci 137:22–31. https://doi.org/10.1016/j.jafrearsci.2017.09.008

Ghavami S, Rajabi M (2021) Investigating the Influence of the Combination of Cement Kiln Dust and Fly Ash on Compaction and Strength Characteristics of High-Plasticity Clays. J Civ Eng Mater Appl 5(1):9–16. https://doi.org/10.22034/JCEMA.2020.250727.1040

Guo S, Zhang Y, Iraji A, Gharavi H, Deifalla AF (2023) Assessment of rock geomechanical properties and estimation of wave velocities. Acta Geophys 71(2):649–670. https://doi.org/10.1007/s11600-022-00891-8

Hadi F, Nygaard R (2023) Estimating unconfined compressive strength and Young's modulus of carbonate rocks from petrophysical properties. Pet Sci Technol 41(13):1367–1389

Hasheminezhad A, Sadeghi A (2023) Indirect Estimation of Uniaxial Compressive Strength of Limestone Using Rock Index Tests Through Computational Methods. JCEMA 7(3):161–168. https://doi.org/10.22034/JCEMA.2023.393871.1107

He MY, Dong JB, Jin Z, Liu CY, Xiao J, Zhang F, Sun H, Zhao ZQ, Gou LF, Liu WG, Luo CG (2021) Pedogenic processes in loess-paleosol sediments: Clues from Li isotopes of leachate in Luochuan loess. GCA 299:151–162. https://doi.org/10.1016/j.gca.2021.02.021

Hussan BK, Rashid ZN, Zeebaree SR, Zebari RR (2023) Optimal deep belief network enabled vulnerability detection on smart environment. JSIoT 2022(1):146–162

ISRM (1981) Rock characterization testing and monitoring. In: Brown, E.T. (Ed.), ISRM Suggested Methods. Pergamon Press, Oxford

Joseph J, Swalih CKA (2023) Implementation of Machine Learning in Structural Reliability Analysis. JCEMA 7(3):1–9. https://doi.org/10.22034/JCEMA.2023.396301.1108

Kafash Bazari A (2023) A Case Study of Cement Performance in Different Concretes. JCEMA 7(1):23–40. https://doi.org/10.22034/JCEMA.2023.365342.1098

Khajevand R (2022) Soft computing approaches for evaluating the slake durability index of rocks. Arab J Geosci 15:1698. https://doi.org/10.1007/s12517-022-10997-4

Khajevand R (2023a) Estimating Geotechnical Properties of Sedimentary Rocks Based on Physical Parameters and Ultrasonic P-Wave Velocity Using Statistical Methods and Soft Computing Approaches. Iran J Sci Technol Trans Civ Eng 47:3785–3809. https://doi.org/10.1007/s40996-023-01148-0

Khajevand R (2023b) Prediction of the Uniaxial Compressive Strength of Rocks by Soft Computing Approaches. Geotech Geol Eng 41:3549–3574. https://doi.org/10.1007/s10706-023-02473-x

Khajevand R (2023c) Determining Dry and Saturated Strength of Rocks by Using the Schmidt Hammer. Iran J Sci 47:779–790. https://doi.org/10.1007/s40995-023-01436-4

Khan NM, Cao K, Yuan Q, Bin Mohd Hashim MH, Rehman H, Hussain S, Emad MZ, Ullah B, Shah KS, Khan S (2022) Application of machine learning and multivariate statistics to predict uniaxial compressive strength and static Young's modulus using physical

properties under different thermal conditions. Sustainability 14(16):9901

Kheirollahi H, Manaman NS, Leisi A (2023) Robust estimation of shear wave velocity in a carbonate oil reservoir from conventional well logging data using machine learning algorithms. J Appl Geophys 211:104971. https://doi.org/10.1016/j.jappgeo.2023.104971

Khosravi M, Tabasi S, Eldien HH, Motahari MR, Alizadeh SM (2022) Evaluation and prediction of the rock static and dynamic parameters. J Appl Geophys 199:104581. https://doi.org/10.1016/j.jappgeo.2022.104581

Köken E (2021) Assessment of deformation properties of coal measure sandstones through regression analyses and artificial neural networks. Arch Min Sci 66(4):523–542. https://doi.org/10.24425/ams.2021.139595

Kookalani S, Cheng B (2021) Structural analysis of GFRP elastic grid shell structures by particle swarm optimization and least square support vector machine algorithms. J Civ Eng Mater Appl 5(2):139–150. https://doi.org/10.22034/JCEMA.2021.304981.1064

Kotsanis D, Nomikos P, Rozos D (2021) Comparison of Static and Dynamic Young's Modulus of Prasinites. Mater Proc 5(1):54. https://doi.org/10.3390/materproc2021005054

Kurtulus C, Bozkurt A, Endes H (2012) Physical and Mechanical Properties of Serpentinized Ultrabasic Rocks in NW Turkey, Pure Appl. Geophys 169:1205–1215

Li C, Dias D (2023) Assessment of the Rock Elasticity Modulus Using Four Hybrid RF Models: A Combination of Data-Driven and Soft Techniques. Appl Sci 13(4):2373

Li X, Fan G (2024) On strain localization of aeolian sand in true tri-axial apparatus. Acta Geotech 19(5):3115–3128. https://doi.org/10.1007/s11440-024-02273-4

Li J, Zhang Y, Lin L, Zhou Y (2023) Study on the shear mechanics of gas hydrate-bearing sand-well interface with different roughness and dissociation. Bull Eng Geol Environ 82(11):404. https://doi.org/10.1007/s10064-023-03432-9

Liang S, Zhao Z, Li C, Yin Y, Li H, Zhou J (2024) Age and petrogenesis of ore-forming volcanic-subvolcanic rocks in the Yidonglinchang Au deposit, Lesser Xing'an Range: Implications for late Mesozoic Au mineralization in NE China. Ore Geol Rev 165:105875. https://doi.org/10.1016/j.oregeorev.2024.105875

Liu W, Zhou H, Zhang S, Zhao C (2023) Variable Parameter Creep Model Based on the Separation of Viscoelastic and Viscoplastic Deformations. Rock Mech Rock Eng 56(6):4629–4645. https://doi.org/10.1007/s00603-023-03266-7

Mahmoud AA, Elkatatny S, Al Shehri D (2020) Application of machine learning in evaluation of the static young's modulus for sandstone formations. Sustainability 12(5):1880

Maleki MA, Emami M (2019) Application of SVM for investigation of factors affecting compressive strength and consistency of geopolymer concretes. J Civ Eng Mater Appl 3(2):101–107. https://doi.org/10.22034/JCEMA.2019.92507

Manda SR, Raj A, Andraju N (2023) A machine learning approach to predict geo-mechanical properties of rocks from well logs. IJDSA 1–18. https://doi.org/10.1007/s41060-023-00451-3

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

Motahari MR, Amini O, Iraji A, Mahdizadeh Gohari O, Saffarian M (2022) Comparison of dynamic and static properties of sandstone and estimation of shear wave velocity and Poisson's ratio. Bull Eng Geol Environ 81(9):384–396. https://doi.org/10.1007/s10064-022-02867-w

Najibi A, Mohammadreza M, Ajal Louian R, Gholam Abbas S (2011) Estimation of mechanical properties of limestone using

petrophysical data. J Eng Geol 5(1):1–18. https://civilica.com/doc/280877 (In Persian)

Onaloa D, Oloruntobi O, Adedigba S, Khan F, James L, Butt S (2018) Static Young's modulus model prediction for formation evaluation. J Pet Sci Eng. 171:394–402. https://doi.org/10.1016/j.petrol.2018.07.020

Pappalardo G, Mineo S (2022) Static elastic modulus of rocks predicted through regression models and Artificial Neural Network. Eng Geol 308:106829

Pereira ML, da Silva PF, Fernandes I, Chastre C (2021) Characterization and correlation of engineering properties of basalts. Bull Eng Geol Environ 80:2889–2910. https://doi.org/10.1007/s10064-021-02107-7

Rahman T, Sarkar K (2023a) Correlations between uniaxial compressive strength and dynamic elastic properties for six rock types. Int J Geomech 23(6):04023064

Rahman T, Sarkar K (2023b) Empirical correlations between uniaxial compressive strength and density on the basis of lithology: implications from statistical and machine learning assessments. Earth Sci Inform 16:1389–1403. https://doi.org/10.1007/s12145-023-00969-x

Rastegarnia A, Alizadeh SMS, Esfahani MK, Amini O, Utyuzh AS (2020) The effect of hydrated lime on the petrography and strength characteristics of Illite clay. Geomech Eng. 22(2):143. https://doi.org/10.12989/gae.2020.22.2.143

Rastegarnia A, Lashkaripour GR, Sharifi T, Teshnizi E, Ghafoori M (2021) Evaluation of engineering characteristics and estimation of static properties of clay-bearing rocks. Environ Earth Sci 80(18):1–24. https://doi.org/10.1007/s12665-021-09914-x

Rastegarnia A, Ghafoori M, Moghaddas NH, Lashkaripour GR, Shojaei H (2022) Application of cuttings to estimate the static characteristics of the dolomudstone rocks. Geomech Eng 29(1):65–77

Salehi R, Ajallouian R, Hashemi M (2011) Comparison of Modulus of Dynamic and Static Elasticity of Bazaft Dam Stones, 4th National Geological Conference, Payame Noor University of Mashhad, Iran

Shahani NM, Zheng X, Guo X, Wei X (2022) Machine learning-based intelligent prediction of elastic modulus of rocks at thar coalfield. Sustainability 14(6):3689

Shakir AK (2023) Optimal Deep Learning Driven Smart Sugarcane Crop Monitoring on Remote Sensing Images. JSIoT 22(1):163–177

Shamsashtiany R, Ameri M (2018) Road Accidents Prediction with Multilayer Perceptron MLP modelling Case Study: Roads of Qazvin. Zanjan and Hamadan JCEMA 2(4):181–192

Sharifi J, Moghaddas NH, Saberi MR, Mondol NH (2023) A novel approach for fracture porosity estimation of carbonate reservoirs. Geophys Prospect 71(4):664–681. https://doi.org/10.1111/1365-2478.13321

Sharifi J, Nooraiepour M, Mondol NH (2021) Application of the Analysis of Variance for Converting Dynamic to Static Young's Modulus. In: 82nd EAGE annual conference and exhibition, 1(2):1–5. https://doi.org/10.3997/2214-4609.202012000

Shirnezhad Z, Azma A, Foong LK, Jahangir A, Rastegarnia A (2021) Assessment of water resources quality of a karstic aquifer in the Southwest of Iran. Bull Eng Geol Environ 80:71–92

Song F, Liu Y, Shen D, Li L, Tan J (2022) Learning Control for Motion Coordination in Water Scanners: Toward Gain Adaptation. IEEE Trans Ind Electron 69(12):13428–13438. https://doi.org/10.1109/TIE.2022.3142428

Soustelle V, ter Heege J, Wassing B (2023) Relationship between static and dynamic elastic moduli in carbonates. In ARMA US Rock Mechanics/Geomechanics Symposium. ARMA. Paper Number: ARMA- 2023–0331, https://doi.org/10.56952/ARMA-2023-0331

Taheri S, Ziad H (2021) Analysis and Comparison of Moisture Sensitivity and Mechanical Strength of Asphalt Mixtures Containing Additives and Carbon Reinforcement. JCEMA 5(1):01–08. https://doi.org/10.22034/jcema.2021.128042

Tashayo B, Honarbakhsh A, Azma A, Akbari M (2020) Combined fuzzy AHP–GIS for agricultural land suitability modeling for a watershed in southern Iran. Environ Manag 66:364–376

Teshnizi ES, Golian M, Sadeghi S, Rastegarnia A (2021) Application of analytical hierarchy process (AHP) in landslide susceptibility mapping for Qazvin province. N Iran Comp Earth Environ Sci 2(4):55–95. https://doi.org/10.1016/B978-0-323-89861-4.00041-5

Tofighkhah M, Hashemidanesh N, Ameri M (2023) Investigating the Resistance Behavior of the Clayey Sand Soil Improved with Nano-Silica and Carbon Fibers. J. Civ. Eng. Mater. Appl, 7(1).https://doi.org/10.22034/JCEMA.2023.171494

Wang M, Wu Y, Song B, Xu W (2024b) Point Load Strength Test Power Index of Irregular Sandy Dolomite Blocks. Rock Mech. Rock Eng. https://doi.org/10.1007/s00603-023-03733-1

Xiao D, Liu M, Li L, Cai X, Qin S, Gao R, Liu J, Liu X, Tang H, Li G (2023) Model for economic evaluation of closed-loop geothermal systems based on net present value. Appl Therm Eng 231:121008. https://doi.org/10.1016/j.applthermaleng.2023.121008

Xie Y, Wang L, Gu Y, Gu X, Chen S, Khajehzadeh M, Hosseini S (2024) Prediction of Rock's Brittleness and Dynamic Properties Utilizing Effective Artificial Intelligence Approaches. Period. Polytech.: Civ. Eng. https://doi.org/10.3311/PPci.23156

Yasar E, Erdogan Y (2004) Correlating sound velocity with the density, compressive strength and Young's modulus of carbonate rocks. Int J Rock Mech Min Sci 41(5):871–875

Yu J, Zhu Y, Yao W, Liu X, Ren C, Cai Y, Tang X (2021) Stress relaxation behaviour of marble under cyclic weak disturbance and confining pressures. Measurement 182:109777. https://doi.org/10.1016/j.measurement.2021.109777

Zhang Y, Gu H, Zhou L, Yang L (2023) Extracting static elastic moduli of rock through elastic wave velocities. Acta Geophys 72:915–931. https://doi.org/10.1007/s11600-023-01139-9

Zhao Z, Zhao X, Yin Y, Li C, Yang Y, Wang Y (2024) Identification of Geochemical Anomalies Based on RPCA and Multifractal Theory: A Case Study of the Sidaowanzi Area, Chifeng, Inner Mongolia. ACS Omega. https://doi.org/10.1021/acsomega.4c02078