Rock Mechanics and Rock Engineering (2020) 53:4691–4715
https://doi.org/10.1007/s00603-020-02184-2

ORIGINAL PAPER

Machine Learning Approach to Model Rock Strength: Prediction and Variable Selection with Aid of Log Data

Mohammad Islam Miah[1,2] · Salim Ahmed[1] · Sohrab Zendehboudi[1] · Stephen Butt[1]

Received: 18 February 2020 / Accepted: 17 June 2020 / Published online: 28 June 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

Comprehensive knowledge and analysis of in situ rock strength and geo-mechanical characteristics of rocks are crucial in hydrocarbon and mineral exploration stage to maximize wellbore performance, maintain wellbore stability, and optimize hydraulic fracturing process. Due to the high cost of laboratory-based measurements of rock mechanics properties, the log-based prediction is a viable option. Nowadays, the machine learning tools are being used for estimation of the in situ rock properties using wireline log data. This paper proposes a machine learning approach for rock strength (uniaxial compressive strength) prediction. The main objectives are to investigate the performance of data-driven predictive model in determining this vital parameter and to select features of predictor log variables in the model. The backpropagation multilayer perception (MLP) artificial neural network (ANN) with Levenberg–Marquardt training algorithm as well as the least squares support vector machine (LS-SVM) with coupled simulated annealing (CSA) optimization technique is employed to develop the dynamic data-driven models. Capturing nonlinear, high dimensional, and complex nature of real field log data, the rock strength models' performances are evaluated using statistical criteria to ensure concerning the model reliability and accuracy. The model predictions are compared and validated against the measured values as well as the results obtained from existing log-based correlations. Both the MLP-ANN and the CSA-based LS-SVM connectionist strategies are able to predict the rock strength so that there is a very good match between the model results and corresponding measured values. The input log parameters are ranked based on their contributions in prediction performance. The acoustic travel time and gamma ray are found to have the highest relative significance in estimating rock strength. New correlations are also developed to obtain the in situ rock strength of the siliciclastic sedimentary rocks using the most important log parameters such as dynamic sonic slowness, formation electron density, and shalyness effect. The developed correlations can be used to obtain quick estimation of dynamic uniaxial compressive strength profile using wireline logging data, instead of static data from the surface measurements or laboratory data. It is expected that the proposed models and tools will enable oil and gas engineers to better predict rock strength and thus enhance wellbore performance.

Keywords Log variables · Connectionist models · Uniaxial compressive strength · Variable selection · Rock stability

Abbreviations

AAPE	Average absolute percentage error
ANN	Artificial neural network
BTS	Brazilian tensile strength
DT	Sonic travel time (μs/ft)
FL	Fuzzy logic
GEP	Gene expression programming
GR	Gamma ray (API)
LS-SVM	Least square support vector machine
LM	Levenberg–Marquardt
MAPE	Maximum absolute error percentage
MSE	Mean square error
MLP	Multilayer perception
MVRE	Multivariate regression analysis
RBF	Radial basis kernel function
RMSE	Root mean square error
UCS	Unconfined compressive strength (MPa), rock strength

✉ Mohammad Islam Miah
mim625@mun.ca

[1] Department of Process Engineering, Memorial University, St. John's, NL, Canada

[2] Department of Petroleum and Mining Engineering, Chittagong University of Engineering and Technology, Chittagong 4349, Bangladesh

List of symbols

b	Bias
dd	Dry density
E	Young's modulus (MPa)
G	Shear modulus (MPa)
GR_{log}	Gamma-ray value of the zone of interest
GR_{max}	Maximum value of gamma-ray log over the entire log
GR_{min}	Minimum value of gamma-ray log over the entire log
I_d	Slake durability index
$Is_{(50)}$	Point load index test
I_{GR}	Shale Index (Clay index)
K	Bulk modulus (MPa)
N	Number of neurons
NPHI	Neutron porosity
PHI	Effective porosity
PHIN	Neutron porosity (frac.)
PHIND	Porosity from the combination of density and neutron log
R^2	Coefficient of determination
RB	Formation bulk density
RT	True (Deep) resistivity (Ω m), Rt
SRn	Schmidt hammer rebound number harness number (SRn)
Vp	Compressional wave velocity (km/s)
Vs	Shear wave velocity (km/s)
Vsh	Shale volume (shaliness)
wc	Water content
x_i	Input variables
y_p	Predicted value (y)
y_m	Target (actual) output variable (y)

Greek letters

ϕ	True porosity (frac.)
ϕ_e	Effective porosity (frac.)
$\phi_{D,e}$	Effective density porosity (frac.)
$\phi_{N,e}$	Effective neutron porosity (frac.)
$\phi_{N,sh}$	Neutron porosity of the adjacent shale zone (frac.)
ν	Poison's ratio
ρ_b	Bulk density (g/cm^3)
$\rho_{b\cdot c}$	Clay corrected density porosity (g/cm^3)
ρ_{fl}	Fluid density (g/cm^3)
ρ_{ma}	Matrix density (g/cm^3)
ω_{ij}	Connected weight between the ith neuron and jth neuron

1 Introduction

1.1 Background

Reservoir rock mechanical properties (RMPs) play an important role in making decisions regarding exploration development, wellbore stability analysis, sanding potentiality, and safe drilling operations. The RMPs are closely related to the rock physical properties such as rock density, porosity, and pore fluid saturation. The uniaxial compressive strength (UCS) is one of the key rock strength parameters that can be used in wellbore stability assessment, in situ stress analysis, drilling optimization as well as reservoir sanding likelihood prediction (Nouri et al. 2006; Crawford et al. 2011). It is also utilized for mining and geotechnical engineering applications including the rock mass characterization, underground rock excavation, and geotechnical project design (Brown 1981). The reservoir rock strength provides the limit of axial stress in situ condition before rock failing (Fjær et al. 2008). The most common techniques for determining rock strength are in situ measurements using geophysical log-based models and laboratory measurements using core samples. The direct and indirect measurements of rock strength include UCS test, triaxial test, Brazilian test, and basic mechanical tests such as point load index, non-destructive ultrasonic test, Schmidt hammer rebound test, impact strength test, scratch test, indentation test, and Equotip hardness tester (Miller 1965; Broch and Frankline 1972; Hoek and Brown 1997; Fener et al. 2005; Yilmaz 2013; Mousavi et al. 2018). The Rock Mechanical Laboratory (RML) testing on real core samples is the most reliable method to determine the rock strength using the procedures recommended by the American Society for Testing Materials (ASTM 1986) and/or the International Society of Rock Mechanics (Brown1981). The RML testing cannot provide continuous rock strength profile along with the in situ wellbore conditions. Also, the core sample preparation is tedious, time consuming, expensive, and very sensitive to stress unloading (Raaen et al. 1996). Most of the times, high-quality cylindrical core specimens preparation with regular geometry is difficult due to the nature of fractured thinly bedded and clay-bearing sedimentary rocks. In addition, a large-scale core sample preparation is often challenging due to time limitation, complex variations in rock composition, and geological behavior and heterogeneity of the formation. When the formation rock sample is not attainable from the deeper depth of oil and gas and/or mining fields, the well logging data can be used to estimate the rock mechanical and physical properties. Over the past few decades, numerous studies have been performed to develop empirical models for estimation of rock strength using

petrophysical logs such as the sonic and density porosity logs. The log-based existing models or correlations for rock strength were listed/summarized by different authors such as Chang et al. (2006) and Odunlami et al. (2011). A continuous in situ UCS profile is essential to conduct rock failure analysis during a drilling operation. Nowadays, the artificial intelligence (AI) approaches such as artificial neural network (ANN), functional network (FN), adaptive neuro-fuzzy inference system (ANFIS), decision tree (DT), classic support vector machine (SVM), and/or least square support vector machine (LS-SVM) are extensively used to develop the data-driven models for rock mechanics related to petroleum and mining (Asoodeh and Bagheripour 2012; Ocak and Seker 2012; Khandelwal and Monjezi 2013; Ashena and Thonhauser 2015; Dehghan et al. 2010; Koolivand-Salooki et al. 2017; Ceryan and Can 2018; Anemangely et al. 2019; Miah et al. 2019). Recently, ANN and LSSVM are becoming more popular strategies for data-driven model development, prediction of formation elastic properties, and failure analysis to save the operational expenses and time.

1.2 Literature survey

Several empirical correlations exist for calculation of UCS along the wellbore using rock porosity (ϕ), sonic travel time (DT), elastic moduli, and other formation properties such as shale volume and resistivity (Fjaer et al. 1992; Edimann et al. 1998; Sonmez et al. 2004; Chang et al. 2006; Odunlami et al. 2011; Singh et al. 2012; Rabbani et al. 2012). The rock strength can be estimated from the empirical correlations considering the lithology type (e.g., carbonate or clastic rocks) and availability of log variables. In addition, there are so-called "worldwide" models, which can be used irrespective of the rock type (Chang et al. 2006). Well-known correlations for estimating rock strength (UCS) using log data in the oil and gas industry are listed in Table 1. Onyia (1988) investigated the in situ rock strength and developed a correlation to estimate ultimate rock strength using drilling and logging data such as gamma-ray (GR) log, resistivity log, density log, and borehole compensated sonic logs. Fjaer et al. (1992) developed a correlation that is most commonly used to obtain the rock strength for shaly sand reservoirs using three rock parameters, namely, Poisson's ratio (ν), sonic velocity (Vp), and shale volume (Vsh). The shale volume measurement is a costly and time-consuming approach through core analysis, while it can be found from the wireline log data including GR log. For instance, the shale volume can be obtained using the entire GR log of formation to capture the maximum and minimum magnitudes of GR if the full-length log is available. Moos et al. (2001) developed the UCS correlation for the clastic sandstone rocks using sonic travel time or acoustic velocity, and formation bulk density. Also, Chang et al. (2006) took into account only sonic travel time to introduce the UCS correlation. Both correlations of Moos et al. (2001) and Chang et al. (2006) did not consider the shale content effect (e.g., GR parameter) to estimate the in situ UCS parameter. A number of studies have investigated the influences of various rock properties on rock strength; however, the ranking of variables was not discussed in their studies (Rajabzadeh et al. 2012; Haftani et al. 2015; Jamshidi et al. 2016, 2018; Kong and Shang, 2018). The existing UCS correlations have been developed considering static formation properties from core analysis, as well as a few influential log variables.

In situ UCS models have not been developed yet to capture the lithology indicator (GR), number of electron density, and acoustic travel time of the porous shay sand formation. Thus, it seems necessary to develop new correlation(s) for estimation of realistic continuous in situ rock strength profile using dynamic wireline logging data in the absence of drilling or laboratory-based data.

The connectionist models, for example ANN and LS-SVM, are being increasingly used to predict rock mechanical properties for wellbore failure analysis, reservoir geomechanics, underground excavation as well as rock mass characterization (Yang and Zhang 1997; Yılmaz and Yuksek 2008; Ocak and Seker 2012; Asoodeh and Bagheripour 2012; Khandelwal and Monjezi 2013; Barzegar et al. 2016; Negara et al. 2017; Matin et al. 2018; Abdi et al. 2018; Onalo et al. 2018; Ceryan and Can 2018; Onalo 2019). Although several research works have attempted to determine the elastic or geomechanical characteristics of rocks using experimental core and log data, only a few studies have forecasted the rock strength (UCS) using wireline logs with AI or machine learning tools. The existing literature reveals that several input variables are adopted in soft computing smart models to obtain the UCS or Young's modulus (E) using rock physical properties or features such as porosity, unit weight (uw), dry density (dd), compressional wave velocity (Vp), water absorption (wa), water content (wc), Schmidt hammer rebound number or hardness number (SRn), Brazilian tensile strength (BST), point load index ($Is_{(50)}$), and slake durability index (I_d). The main studies conducted for rock strength prediction are listed in Table 2 with a focus on the corresponding research methods and shortcomings. Only a few studies investigated ranking of the input variables while determining the rock strength (UCS) (Majidi and Rezaci 2013; Torabi-Kaveh et al. 2015; Matin et al. 2018). Also, the use of AI approaches to estimate elastic properties and rock strength parameter was found to be limited (Sharma et al. 2010; Tariq et al. 2017; Onalo et al. 2018). Sharma et al. (2010) introduced a correlation between UCS and physical properties of weakly cemented sandstone rocks such as RB, synthetic Vp and Vs, and Vsh

Table 1 Selected empirical correlations to estimate rock strength (UCS, MPa) for sandstone reservoir rocks

References	Input variables	Correlation for UCS	Region where developed	Comments/remarks
Freyburg (1972)	Vp	$0.035V_p - 31.5$	Germany	It was developed for the relatively strong rocks
McNally (1987)	DT	$1200\exp(-0.036\ \mathrm{DT})$	Bowen Basin, Australia	It is applicable to sandstones (fine grained, both consolidated and unconsolidated)
Onyia (1988)	R_t, GR, RB, DT	$0.98(P + Q\log_{10}(R_t) + R \times \mathrm{GR} + S \times \mathrm{RB} + T \times \mathrm{DT})$	Rogers Country, Oklahoma	The correlation was developed based on multivariate regression analysis (MVRA) incorporating log properties
Fjaer et al. (1992)	ν, Vsh, Vp	$3.3 \times 10^{-20}\rho^2 V_p^4(1-2\nu)\{(1+\nu)^2/((1-\nu)^2\}(1+0.78\mathrm{Vsh})$	Gulf coast	This correlation is applicable for sandstones if UCS > 30 MPa Shale effect was considered; it is applicable for shaly sand reservoirs
Vernik et al. (1993)	ϕ	$254(1-2.7\phi)^2$	Global	It can be used for very clean, well-consolidated sandstones with $0.3 > \phi$
Sarda et al. (1993)	ϕ	a. $(357\exp(-10.8\phi)$ b. $(258\exp(-9\phi)$ c. $115\exp(-11.6\phi)$	–	It is developed using Germingy-sous-Coulombs reservoir It is applicable for the range of $0.07 > \phi > 0$ (Eq. a)., $0.3 > \phi$ (Eq. b), and $\phi > 0.3$ (Eq. c)
Farquhar et al. (1994)	ϕ	$(208.08\exp(-0.074\phi)$	North Sea	It was developed for sandstones
Weingarten and Perkins (1995)	E, K, Vsh	$(114 + 97 \times \mathrm{Vsh})K \times \mathrm{Ed}$	USA	It is applicable for unconsolidated sandstone reservoir
Raaen et al. (1996)	DT	$1.40 - 2.15\mathrm{DT} + 0.0083 \times \mathrm{DT}^2$	Norway, North Sea	It was developed for relatively clean sandstones It is applicable for $140 > \Delta t > 90$ µs/ft
Raaen et al. (1996)	ϕ	$\left(43 - 140\phi + 63\phi^2\right)$	Norway, North Sea	It is applicable for the range of $0.35 > \phi > 0.2$ It is valid for relatively clean sandstone reservoirs
Edimann et al. (1998)	ϕ	$(129.54 - 3.225\phi)$	North Sea	It was developed for sandstones
Bradford et al. (1998)	E	$2.28 + 4.1089E$	Worldwide	It is a more representative model for the sandstone reservoir of the Chicontepec basin
Moos et al. (2001)	Vp, ρ_b	$1.745 \times 10^{-9} \times \mathrm{RB} \times V_p^{-3} - 21$	Cook Inlet, Alaska	It was developed for coarse-grained sandstones and conglomerates
Chang et al. (2006)	DT	$1.4138 \times 10^7 \mathrm{DT}^{-3}$	Gulf Coast	It is applicable for the weak and unconsolidated sandstone rocks

Table 2 Applications of AI-based connectionist tools for rock strength and mechanical properties prediction

References	Input variables	AI approach	Variable ranking	Comments/Remarks
Yilmaz and Yuksek (2009)	SRn, $I_{S(50),}$ wc, sonic velocity	ANN, ANFIS	No	Authors used tansig transfer function for the feedforward neural network that consists of four neurons for an input layer, nine neurons for the hidden layer, and one neuron for the output layer The predictive model performance was acceptable with a root means square error (RMSE) of 5.98% and an correlation coefficient (R^2) of 0.887
Sharma et al. (2010)	RB, Vsh, Vp, Vs	ANN	No	It was concluded that the neural network is better at capturing the complex relationships between geophysical properties and the strength of rock
Rabbani et al. (2012)	Total ϕ, RB, Sw	ANN	No	To train the ANN network, the optimal number of neurons was used for the first layer (20 perceptions) and the second hidden layer (25) using the trial and error approach The correlation coefficient of about 0.903 was found for the UCS deterministic model
Yagiz et al. (2012)	SRn, ϕ, I_d, Vp, uw	ANN	No	It was revealed that ANN gives relatively more precise results than a regression approach
Majdi and Rezaei (2013)	Rock type, SRn, RB,ϕ	ANN	Yes	Authors constructed a multilayer neural network model using different rock samples It offered better results, close to laboratory results, with an R^2 of 0.97, compared to the MVRA It was claimed that the most effective variables are density and SRn
Nabaei and Shahbazi (2012)	Well trajectory	ANN	No	The feedforward backpropagation network applied to predict the output using tansig transfer function for the first two layers
Ceryan et al. (2013)	Vp, Vm, total ϕ	ANN	No	Authors constructed the Levenberg–Marquardt algorithm-based network using input variables including solid part (Vm) of the carbonate rock samples
Ceryan (2014)	P-wave durability, ϕ	ANN, SVM	No	Ceryan (2014) concluded that the vector machine approach performs better than the ANN model The model's performance was investigated by adopting porosity and microstructural properties
Momeni et al. (2015)	dd, $Is_{(50)}$, SRn, Vp	ANN	No	It was concluded that the particle swarm optimization (PSO)-based ANN predictive model exhibits greater reliability than the conventional ANN model for limestone and granite It was revealed that the dd and SRn have more importance than other variables ($Is_{(50)}$ and Vp) to predict UCS

Table 2 (continued)

References	Input variables	AI approach	Variable ranking	Comments/Remarks
Mohamad et al. (2015)	Density, BTS, Vp, $Is_{(50)}$	ANN	No	The PSO-based ANN model was employed to find the relationship between input variables and UCS of soft rocks The predictive model showed a good match with real date where R^2 was equal to 0.971
Torabi-Kaveh et al. (2015)	RB, Vp,ϕ	ANN	Yes	The researchers evaluated the performance of the ANN predictive model, and linear and nonlinear regression equations using predictor variables of limestone rocks It was revealed that the Vp has a higher contribution than other input variables
Barzegar et al. (2016)	RRn, Vp,ϕ	ANN, ANFIS, SVM	No	Different AI models were utilized to analyze the model performance It was claimed that SVM performs better than other models for travertine rocks
Behnia et al. (2017)	Quartz content, dd,ϕ	GEP	No	Gene expression programming (GEP) tool was employed to predict the intact rock strength parameters; the statistical criteria were used to assess the models
Tariq et al. (2017)	Ed, RB, Vs, Vp	ANN ANFIS SVM	No	The neural network was developed with tansig transfer function between input layer (4 neurons) and hidden layer (20 neurons) It was concluded that Ed has a higher relative importance than wave velocities and RB in carbonate rocks
Negara et al. (2017)	Grain density, ϕ, elemental spectroscopy	SVR	No	They applied the SVR technique using the core data to build the predictive model It was demonstrated that elemental spectroscopy has a significant impact to estimate the UCS
Rastegarnia et al. (2018)	$Is_{(50)}$, Vp, wc,ϕ	ANN	No	The sigmoidal and linear transfer functions were used between the input and hidden layers, and the hidden and output layers to train the network Levenberg–Marquardt feedforward backpropagation algorithm was adopted to train the data
Matin et al. (2018)	SRn, $Is_{(50)}$, Vp,ϕ	SVR, DT	Yes	It was found that a rain forest is a powerful tool for variable selection, and Vp is the most effective variable in predicting UCS
Onalo et al. (2018)	GR, RB, Vsh	ANN	No	They utilized the ANN model to predict compressional, shear travel time, and elastic moduli The sanding potentiality was also estimated and the results of the ANN-based predictive model and real data are compared The study did not examine variable ranking while estimating UCS with the predictive model

Table 2 (continued)

References	Input variables	AI approach	Variable ranking	Comments/Remarks
Abdi et al. (2018)	dd, wa, Vp,ϕ	ANN	No	It was found that the neural network model is more powerful than the MLR technique to determine the UCS and/or E of sedimentary rocks They proposed the empirical correlations to predict UCS or E using the input variables

(volume of shale). An extended version of the correlation was built through employing multiple regression analysis; the model is valid for the range of 0–4 MPa of UCS using synthetic velocities and multiple parameters from wireline log data. The performance of the regression-based model was also compared with that of an ANN model. Torabi-Kaveh et al. (2015) utilized the ANN model and regressive equation for predicting the rock strength and elasticity modulus of carbonate and limestone rocks. They claimed that higher accuracy can be obtained from the ANN predictive model, compared to the multiple regression analysis (MVRA). Tariq et al. (2017) investigated the applicability of machine learning approaches such as ANN, ANFIS, and SVM to predict UCS. It was concluded that ANN outperforms the other techniques; an empirical correlation to predict the UCS was also proposed that gives better results, compared to a number of existing correlations. It was claimed that four geomechanical parameters such as RB, BP, Vs, and dynamic Young's modulus (E) are sufficient to forecast the rock strength; however, the relative contribution of the variables was not investigated. Onalo et al. (2018) constructed an MLP-based ANN model using input log variables such as GR, Vsh, and RB to predict sonic acoustic velocities. In their study, they obtained the elastic moduli using correlations; the results obtained from the correlations were also compared with ANN predictions. The researchers also investigated sanding possibility as a case study of shaly sand reservoir, Niger delta basin. In their study, ranking of input variables for the predictive ANN model was not discussed. Abdi et al. (2018) applied the ANN and the MLR approaches to obtain the sedimentary rock strength parameters using experimental data sets. It was found that ANN results in higher accuracy than MLR. They introduced a correlation to estimate UCS using data sets for different rock samples of limestone, sandstone, conglomerate, and marl. In their research study, the correlation did not include the lithology type and they did not rank the input variables. Based on the literature review, further comprehensive investigations are required to identify relative contributions of the input variables in prediction of rock strength as well as mechanical properties of shaly sand formations.

1.3 Research objectives

To the best of our knowledge, data-driven predictive model and feature ranking of logging variables to study the rock strength with the aid of logging data have not been investigated systematically. The variable selection and ranking of logging parameters appear to be an important challenge for not only petroleum and mining engineers and drillers but also geologists to estimate ultimate rock strength profile using the wireline log data with acceptable accuracy. The logging data are expensive to assemble and depict the entire formation profile. The current study is planned to fill in the knowledge gap by finding the most contributing predictor log variables, corresponding to their relative performance while predicting the ultimate rock strength. It is believed that the research strategies employed in this study can be less time consuming and cost effective for efficient rock formation evaluation. The objectives of this research work are listed as follows:

- To investigate the performance of rock strength models and compare the predictive performance of deterministic tools through statistical analysis.
- To perform comprehensive parametric sensitivity analysis and find the most contributing variables for predicting UCS.
- Develop new correlations to estimate in situ formation UCS, capturing most effective log parameters.

2 Theory and Model Development

2.1 Log-Based Model to Predict Rock Strength

The strength of rock is a function of rock mineralogy and lithology, acoustic properties, density, porosity, geometric factors, depositional environments, and compaction level.

The most common logging types are the gamma-ray log, resistivity log, and porosity log. The comprehensive information on well logging principles, advantages, limitations and applications can be found in the available literature (Asquith and Krygowski 2004; Mondol 2015). The presence of clay minerals in a formation may strongly affect the rock material properties. The GR log is extensively used to identify the lithology as well as to estimate the shale volume (clay content); this corrects their presence in shaly sand rocks, yielding more accurate effective porosity (ϕ_e). Rock porosity (void space in the rock volume) can be obtained from the responses of neutron log (counts the hydrogen concentration), density log (measures the bulk electron density), and sonic log (estimates acoustic travel time) in a formation. Also, the rock porosity varies primarily with particle size distribution, grain shape, packing arrangement, compaction, cementation, and clay content (Asquith and Krygowski 2004). The conductivity (inverse of resistivity) of a rock can be obtained from the deep induction resistivity log, which converts to true resistivity (RT, Ω m) of the virgin zone in a formation. In the present study, several well log responses are incorporated to estimate the in situ shale volume, acoustic primary velocity, and effective porosity; these rock properties are then used to estimate the rock strength.

The shale index (I_{GR}) is obtained in the current study by the following equation (Schlumberger 1998):

$$I_{GR} = \frac{GR_{log} - GR_{min}}{GR_{max} - GR_{min}}, \tag{1}$$

where GR_{log} is the gamma-ray value of the zone of interest; and GR_{max} and GR_{min} represent the maximum and minimum magnitudes of gamma-ray log over the entire formation, respectively. The clay content or shale volume (V_{sh}) is estimated from the following equation to simulate the nonlinear response of the tertiary rock formation (Larionov 1969):

$$V_{sh} = 0.083(2^{3.7I_{GR}} - 1). \tag{2}$$

The density porosity (ϕ_D) is determined using the formation bulk density response (ρ_b) by the following equation:

$$\phi_D = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_{fl}}. \tag{3}$$

In Eq. (3), ρ_{fl} and ρ_{ma} refer to the density of the drilling mud, and the rock matrix density, respectively.

The clay corrected density porosity ($\phi_{D,e}$) and neutron porosity ($\phi_{N,e}$) are calculated as follows (Miah 2014):

$$\rho_{b\cdot c} = \rho_b + V_{sh}\left(\rho_{ma} - \rho_{cl}\right), \tag{4}$$

$$\phi_{D,e} = \frac{\rho_{ma} - \rho_{b\cdot c}}{\rho_{ma} - \rho_{fl}}, \tag{5}$$

$$\phi_{N,e} = \phi_N - V_{sh}\phi_{N,sh}, \tag{6}$$

in which, $\rho_{b\cdot c}$ refers to the clay corrected formation bulk density (g/cm^3); ϕ_N represents the neutron porosity (fraction) from the neutron log; and $\phi_{N,sh}$ and ρ_{cl} denote the neutron porosity (fraction) and formation bulk density (g/cm^3) of the shale zone, respectively.

The clay corrected effective porosity (ϕ_e) is determined through the combination of clay corrected neutron and density porosities in the porous rock formation using the following equation:

$$\phi_e = \sqrt{\frac{\phi_{D,e}^2 + \phi_{N,e}^2}{2}}. \tag{7}$$

Although, the non-destructive testing for a rock sample is a more accurate way to predict the sonic primary velocity, it is a time consuming, expensive, and tedious method. The sonic log is employed to obtain the profile of sonic acoustic compressional wave velocity (Vp). The compressional sonic acoustic time (DT, μs/ft) is inversely proportional to the sonic primary wave velocity (Vp, km/s) of the formation, as expressed below (Gaviglio 1989; Anemangely et al. 2017):

$$V_p = \frac{304.8}{DT}. \tag{8}$$

When the shear sonic slowness is not available from the wireline log data, the synthetic shear wave velocity (Vs) can be estimated from the available correlations for various rock formations (Pickett 1963; Caroll 1969; Krishna et al. 1989; Greenberg and Castagna 1992; Vernik et al. 1993; Brocher 2005; Hossain et al. 2012; Bailey and Dutton 2012; Lee 2013; Ojha and Sain 2014; Anemangely et al. 2017). A limited number of studies have attempted to obtain the shear wave velocity for the siliciclastic or shaly sand rock formations (Castagna et al. 1985; Han et al. 1986; Williams 1990; Miller and Stewart 1990; Ramcharitar and Hosein 2016). Recently, Oloruntobi and Butt (2020) have proposed an empirical correlation to estimate V_s using V_p and ρ_b for any type of sedimentary rocks such as shaly sand reservoirs, as given below:

$$V_s = A\left[\frac{V_p}{\sqrt{\rho_b}}\right]^4 + B\left[\frac{V_p}{\sqrt{\rho_b}}\right]^3 + C\left[\frac{V_p}{\sqrt{\rho_b}}\right]^2 + D\left[\frac{V_p}{\sqrt{\rho_b}}\right] + E. \tag{9}$$

In Eq. (9), A (0.094), B (−0.881), C (2.605), D (−1.415), and E (−0.435) are the coefficients of the polynomial expression.

In the literature, there are several empirical correlations to calculate the rock strength as well as elastic properties using sonic acoustic travel time and formation density. Also, elastic properties such as Young's modulus (E), shear modulus (G), bulk modulus (K), and Poisson's ratio (υ) can be estimated using density and sonic logs. The following equations can be utilized to obtain the four main isotropic elastic properties at dynamic conditions using bulk density (ρ_b), sonic acoustic compressional (Vp), and shear wave velocity or V_s (Li and Fjær 2012):

$$\nu = 0.5\left(V_p^2 - 2V_s^2\right)/\left(V_p^2 - V_s^2\right), \tag{10}$$

$$E = \rho_b V_s^2(3V_p^2 - 4V_s^2)/(V_p^2 - V_s^2), \tag{11}$$

$$G = \rho_b\left(V_s^2\right), \tag{12}$$

$$K = \rho_b\left(V_p^2 - \frac{4}{3}V_s^2\right) = \frac{E}{3(1-2\nu)}. \tag{13}$$

Any of the elastic properties (such as E) can be calculated using two remaining properties (e.g., K and ν) (Fjær et al. 2008). The dynamic moduli are always larger than the static ones, since the dynamic strains are always smaller than static strains of rocks (Demirdag et al. 2010; Najibi et al. 2015). The conversion factor from static to dynamic moduli is dependent on the formation porosity, confining pressure, degree of loading, and other lithological factors (Rasouli et al. 2011). In the current study, the UCS parameter is estimated using the following model for the shaly sand formations (Fjær et al. 2008; Rabbani et al. 2012):

$$\mathrm{UCS} = 3.3 \times 10^{-20}\rho_b^2 V_p^4(1-2\nu)\frac{(1+\nu)^2}{(1-\nu)^2}\left(1 + 0.78V_{sh}\right). \tag{14}$$

2.2 Log Data Collection and Quality Assurance

A combination of several wireline logs such as gamma ray, deep resistivity, porosity, and caliper logs is used to identify the lithology, and to estimate the formation thickness, and rock physical and mechanical properties. The gamma-ray (GR) log is a complementary log that is able to identify the formation rock type, and to measure the amount of radioactive elements (API) and clay content. The density log measures the electron density (bulk density, RB) to estimate the formation true porosity, while the neutron log counts the hydrogen concentration (NP) as well as porosity of fluid-filled formation. The sonic log helps to measure the acoustic wave in the form of compressional travel time (DT) or shear slowness. On the other hand, the resistivity log measures the conductivity (inverse of resistivity, RT) in the rock formation. All of these log variables are used in this study to develop the data-driven model; field log data are collected from a shaly sand reservoir located in the Bengal basin. The lithology of the formation studied in this investigation have been characterized in a previous study (Islam 2010). The main lithology is sandstone, and sand-shale alteration. The most common detrital grains are quartz, feldspars, and rock fragments. Quartz is the most abundant constituent. The point count data indicated that quartz, feldspars, and lithic fragments ranges (%) are 65.4–71.1, 12.3–20, and 9.2–18.9, respectively. The texture is very fine to medium grain size, and well sorted. The grain (matrix) density of the sandstone ranges from 2.66–2.78 g/cm^3, while the thin section porosity is about 13%. The available field well log data (such as GR, RT, RB, and DT) and neutron-density porosity are utilized to predict the in situ rock strength profile in the shaly sand rock formations. A total of 182 data samples are used for each log variable in the study. The log data quality is also confirmed to ensure the reliability of each log dataset variable by checking the depth shift and borehole conditions. The output and input variables are classified into training, testing, and validation phases with 75%, 15%, and 10%, respectively. All the programming tasks related to this study are carried out using Matlab (R2017b) programming environment.

2.3 Development of Connectionist Predictive Models

ANN model: The ANN is composed of different components such as artificial neurons (connected in each layer), weight factors, bias term, and transfer function in a connectionist process unit system. The ANN can be adopted with a single-layer perception, multilayer perception (MLP), and/or radial basis function networks (Ali 1994; Mohaghegh et al. 1996; Razavi and Tolson 2011). The MLP is one of the conceptually attractive feedforward neural network approach employed in estimation of rock formation properties, rock mechanics, stability of underground openings, wellbore failure analysis, and rock engineering (Yang and Zhang 1997; Meulenkamp and Grima 1999; Helle and Bhatt 2002; Yilmaz and Yuksek 2009; Al-Bulushi et al. 2009; Ocak and Seke 2012; Khandelwal and Monjezi 2013). The deep learning MLP model consists of at least four layers with one input layer, at least two hidden layers, and an output layer for target variables. The training algorithm and activation (transfer) functions are the most important components of an ANN structure. In this study, the ANN model is processed with the Levenberg–Marquardt (LM) algorithm for training function (TRAINLM) to adjust the weight factors through simulation in the connectionist model. The LM algorithm is usually faster and more reliable in the backpropagation system for ANN model than other standard backpropagation methods

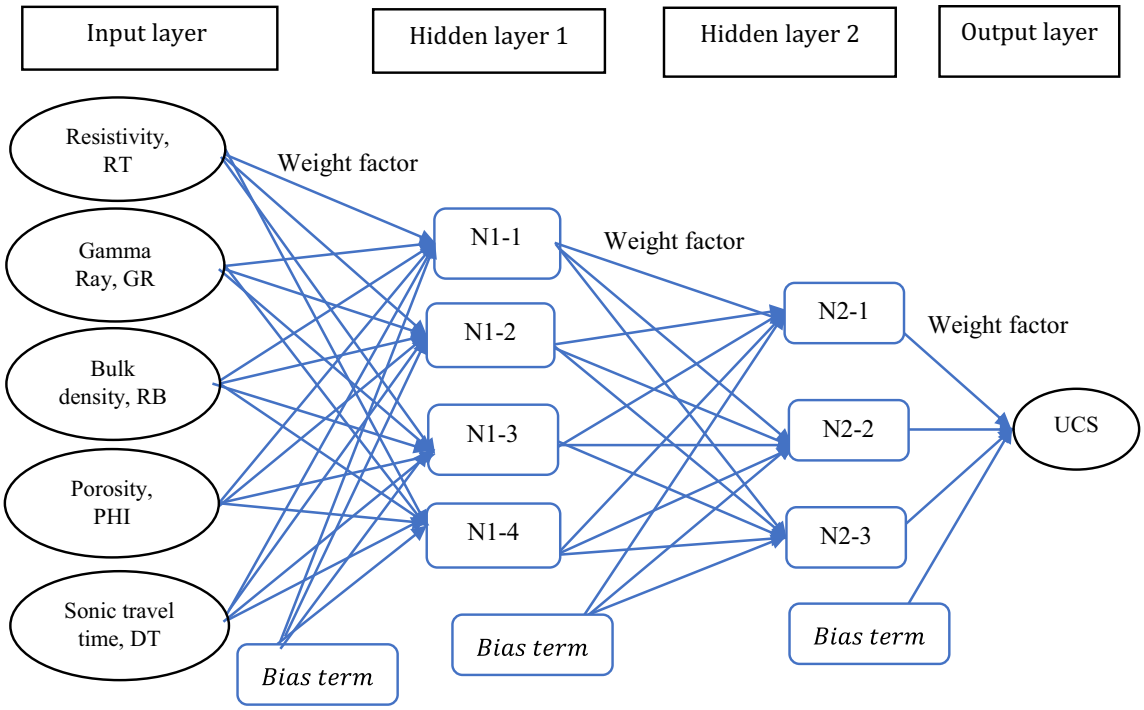


Fig. 1 Schematic of ANN architecture employed in the current work

(Ceryan et al. 2013). The Tansig-type activation function is used between input and hidden layers while Purelin (linear) transfer function is employed between the last hidden layer and output layer in the model. Taheri-Garavand et al. (2015) claimed that a hidden layer with a lower number of neurons is desired in an ANN model because of shrinking the neural network structure as well as increasing the learning potential. In the current study, the number of hidden layers is chosen by a trial and error strategy to optimize the ANN structure such that the maximum regression coefficient and minimum mean squared error (MSE) as the selection criteria are attained to find the best configuration in the smart approach. The number of neurons in the hidden layer is optimized to enhance the network performance as well as to save the computational training time (Zendehboudi et al. 2014). An ANN structure is shown in Fig. 1. A flowchart is also depicted in Fig. 2 to present the ANN model development steps and the optimization strategy in the study. In Fig. 1, N1 and N2 refer to the number of neurons in the hidden layers 1 and 2, respectively.

The activation function is Purlin for the output neuron. The generalized mathematical expression for MLP-based neural network with Tansig transfer function for the hidden neurons is given as follows:

$$y_m = \left[\sum_{j=1}^{m} \omega_{jm} \text{Tansig} \times \left(\sum_{i=1}^{n} \omega_{ij} x_i + b_j\right) + b_m\right]. \quad (15)$$

In Eq. (15), y_m stands for the output variables; x_i introduces the vector of target variables ($i = 1, 2, 3, 4, 5\ldots$); b_m resembles the bias term for output layers; ω_{ij} denotes the connection weight on the link from i to j node between input and hidden layers; m is the number of hidden nodes; and n refers to the number of input variables.

The common advantages of ANN technique are: (a) it generally provides a high prediction accuracy where it captures better the nonlinear and high dimensional relationships among data points, compared to the statistical regression models, (b) it can effectively model randomly changing variables with a non-constant variance, and (c) it offers a more convenient and user-friendly modeling approach, compared to analytical and numerical modeling strategies (Esene et al. 2020). However, the ANN tool has some inherent limitations such as slow convergence speed, low generalization performance, reaching local optima, and overfitting problem.

LS-SVM model: The SVM is one of the powerful artificial intelligence approaches used in data classification and regression analysis. The SVM uses a subset of training points in the support vectors (decision function), which was first introduced by Vapnik (1995). The LS-SVM is a modified version of the classic SVM algorithm proposed by Suykens and Vandewalle (1999). This modified version is less complex than the classic SVM algorithm. It helps to reach the solution of a worsening problem (with less data points) more efficiently by setting up a linear set of equations through employing SVM instead of the quadratic

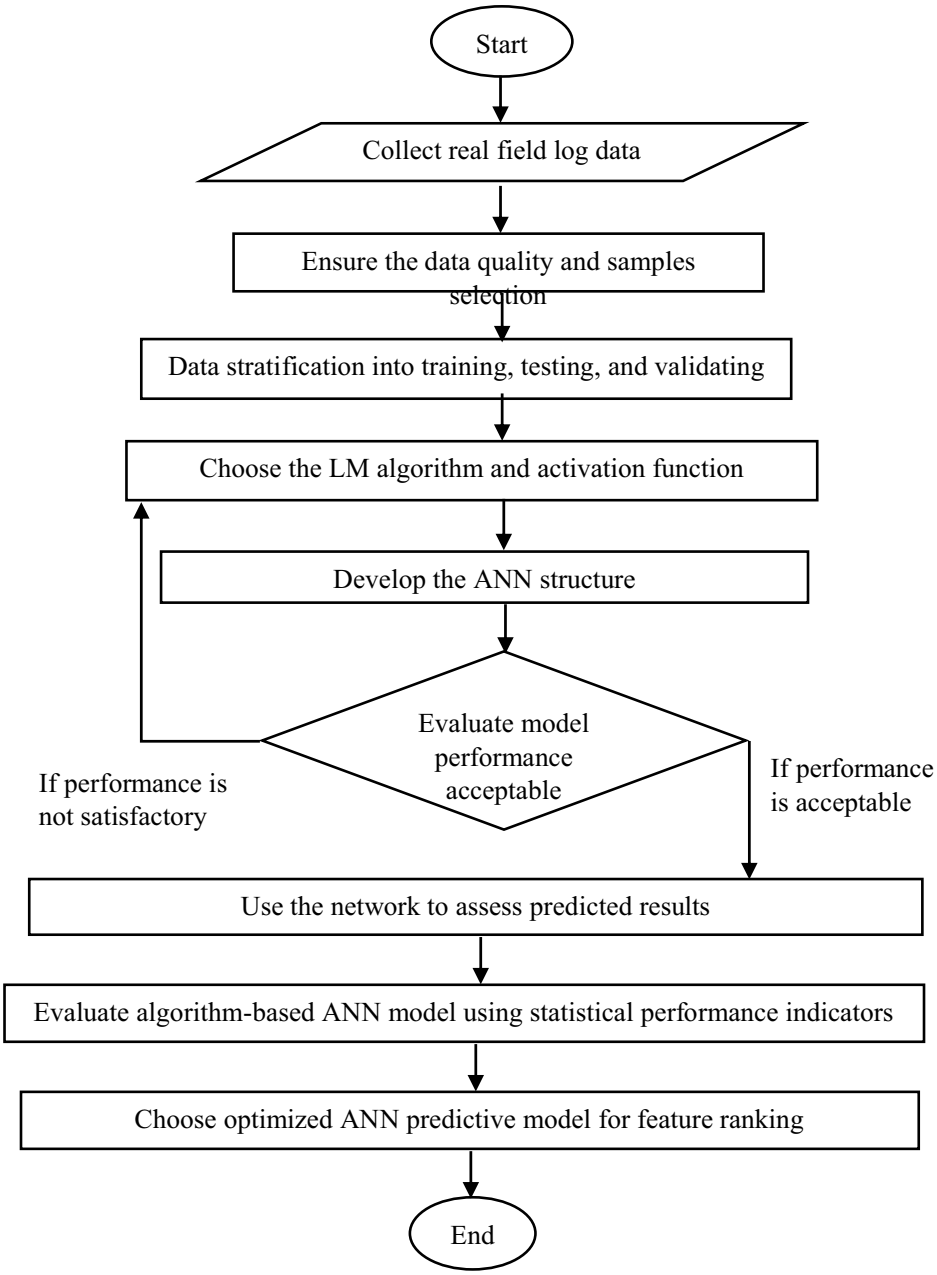


Fig. 2 A flowchart for ANN model development and prediction of the output variable

programming (Suykens and Vandewalle 1999). Compared to SVM, the LS-SVM learning method is less time consuming. Unlike the ANN tool, the LS-SVM can be used with a limited number of data points, while it has a higher generalization and training efficiency (Ghiasi et al. 2014; Kamari et al. 2014; Nejatian et al. 2014; Zendehboudi et al. 2018). However, the LS-SVM is very sensitive to outliers as well as the magnitudes of kernel and tuning parameters. Also, there is a lack of sparsity, which might limit its utilization for large-scale problems. More information regarding the theory and algorithms with different features of SVM or LS-SVM can be found in the literature (Smola et al. 2004; Sebtosheikh et al. 2015; Esfahani et al. 2015; Miah et al. 2020). The following equation is used in the LS-SVM:

$$y(x_i) = \omega^T \varphi(x_i) + b \quad \text{where} \quad x_i \in R^n \text{ and } y_i \in R, \tag{16}$$

where the nonlinear function $\varphi(.)R^n \rightarrow R^{n_i}$ represents the primal space to a feature space with higher dimensions. The dimension n_i of this space is only defined in an implicit way; b is a bias term; and $\omega \in R^{n_i}$ introduces the weight factor. The optimization problem can be written for function estimation in the LS-SVM (Suykens et al. 2002; Esfahani et al. 2015), as shown below:

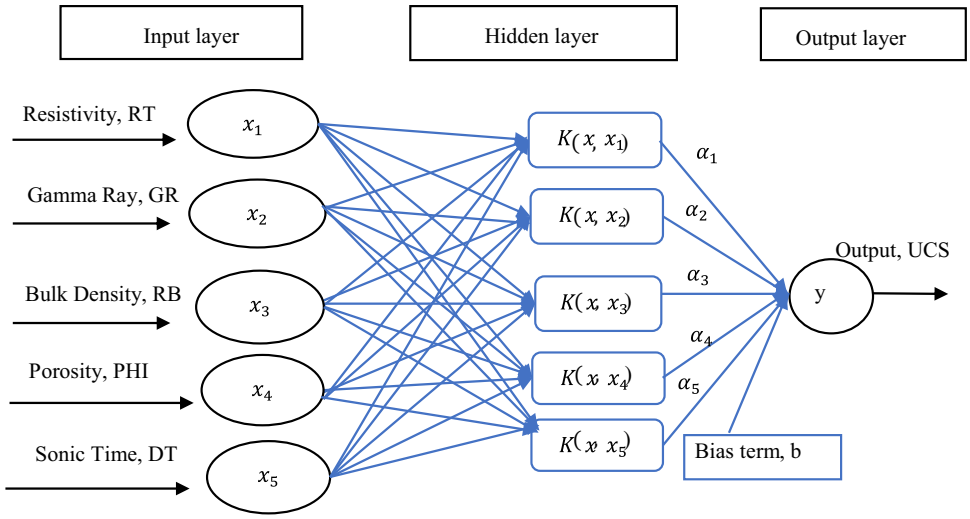


Fig. 3 A generalized structure of LS-SVM proposed in the study

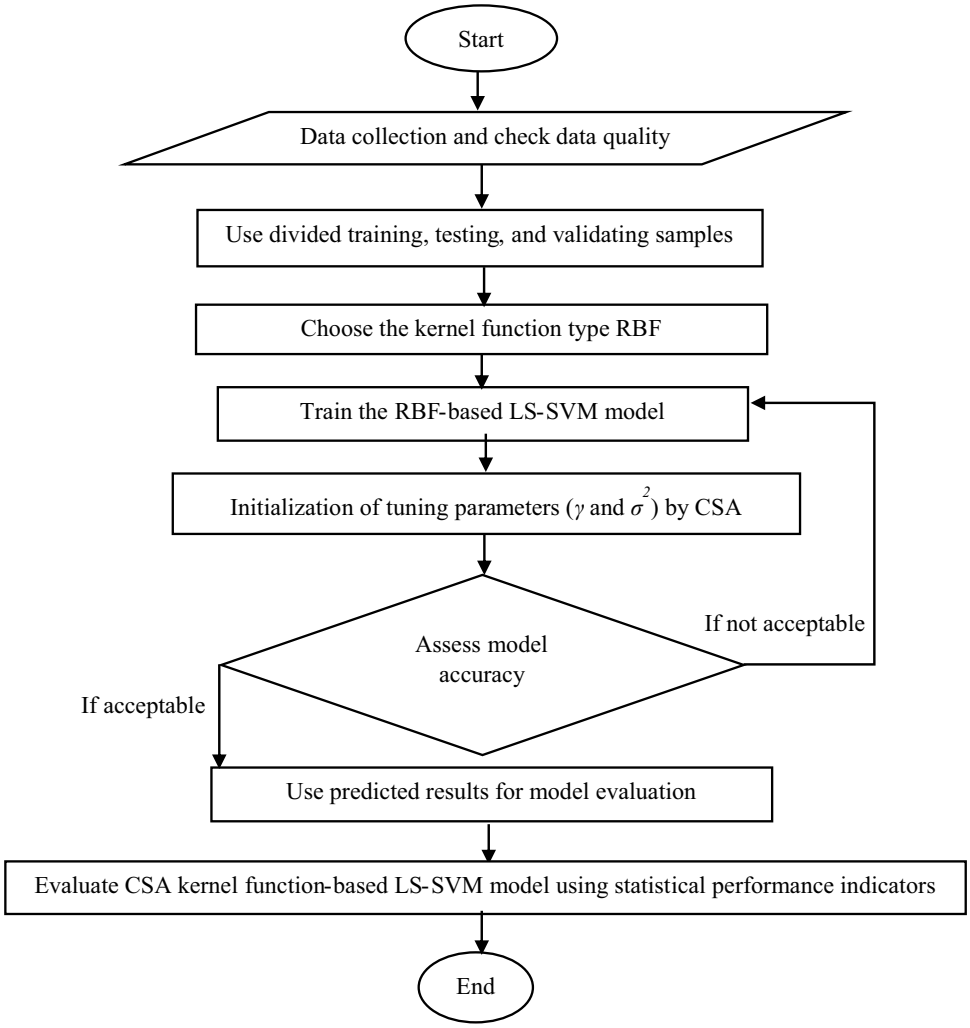


Fig. 4 A flowchart for kernel-based LS-SVM model development

$$\text{Minimize} \quad J(\omega, e_i) \cong \frac{1}{2}\omega^T\omega + \gamma \sum_{i=1}^{n} e_i^2. \tag{17}$$

Subject to the following constraint:

$$y_i = \omega^T \varphi(x_i) + b + e_i, i = 1, 2, 3, \ldots, n. \tag{18}$$

In Eq. (18), e_i represents the error variable, and γ resembles the regularization (annealing) parameter to prevent overfitting.

The kernel function is $K(x, x_i) = \varphi(x_i)^{\mathrm{T}} \varphi(x_i)$, which needs to be satisfied with Mercer's condition. After eliminating ω and e_i, the final expression can be formulated for the LS-SVM function estimation as follows:

$$y(x_i) = \sum_{i}^{n} \alpha_i K(x, x_i) + b, \tag{19}$$

where b and α are the solutions to the linear system expressed through Eq. (19). The α (weight factor) is a vector with the size of $n \times 1$; x is the training sample; and x_i refers to the support vector.

The kernel function-based LS-SVM structure used in this study is illustrated in Fig. 3. A summarized methodology for development of kernel-based LS-SVM model is shown in Fig. 4.

There are many kernel functions that can be used in LS-SVM such as linear, polynomial, spinal, radial basis, and sigmoidal (Zendehboudi et al. 2018). Among all kernel functions, the Gaussian radial basis kernel function (RBF) is mostly used in the LS-SVM learning strategy to find the best output (Suykens et al. 2002; Samui 2008; Zendehboudi et al. 2018) due to its computational simplicity and other features (e.g., capable of solving nonlinear problems). The kernel function-based LS-SVM structure used in this study is shown in Fig. 3. The Gaussian RBF can be defined mathematically as follows (Samui 2008):

$$K(x, x_i) = \exp\left(-\frac{\left(||x_i - x||\right)^2}{2\sigma^2}\right). \tag{20}$$

In the kernel function of RBF-based LS-SVM, the regularization and kernel parameters (also named as tuning parameters) γ and σ^2 are adjusted through a global optimization technique, namely, the coupled simulated annealing (CSA) (Xavier-de-Souza et al. 2009). The CSA optimization process is proven to be more effective than multi-start gradient descent technique (Suykens et al. 2002; Rostami et al. 2019). Similar to the MLP-ANN model, the databank for the log data is divided randomly into three sub-datasets to construct the LS-SVM models using different kernel functions. The total samples are categorized into three groups including 75% for training, 15% for testing, and 10% for validation in the LS-SVM connectionist model with the CSA optimization approach.

2.4 Model Performance Assessment

Four statistical indicators are used in this study to analyze the predictive model performance. The indicators are root mean square error (RMSE), coefficient of determination (R^2), average absolute percentage relative error (AAPE), and maximum average absolute percentage error (MAPE). The mathematical expressions for all performance indices are listed below (Zendehboudi et al. 2014, 2018):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_{\mathrm{m},i} - Y_{\mathrm{p},i}\right)^2}, \tag{21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_{\mathrm{m},i} - Y_{\mathrm{p},i}\right)^2}{\sum_{i=1}^{n}\left(Y_{\mathrm{m},i} - Y_{\mathrm{m,mean}}\right)^2}, \tag{22}$$

$$\text{AAPE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\left(Y_{\mathrm{m},i} - Y_{\mathrm{p},i}\right)}{Y_{\mathrm{m},i}}\right) \times 100, \tag{23}$$

$$\text{MAPE} = \max .\left|\frac{\left(Y_{\mathrm{m}} - Y_{\mathrm{p}}\right)}{Y_{\mathrm{m}}}\right| \times 100. \tag{24}$$

In the preceding equations, n indicates the total number of samples; Y_{m} resembles the measured variable; $Y_{\mathrm{m,mean}}$ is the mean value of Y_{m}; and Y_{p} represents the predicted output variable. The accuracies of the data-driven models are analyzed on the basis of the low or high value of statistical indices. In the study, the model is best for the high value of R^2 (close to 1) and low values of RMSE, AAPE, and MAPE.

2.5 Sensitivity Analysis and Variable Selection

In this study, a systematic strategy is employed to perform the parametric sensitivity analysis as well as to find the relative importance of the input log variables in the AI-based predictive rock strength models (see Fig. 5). Furthermore, the optimized AI models predict the UCS; the model performance is assessed using the statistical criteria. Based on the contribution of an input variable to the predictive models, the input parameters are ranked. The variable ranking approach is called 'single variable elimination' of the data-driven model, while it has selective multiple input variables to predict the output variable using the optimized AI structure. In the log parameter ranking through the data-driven model, if the model results in high AAPE, MAPE, and RMSE, and low R^2, it implies that the eliminated variable

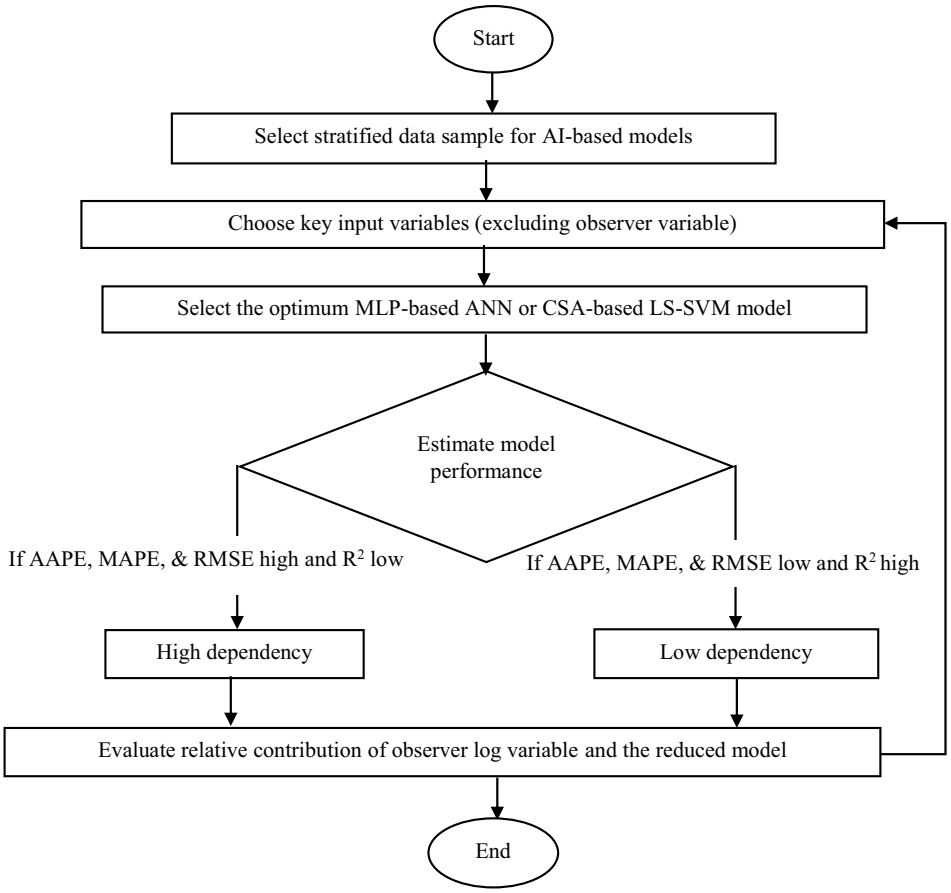


Fig. 5 Flowchart to conduct parameter sensitivity analysis and to rank the input parameters

Table 3 Summary of the statistical values of the used log data

Log parameters	Max	Min	Mean	St. Dev	Sample var
GR (API)	157.82	76.28	100.19	13.87	192.51
RT (Ω m)	39.70	13.70	22.67	4.96	24.58
RB (g/cm^3)	2.53	2.30	2.37	0.0425	0.0018
NPHI (v/v)	0.2039	0.1455	0.17	0.013	0.0002
DT (μs/ft)	97.40	85.84	92.90	2.43	5.89

Table 4 Summary of log-based formation properties magnitudes in the study

Formation properties	Maximum	Minimum	Mean
Compressional wave, Vp (km/s)	3.55	3.13	3.28
Shear wave, Vs (km/s)	1.59	1.51	1.54
Poison's ratio, ν (fraction)	0.31	0.25	0.29
Porosity, PHI (v/v))	0.18	0.06	0.15
Rock strength, UCS (MPa)	78.82	24.40	33.19

has high impact on the model. It is worth noting that only most influential parameters are used to develop the new correlations for obtaining continuous in situ UCS profile for the clastic sedimentary reservoir rocks through multivariable regression analysis with real field application.

3 Results and Discussions

3.1 Data Analysis

The radioactivity properties including gamma ray, deep resistivity, formation bulk density, and sonic log responses considerably change with depth throughout a shaly sand formation. For the data set under study, the minimum and maximum values of gamma ray over the entire lithology log of the formation are 77 and 155 API, respectively, which are used to calculate the shale volume. The average shale volume extent is 11.84%, while the minimum value is 0.26% in the studied depth interval. The statistical information on the log data samples and the estimated values of rock characteristics are presented in Tables 3 and 4 over the entire depth of interest reservoir rock zone, respectively. Note that the formation properties in Table 4 are derived from the statistical

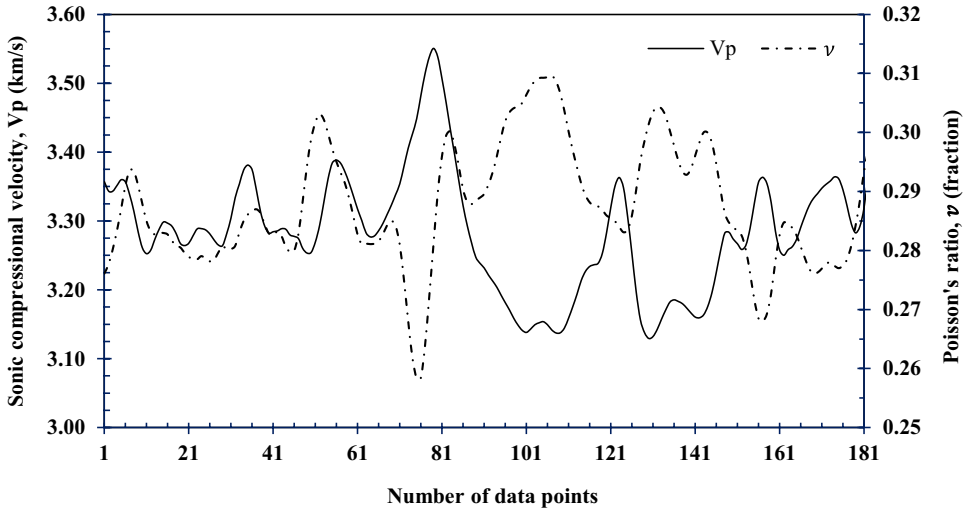


Fig. 6 Profile of compressional wave and Poison's ratio in the rock formation

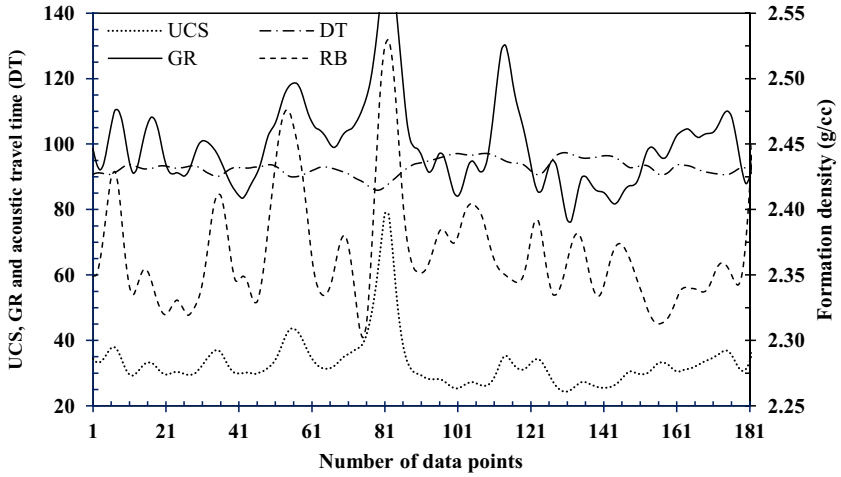


Fig. 7 Variation of rock strength with formation gamma ray, sonic travel time, and bulk density, depending on the number of data points

Table 5 Correlation matrix between rock strength and formation characteristics

	UCS	RT	PHI	RB	DT	GR
UCS	1					
RT	−0.44106	1				
PHI	−0.84662	0.473699	1			
RB	0.664292	−0.59155	−0.84257	1		
DT	−0.77644	0.311042	0.464583	−0.29861	1	
GR	0.852674	−0.37541	−0.8225	0.523704	−0.58251	1

values in Table 3 by following the methodology outlined in Sect. 2.1.

The clay corrected effective porosity (PHI) is estimated using neutron and density porosity; it is then used to obtain the in situ UCS of shaly sand formation using an empirical correlation presented in Fjær et al. (2008). The clay corrected porosity of a siliciclastic rock changes with the heterogeneity of pore diameter, size and shape of pores, and formation density along the direction of vertical depth. Also, the primary acoustic velocity (inverse of sonic compressional travel time) and in situ Poisson's ratio vary with respect to the vertical depth (6955.95–7015.33 ft) due to the rock formation compaction, cementation, and heterogeneity, as shown in Fig. 6. The in situ rock strength, UCS, varies vertically due to the overall effect of rock heterogeneity, formation radioactive mineral depositions, compaction, pore structure, grain size, packing, and rock density, as depicted in Fig. 7. The rock strength of a shaly sand formation increases with an increase in gamma ray and bulk density, while it is inversely proportional to the

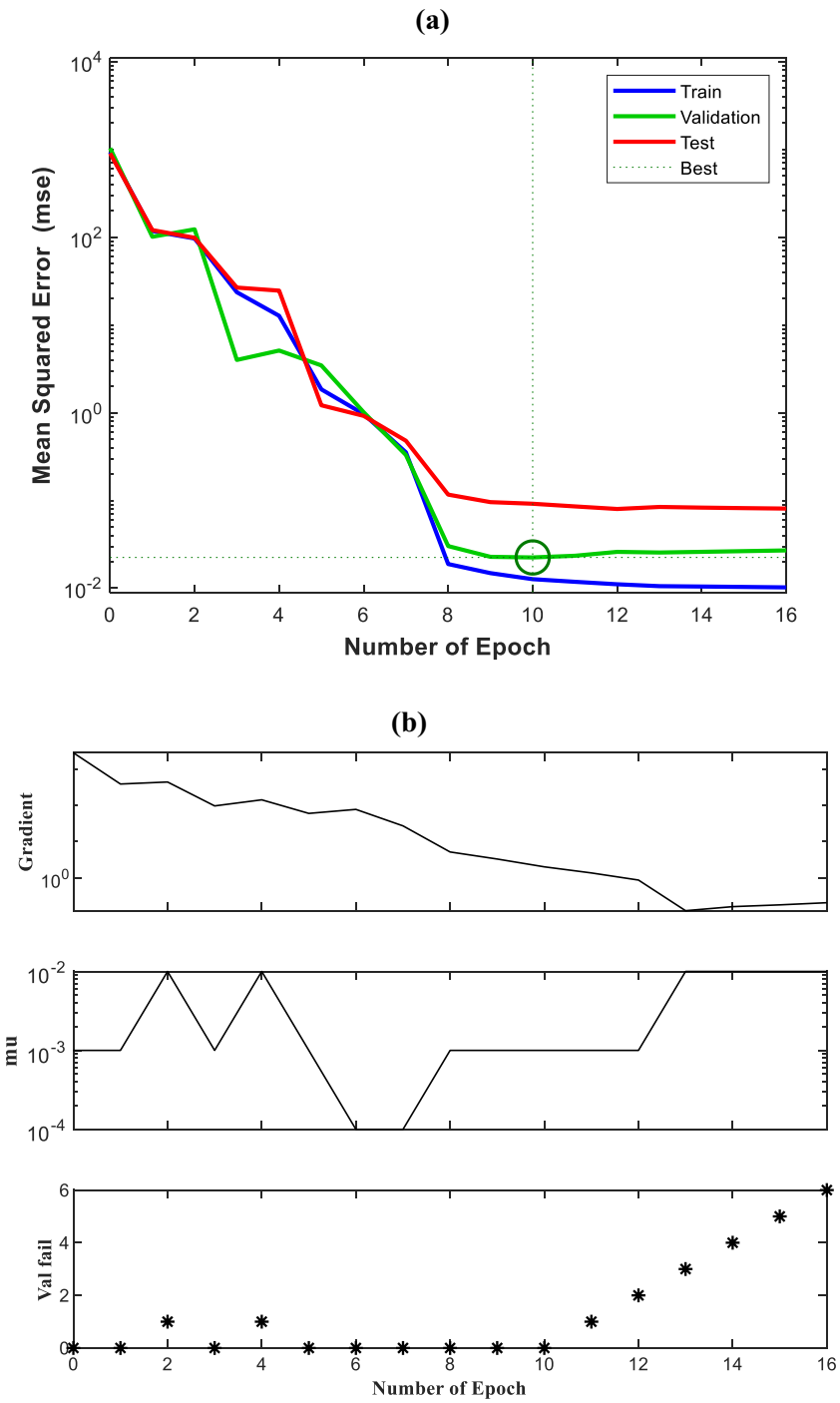


Fig. 8 Graphical representation of **a** validation performance, and **b** training phase for optimized ANN model

acoustic travel time as well as rock resistivity and porosity. The correlation matrix between UCS and other formation properties (e.g., porosity and true resistivity) is shown in Table 5.

3.2 Data-Driven Model Performance

A feedforward ANN is designed using stratified data samples. The optimum structure is obtained with the algorithm of LM; the optimal model has one input layer with five neurons, the first hidden layer with four neurons, the second hidden layer with three neurons, and one output layer with one neuron. The validation performance versus the number of epochs is illustrated in Fig. 8a for the constructed MLP-based ANN model. Figure 8b also shows the performance of training phase in terms of number of epochs. In this study, the validation phase is conducted to tune the model parameters and terminate the neural network training stage before

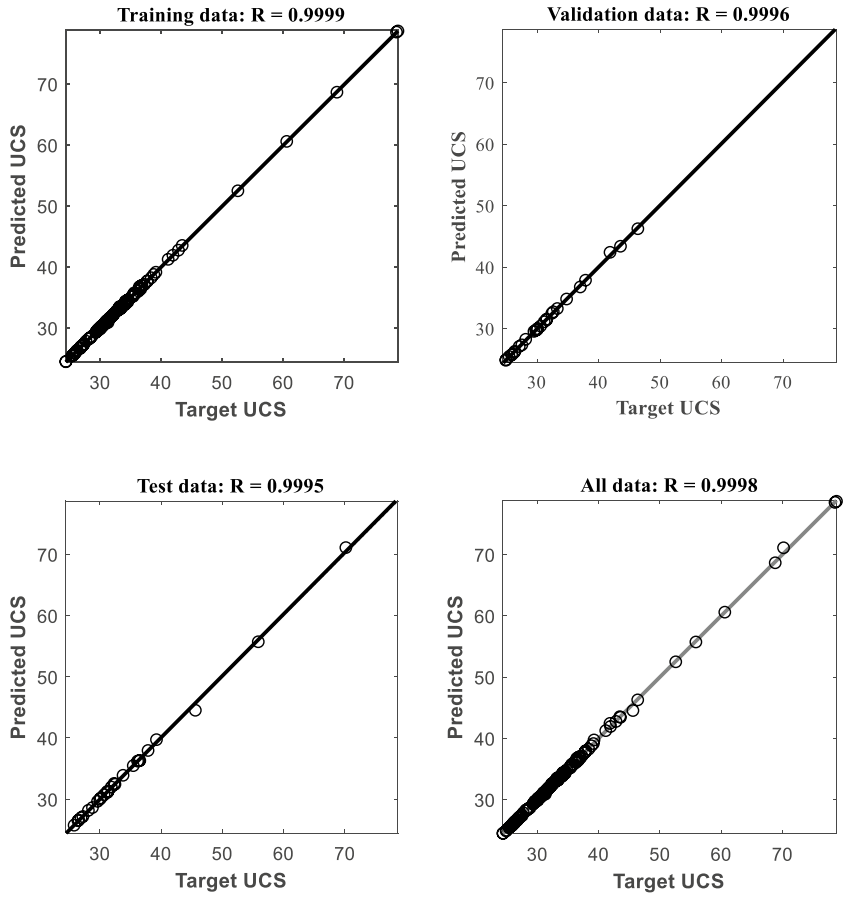


Fig. 9 Predictive performance of the optimized ANN model

overfitting; μ (mu) is the Marquardt parameter in the training step of the network. The best validation performance is found at the epoch number of 10 with a mean square error of 0.02243. The magnitudes of the gradient and mu are 0.02034 and 0.01 at epoch 16, respectively, in the training stage. The correlation coefficient is close to one for both training and testing stages in the optimized ANN model (see Fig. 9). It follows that the predicted results are in good agreement with the actual results for all training, testing, and validation data sets. The MLP-based ANN gives less error (percentage) with an MAPE of 2.3692% and 1.1956% for the training and testing phases, respectively.

In addition, the CSA optimization technique is used in the LS-SVM model as an iterative random search strategy. The optimization procedure is repeated several times to reach the optimum global point. Figure 10 displays the scatter plots (target values versus predictions) of the training, testing, and validation steps for the optimized LS-SVM model. The initial values of the annealing and tuning parameters (γ and σ^2) are 505,962.56 and 104.38, respectively. The final tuning parameters γ and σ^2 of the radial kernel function (Gaussian) based log data-driven model for the UCS are 3.62×10^{11} and 1.74×10^{3}, after 14 iterations. The RBF-based model results in a value of MAPE (%) equal to 0.0476 for the training and 0.2656 for the testing, respectively; the correlation coefficient is close to 1. The RBF-based LS-SVM predictive model has a greater performance in terms of accuracy and reliability; it leads to the lowest RMSE, AAPE and MAPE, and high R^2. The statistical information for both MLP-based ANN and kernel function-based LS-SVM models is summarized in Table 6.

3.3 Parametric Sensitivity Analysis and Variable Selection

The optimized ANN structure in this study has two hidden layers with the topography of (4–4–3–1). The selected ANN model is used to perform parametric sensitivity analysis. Conducting the statistical analysis, the log parameters are ranked based on their relative importance in the predictive UCS model. To further understand the effect of individual variables, one variable is excluded at each time and the remaining four variables are used in both ANN and LS-SVM models. In the model schemes A through E, the UCS predictions are obtained in the absence of one input variable. Due to the less significance of RT and PHI (input parameters excluded from the predictive models), models D and E show better performance (such as lower error and greater

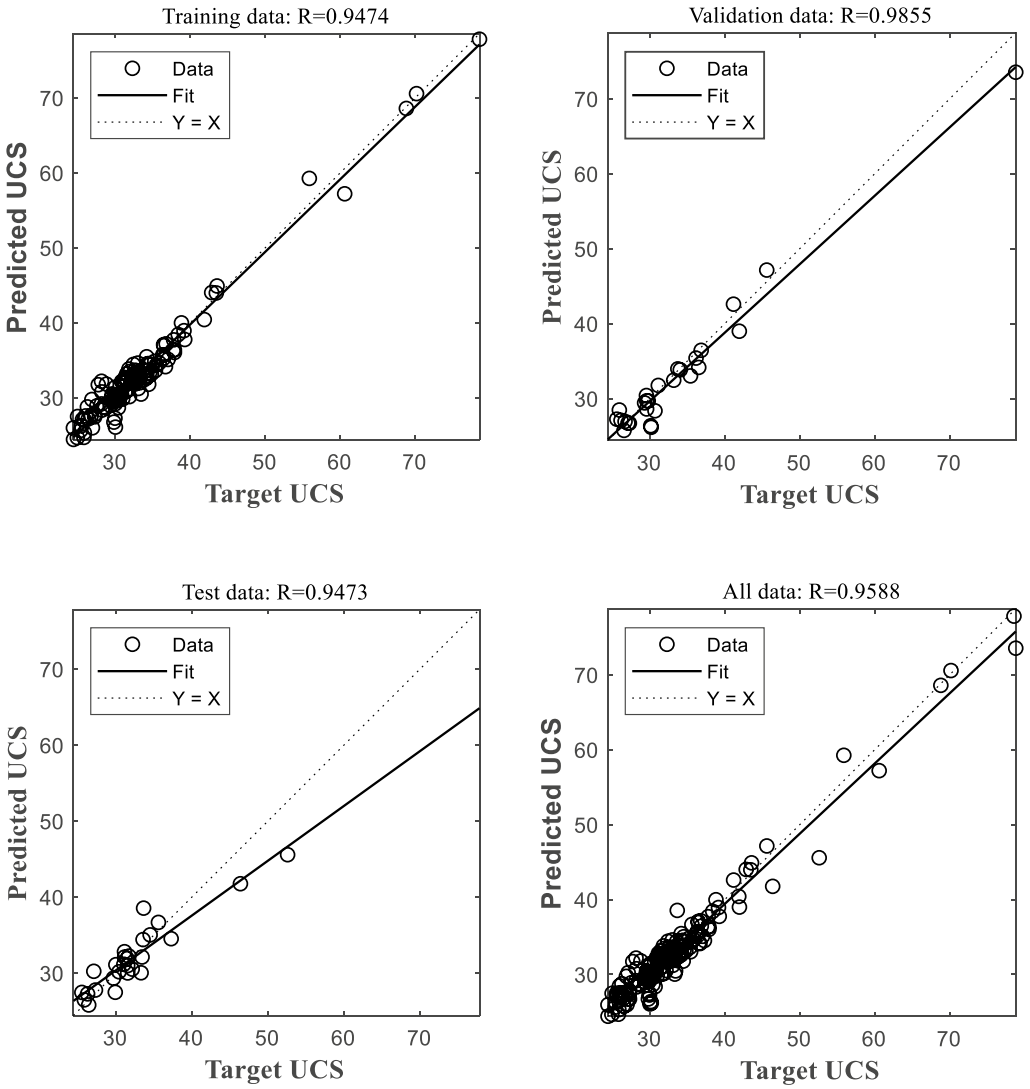


Fig. 10 Performance of model A at different stages in the absence of DT (observer) variable

Table 6 Comparison of data-driven predictive models performances in the study

AI approach	Prediction phase	AAPE	MAPE	RMSE
MLP-ANN	Training	0.2803	2.3692	0.1735
	Testing	0.2779	1.1956	0.1314
LS-SVM	Training	0.0118	0.0476	0.0048
	Testing	0.0306	0.2656	0.0199

correlation coefficient), compared with the other models. Model A leads to a weaker performance with a low value of the correlation coefficient and high magnitudes of statistical errors due to the absence of an important log variable, DT, in the model (Fig. 10).

Tables 7 and 8 present the performance metrics for the developed models. The statistical information of both connectionist techniques is presented in Fig. 11.

According to the statistical analysis conducted on the LM-based ANN approach, the significant input variables are the sonic travel time, gamma ray, formation bulk density, porosity, and true resistivity (high to low ranked order) for prediction of UCS. Also, the same order of variable importance is attained from the RBF-based LS-SVM model. Based on the literature, the sonic porosity log (e.g., sonic travel time) is needed to estimate the rock strength using lithology-based correlations. The current study demonstrates that acoustic compressional travel time (DT) parameter has the highest contribution to the UCS value in a shaly sand formation. The rock resistivity and porosity have minor significance while predicting in situ UCS profile. According to the different testing, validation, and generalization approaches

Table 7 Investigating the impact of excluded variable on the performance of the MLP-based ANN msodel with four layers

Model scheme	Predictor variables	Excluded variable	Phase	AAPE	RMSE	R^2
A	RT, GR, RB, PHI	DT	Training	3.60	1.60	0.947
			Testing	4.84	2.26	0.947
B	RT, RB, PHI, DT	GR	Training	1.85	0.96	0.983
			Testing	1.90	0.84	0.991
C	RT, GR, PHI, DT	RB	Training	0.63	0.29	0.999
			Testing	0.53	0.23	0.999
D	GR, RB, PHI, DT	RT	Training	0.18	0.10	0.999
			Testing	0.16	0.07	0.999
E	RT, GR, RB, DT	PHI	Training	0.005	0.003	1
			Testing	0.004	0.004	1

Table 8 Importance of input variables in the RBF-based LS-SVM model

Model scheme	Predictor variables	Observer variable	Phase	AAPE	RMSE	R^2
A	RT, GR, RB, PHI	DT	Training	2.85	1.20	0.973
			Testing	4.41	2.78	0.938
B	RT, RB, PHI, DT	GR	Training	1.08	0.61	0.994
			Testing	1.42	0.69	0.952
C	RT, GR, PHI, DT	RB	Training	0.44	0.19	0.999
			Testing	1.71	2.38	0.956
D	GR, RB, PHI, DT	RT	Training	0.01	0.002	0.999
			Testing	0.02	0.018	1
E	RT, GR, RB, DT	PHI	Training	0.005	0.003	1
			Testing	0.004	0.004	1

used in the current study, the most significant predictor variables are DT, GR, and RB to estimate the rock strength of siliciclastic rock formations. Also, these variables are vital in capturing dense minerals effect, clay content, acoustic velocity, and bulk density of rocks.

3.4 Development of New Correlations for Rock Strength Estimation

Several correlations have been developed to estimate rock strength (UCS) using either core or wireline log properties data. Thus, it seems that a new log-based UCS correlation is required to obtain in situ rock strength profile that can capture the mineralogical or clay effect, dynamic formation acoustic travel time, and bulk density. A set of 175 in situ data samples of a shaly sand reservoir in the Bengal basin is used in this study to figure out the influences of formation properties on the rock strength as well as to develop a new correlation. The relationships between the rock strength and formation bulk density, gamma ray, and acoustic travel time are illustrated in Figs. 12,13,14. The bulk density is highly heterogeneous with respect to depth in the studied formation. The in situ rock strength increases with an increase in the formation bulk density due to the higher number density of electrons. The rock strength also increases with increasing gamma ray due to the radioactive minerals and clay content in the shaly sand rock. The in situ rock strength decreases with increasing acoustic travel time or decreasing compressional wave in the clastic rock formation.

To capture the lithology indicators, degree of formation density and touchstone of compaction, the formation gamma ray (GR), bulk density (RB), and sonic transit time (DT) are considered to develop the new correlations for predicting the UCS of clastic rocks such as clean or shaly sand rock formation. The following two UCS models are proposed through statistical regression analysis:

$$\text{Model 1: UCS} = 68.158 + 31.347\text{RB} + 0.156\text{GR} - 1.349\text{DT}, \tag{25}$$

$$\text{Model 2: UCS} = 0.149\,\text{GR}^{0.93} + 6.67 \times 10^{10}\text{DT}^{-5} + 0.75\text{RB}^{3.07}. \tag{26}$$

Both models are obtained through multivariate regression analysis where the field data are used considering interrelationships between the parameters. Various forms of equations are examined and the best ones are selected based on

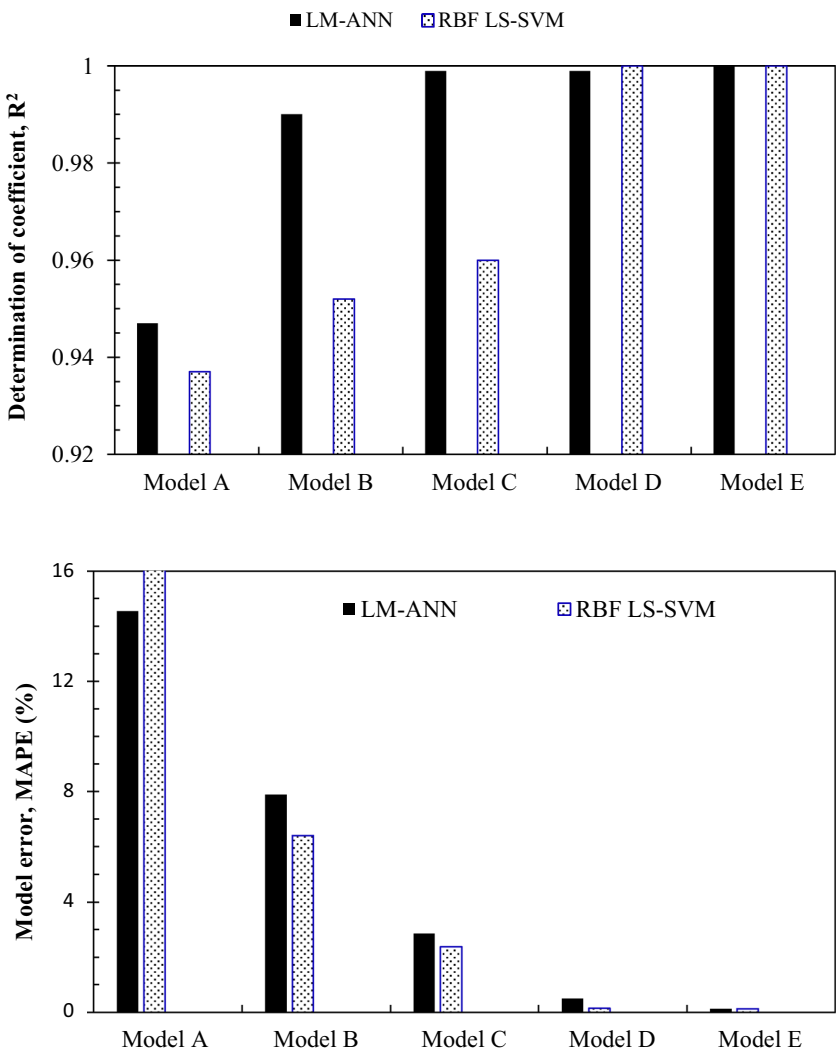


Fig. 11 Comparison of **a** R (top) and **b** statistical errors (bottom) for various model schemes in the absence of an observer variable

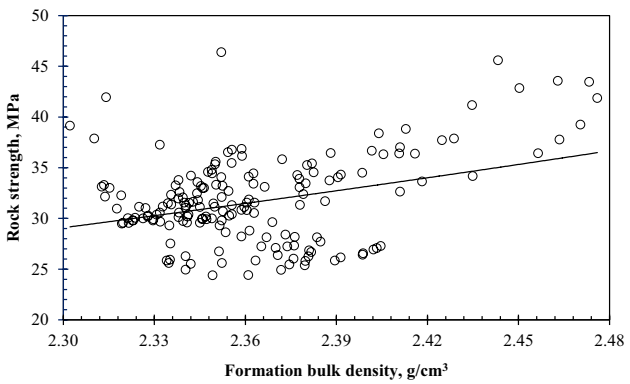


Fig. 12 Relationship between the in situ rock strength and bulk density

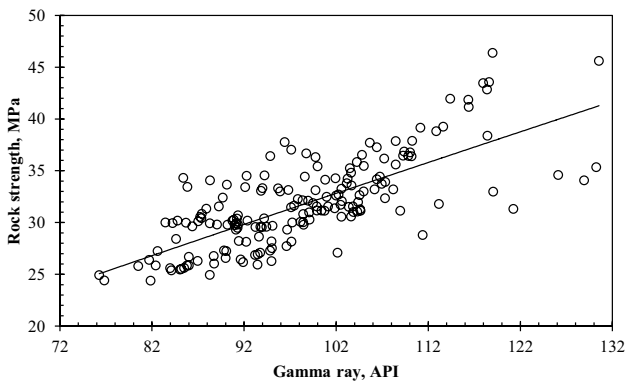


Fig. 13 Relationship between the in situ rock strength and gamma ray

the magnitudes of statistical parameters (high R^2 and low relative error).

The performance of the proposed models is compared with that of other models in Fig. 15. To develop the above models, 175 field log data samples are collected from the shaly sand clastic sedimentary reservoir rock, while it has an average UCS of 31.85 MPa, GR (API) of 98.39, *RB* of 2.36 g/cm^3, and *DT* of 93.12 μs/ft in the depth interval of 6955.5–7015.34 ft. The proposed models have a lower

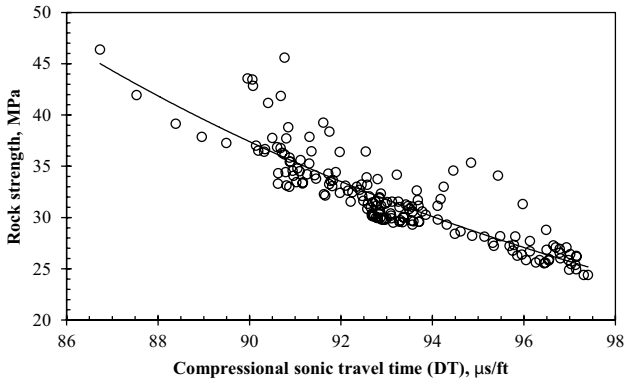


Fig. 14 In situ rock strength versus acoustic travel time

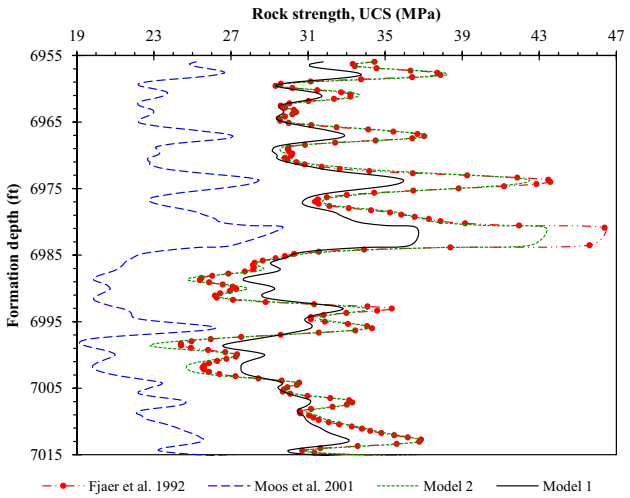


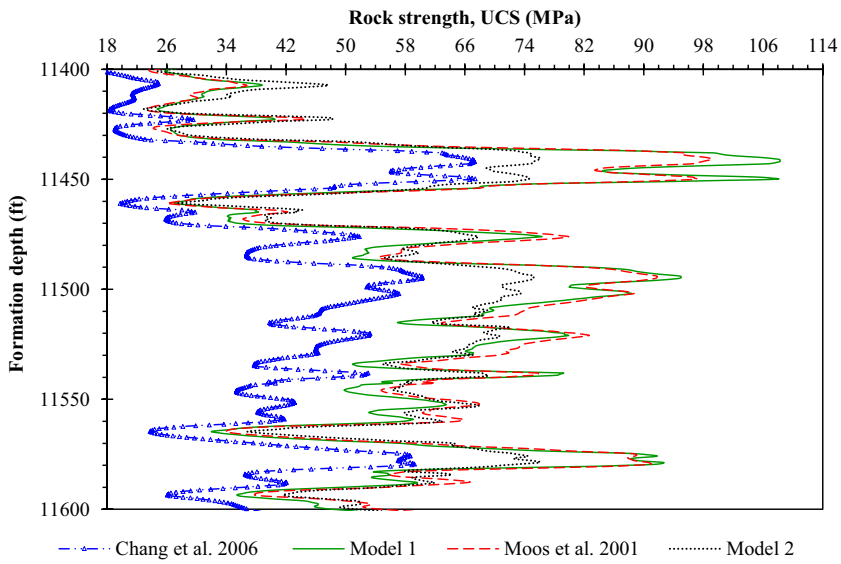
Fig. 15 A comparison of UCS profile obtained from different models

prediction error (a RMSE of 2.593 and an MAPE of 21%), compared to Moos et al. (2001) model (a RMSE of 8.92 and MAPE of 40.63%), while determining the UCS.

To further validate the robustness of the new correlations for UCS, real data taken from Volve field in the Norwegian North Sea are also used. The detailed geological information about the field can be found in the literature (Brekke et al. 2001; Faleide et al. 2010). Due to the availability of well log data in the open sources, Well no. 15/9-F-1-A is chosen to examine the validity of the proposed correlations. The most common lithologies of this well are carbonate rocks (limestone, 9091–10,538 ft), marlstones with the trace of limestone and sandstone (10,538–11,017 ft), claystone (11,017–11,250 ft), and predominantly sandstones with some claystone in the depth interval of 11,250–11,948 ft. The field log data of GR, RB, and DT are employed to estimate continuous UCS profile using proposed models and other correlations, namely Chang et al. (2006) and Moos et al. (2001) where the formation depth is in the range 11,250–11,850 ft of Volve field (Equinor, 2018). A total of 1823 log samples are used from the selected depth of the formation with average values of 75.50 µs/ft, 130.52 API, and 2.48 g/cm^3 for DT, GR, and RB, respectively. The predicted profile of UCS at 200 ft is depicted in Fig. 16.

Based on the UCS profile shown in Fig. 16, Model 2 predictions are closer to the results of Moos et al. (2001)'s model; while Model 1 gives slightly lower values than the outputs obtained from Moos et al. (2001)'s model. Chang et al. (2006)'s model underestimates the in situ rock strength. Using the North Sea fields' log data, Chang et al. (2006)'s model exhibits higher error (an AAPE of 32.52%), compared to Moos et al. (2001)'s model, and Models 1 and 2 introduced in this study.

Fig. 16 A comparison of UCS profile obtained from different models in the Volve field



The trial and error approach for finding proper correlations and theoretical concepts/relevant background reveal that Model 1 in the form of a linear regressive equation offers accurate results for shaly sand formations, while Model 2 that presents a power law equation is a reliable and precise correlation to determine in situ UCS for both clean and shaly sand formations using dynamic log parameters. Chang et al. (2006)'s model does not include the lithology feature (e.g., gamma-ray log) as well as formation density to capture the clay content and formation electron density effect in the model. The drilling engineers and/or geomechanists can employ Model 1 to obtain more reliable in situ UCS profile for the shaly sand rock formations, compared to the model proposed by Moos et al. (2001).

Static formation properties can be obtained using surface measurements or laboratory tests. The static UCS can be measured with different laboratory tests, including unconfined and confined (triaxial) compressive tests. On the other hand, the dynamic rock strength can be obtained using the dynamic wireline log data. The dynamic UCS profile can be used for future wellbore failure analysis to estimate the rate of penetration for safe drilling operations. In this research, we aim to introduce new correlations for determination of the in situ rock strength (or UCS) and also to investigate the importance of log variables involved in UCS where the dynamic nature of reservoir formation is captured. It is worth noting that the relation between the field rock strength and laboratory rock strength (particularly, intact rock strength) is a key issue in geotechnical and rock engineering. Comparison of laboratory and field rock strength data and development of new models based on both types of data would be of high importance for future work.

In the current research investigation, a correlation matrix (Table 5) is introduced to show the relationship between UCS and other formation properties. Furthermore, we employ the machine learning strategies (ANN and LS-SVM) to find the most influential parameters while predicting in situ UCS values. According to the results, only three variables (e.g., DT, RB, and GR) are considered to develop the UCS correlations, which are valid for siliciclastic formations. It is worth noting that DT, RB, and GR represent the formation sonic slowness, electron density (related to porosity), and shale effect, respectively. The true resistivity (RT) of the formation is disregarded due to its low importance. As PHI can be obtained from the density log (RB) and neutron porosity, the PHI parameter is also eliminated to have only independent parameters as the inputs for estimation of UCS. The new models take into account the most contributing log parameters to figure out the lithology effect, number of electron density, and acoustic travel time of the porous formation. The main advantages of the developed models over the existing correlations for siliclastic rocks are that they capture dynamic condition with key formation parameters; they are capable of estimating continuous in situ UCS profile; and they are cost effective and can be used through a timely manner. Furthermore, the obtained UCS profile can be used to investigate the wellbore stability or rock failure criterion, and drilling performance analysis, reducing the exploration costs during the field development phase.

4 Conclusions

In this study, ANN and LS-SVM techniques are employed to predict the continuous profile of in situ rock strength (UCS) of clastic sedimentary rocks using several field log data such as true resistivity, gamma ray, bulk density, porosity, and compressional sonic travel time. The statistical indicators such as AAPE, MAPE, RMSE, and coefficient of determination (R^2) are used to evaluate the AI-based connectionist model performance. The key findings of this study are listed as follows:

- The connectionist models based on ANN with Levenberg–Marquardt training algorithm and the LS-SVM with CSA optimization strategy are capable of accurately estimating the reservoir rock strength (UCS) using log data.
- The compressional sonic travel time is the most influential parameter for determination of continuous in situ rock strength profile of siliciclastic rocks.
- The formation acoustic travel time, gamma ray, and the bulk density are essential to attain an accurate continuous formation UCS profile to capture lithology indicator, dense minerals, number of electron density, and acoustic travel time of the underground formation.
- Two correlations (models) are developed to estimate UCS through multivariate regression analysis by incorporating three influential log parameters. Similar to existing correlations, the introduced models exhibit a good performance.
- It is proven that newly developed log-based correlations can be used to predict actual in situ unconfined rock strength for clastic sedimentary rocks including shaly sand rock formations.
- The deterministic tools, log variable ranking approach, and developed correlations can be useful for field specialists, researchers, and rock engineers while dealing with rock failure analysis, geomechanics, drilling optimization as well as formation evaluation.

Acknowledgements We thank Equinor (formerly Statoil) Canada Ltd., Natural Sciences and Engineering Research Council of Canada (NSERC), and InnovateNL for providing financial support to accomplish this study at the Memorial University, St. John's, NL, Canada.

Author contributions MIM: conceptualization, modeling/simulation, validation, and original draft; SZ: supervision, technical discussion, and writing—reviewing and editing; SA: supervision, technical discussion, and writing—reviewing and editing; SB: technical discussion and reviewing, manuscript editing.

Funding The funding was funded by Equinor (formerly Statoil) Canada Ltd., CA (211162).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

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(19):587

Al-Bulushi N, King PR, Blunt MJ, Kraaijveld M (2009) Development of artificial neural network models for predicting water saturation and fluid distribution. J Petrol Sci Eng 68(3–4):197–208

Ali JK (1994) Neural network: a new tool for petroleum industry. In: Proceedings of SPE European petroleum computer conference, UK, SPE Paper 27561

Anemangely M, Ramezanzadeh A, Tokhmechi B (2017) Shear wave travel time estimation from petrophysical logs using ANFIS-PSO algorithm: a case study from Ab-Teymour Oilfield. J Nat Gas Sci Eng 38:373–387

Anemangely M, Ramezanzadeh A, Amiri H, Hoseinpour SA (2019) Machine learning technique for the prediction of shear wave velocity using petrophysical logs. J Petrol Sci Eng 174:306–327

Ashena R, Thonhauser G (2015) Application of artificial neural networks in geoscience and petroleum industry. Artificial intelligent approaches in petroleum geosciences. Springer, Cham, pp 127–166

Asoodeh M, Bagheripour P (2012) Prediction of compressional, shear, and stoneley wave velocities from conventional well log data using a committee machine with intelligent systems. Rock Mech Rock Eng 45(1):45–63

Asquith G, Krygowski D (2004) Basic well log analysis, 2nd edn. American Association of Petroleum Geologists, Tulsa, p 216

ASTM A (1986) Standard test method of unconfined compressive strength of intact rock core specimens. ASTM Publication, West Conshohocken

Bailey T, Dutton D (2012) An empirical Vp/Vs shale trend for the Kimmeridge Clay of the Central North Sea. In: 74th EAGE Conference and Exhibition incorporating EUROPEC 2012

Barzegar R, Sattarpour M, Nikudel MR, Moghaddam AA (2016) Comparative evaluation of artificial intelligence models for prediction of uniaxial compressive strength of travertine rocks, case study: Azarshahr area, NW Iran. Model Earth Syst Environ 2(2):76

Behnia D, Behnia M, Shahriar K, Goshtasbi K (2017) A new predictive model for rock strength parameters utilizing GEP method. Proc Eng 191:591–599

Bradford IDR, Fuller J, Thompson PJ, Walsgrove TR (1998) Benefits of assessing the solids production risk in a North Sea reservoir using elastoplastic modelling. In: SPE/ISRM rock mechanics in petroleum engineering. Society of Petroleum Engineers

Brekke H, Sjulstad HI, Magnus C, Williams RW (2001) Sedimentary environments offshore Norway—an overview. Norwegian Petroleum Society Special Publications, vol 10. Elsevier, Amsterdam, pp 7–37

Broch E, Franklin JA (1972) The point-load strength test. Int J Rock Mech Min Sci Geomech Abstracts 9(6):669–676

Brocher TM (2005) Empirical relations between elastic wave speeds and density in the Earth's crust. Bull Seismol Soc Am 95(6):2081–2092

Brown ET (1981) Rock characterization testing and monitoring ISRM suggested methods. Pergamon Press, Oxford

Castagna JP, Batzle ML, Eastwood RL (1985) Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. Geophysics 50(4):571–581

Ceryan N (2014) Application of support vector machines and relevance vector machines in predicting uniaxial compressive strength of volcanic rocks. J Afr Earth Sc 100:634–644

Ceryan N, Can NK (2018) Prediction of the uniaxial compressive strength of rocks materials. In: Handbook of research on trends and digital advances in engineering geology, pp. 31–96. IGI Global

Ceryan N, Okkan U, Kesimal A (2013) Prediction of unconfined compressive strength of carbonate rocks using artificial neural networks. Environ Earth Sci 68(3):807–819. https://doi.org/10.1007/s12665-012-1783-z

Chang C, Zoback MD, Khaksar A (2006) Empirical relations between rock strength and physical properties in sedimentary rocks. J Pet Sci Eng 51(3–4):223–237. https://doi.org/10.1016/j.petrol.2006.01.003

Crawford B, Alramahi B, Gaillot P, Sanz P, DeDontney N (2011) Mechanical rock properties prediction: deriving rock strength and compressibility from petrophysical properties 12th ISRM Congress

Dehghan S, Sattari GH, Chelgani SC, Aliabadi MA (2010) Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. Min Sci Technol (China) 20(1):41–46

Demirdag S, Tufekci K, Kayacan R, Yavuz H, Altindag R (2010) Dynamic mechanical behavior of some carbonate rocks. Int J Rock Mech Min Sci 47(2):307–312

Edimann K, Somerville JM, Smart BGD, Hamilton SA, Crawford BR (1998) Predicting rock mechanical properties from wireline porosities. In: SPE/ISRM rock mechanics in petroleum engineering. society of petroleum engineers

Equinor (2018) Volve Data Village. https://data.equinor.com/dataset/Volve. Accessed 05 Feb 2020

Esene C, Zendehboudi S, Shiri H, Aborig A (2020) Deterministic tools to predict recovery performance of carbonated water injection. J Mol Liq 301:111911

Esfahani S, Baselizadeh S, Hemmati-Sarapardeh A (2015) On determination of natural gas density: least square support vector machine modeling approach. J Nat Gas Sci Eng 22:348–358

Faleide JI, Bjørlykke K, Gabrielsen RH (2010) Geology of the Norwegian continental shelf. Petroleum Geoscience. Springer, Berlin, pp 467–499

Farquhar RA, Somerville JM, Smart BGD (1994) Porosity as a geomechanical indicator: an application of core and log data and rock mechanics. In: European petroleum conference. Society of Petroleum Engineers

Fener M, Kahraman S, Bilgil A, Gunaydin O (2005) A comparative evaluation of indirect methods to estimate the compressive strength of rocks. Rock Mech Rock Eng 38(4):329–343

Fjær E, Holt RM, Horsrud P, Raaen AM, Risnes R (1992) Petroleum related rock mechanics, 1st edn. Elsevier, Amsterdam, p 346

Fjær E, Holt RM, Raaen AM, Risnes R, Horsrud P (2008) Petroleum related rock mechanics, vol 53. Elsevier, Amsterdam

Freyburg E (1972) Der Untere und mittlere Buntsandstein SW-Thuringen in seinen gesteinstechnicschen Eigenschaften.

Deustche Gesellschaft Geologische Wissenschaften A; Berlin 176:911–919
Gaviglio P (1989) Longitudinal waves propagation in a limestone: the relationship between velocity and density. Rock Mech Rock Eng 22(4):299–306
Ghiasi MM, Bahadori A, Zendehboudi S (2014) Estimation of triethylene glycol (TEG) purity in natural gas dehydration units using fuzzy neural network. J Nat Gas Sci Eng 17:26–32
Greenberg ML, Castagna JP (1992) Shear-wave velocity estimation in porous rocks: theoretical formulation, preliminary verification and applications 1. Geophys Prospect 40(2):195–209
Haftani M, Bohloli B, Nouri A, Javan MRM, Moosavi M, Moradi M (2015) Influence of penetration rate and indenter diameter in strength measurement by indentation testing on small rock specimens. Rock Mech Rock Eng 48(2):527–534
Han DH, Nur A, Morgan D (1986) Effect of porosity and clay content on wave velocity in sandstones. Geophysics 51(11):2093–2107
Hoek E, Brown ET (1997) Practical estimates of rock mass strength. Int J Rock Mech Min Sci 34(8):1165–1186
Hossain Z, Mukerji T, Fabricius IL (2012) Vp-Vs relationship and amplitude variation with offset modeling of glauconitic greensand. Geophys Prospect 60(1):117–137
Islam MA (2010) Petrography and provenance of subsurface Neogene sandstones of Bengal Basin, Bangladesh. J Geol Soc India 76(5):493
Jamshidi A, Nikudel MR, Khamehchiyan M, Sahamieh RZ (2016) The effect of specimen diameter size on uniaxial compressive strength, P-wave velocity and the correlation between them. Geomech Geoeng 11(1):13–19
Jamshidi A, Zamanian H, Sahamieh RZ (2018) The effect of density and porosity on the correlation between uniaxial compressive strength and P-wave velocity. Rock Mech Rock Eng 51(4):1279–1286
Kamari A, Mohammadi AH, Bahadori A, Zendehboudi S (2014) A reliable model for estimating the wax deposition rate during crude oil production and processing. Pet Sci Technol 32(23):2837–2844
Khandelwal M, Monjezi M (2013) Prediction of backbreak in open-pit blasting operations using the machine learning method. Rock Mech Rock Eng 46(2):389–396
Kong F, Shang J (2018) A validation study for the estimation of uniaxial compressive strength based on index tests. Rock Mech Rock Eng 51(7):2289–2297
Koolivand-Salooki M, Esfandyari M, Rabbani E, Koulivand M, Azarmehr A (2017) Application of genetic programing technique for predicting uniaxial compressive strength using reservoir formation properties. J Petrol Sci Eng 159:35–48
Krishna KS, Rao DG, Murty GPS, Ramana YV (1989) Sound velocity, density, and related properties along a transect across the Bay of Bengal. Geo Mar Lett 9(2):95–102
Larionov VV (1969) Radiometry of boreholes. Nedra, Moscow **(in Russian)**
Lee MW (2013) Comparison of methods for predicting shear-wave velocities of unconsolidated shallow sediments in the gulf of mexico. US Department of the Interior, US Geological Survey, Reston
Li L, Fjær E (2012) Modeling of stress-dependent static and dynamic moduli of weak sandstones. J Geophys Res Solid Earth117(B5)
Majdi A, Rezaei M (2013) Prediction of unconfined compressive strength of rock surrounding a roadway using artificial neural network. Neural Comput Appl 23(2):381–389
Matin SS, Farahzadi L, Makaremi S, Chelgani SC, Sattari G (2018) Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. Appl Soft Comput 70:980–987
McNally GH (1987) Estimation of coal measures rock strength using sonic and neutron logs. Geoexploration 24(4–5):381–395
Meulenkamp F, Grima MA (1999) Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness. Int J Rock Mech Min Sci 36(1):29–39
Miah MI (2014) Porosity assessment of gas reservoir using wireline log data: a case study of bokabil formation, Bangladesh. Proc Eng 90:663–668
Miah MI, Ahmed S, Zendehboudi S (2019) Connectionist and mutual information tools to determine water saturation and rank input log variables. J Pet Sci Eng 190:106741
Miah MI, Zendehboudi S, Ahmed S (2020) Log data-driven model and feature ranking for water saturation prediction using machine learning approach. J Pet Sci Eng. https://doi.org/10.1016/j.petrol.2020.107291
Miller RP (1965) Engineering classification and index properties for intact rock. PhD Thesis, University of Illinois
Miller SL, Stewart RR (1990) Effects of lithology, porosity and shaliness on P-and S-wave velocities from sonic logs. Can J Explor Geophys 26(1–2):94–103
Mohaghegh S, Arefi R, Ameri S, Aminiand K, Nutter R (1996) Petroleum reservoir characterization with the aid of artificial neural networks. J Petrol Sci Eng 16(4):263–274
Mohamad ET, Armaghani DJ, Momeni E, Abad SVANK (2015) Prediction of the unconfined compressive strength of soft rocks: a PSO-based ANN approach. Bull Eng Geol Env 74(3):745–757
Momeni E, Jahed Armaghani D, Hajihassani M, Amin MFM (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
Mondol NH (2015) Well logging: principles, applications and uncertainties. Petroleum geoscience. Springer, Berlin, pp 385–425
Moos D, Zoback MD, Bailey L (2001) Feasibility study of the stability of openhole multilaterals, Cook Inlet. Alaska SPE Drill Completion 16(03):140–145
Mousavi E, Cheshomi A, Ashtari M (2018) Estimating elasticity modulus and uniaxial compressive strength of sandstone using indentation test. J Petrol Sci Eng 169:157–166
Nabaei M, Shahbazi K (2012) A new approach for predrilling the unconfined rock compressive strength prediction. Pet Sci Technol 30(4):350–359
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 in Iran. J Petrol Sci Eng 126:78–82
Negara A, Ali S, AlDhamen A, Kesserwan H, Jin G (2017) Unconfined compressive strength prediction from petrophysical properties and elemental spectroscopy using support-vector regression. In: SPE Kingdom of Saudi Arabia annual technical symposium and exhibition. Society of Petroleum Engineers
Nejatian I, Kanani M, Arabloo M, Bahadori A, Zendehboudi S (2014) Prediction of natural gas flow through chokes using support vector machine algorithm. J Nat Gas Sci Eng 18:155–163
Nouri A, Vaziri H, Kuru E, Islam R (2006) A comparison of two sanding criteria in physical and numerical modeling of sand production. J Petrol Sci Eng 50(1):55–70
Ocak I, Seker SE (2012) Estimation of elastic modulus of intact rocks by artificial neural network. Rock Mech Rock Eng 45(6):1047–1054
Odunlami T, Soroush H, Kalathingal P, Somerville J (2011) Log-based rock property evaluation-A new capability in a specialized log data management platform. In: SPE/DGS Saudi Arabia section technical symposium and exhibition. Society of Petroleum Engineers

Ojha M, Sain K (2014) Velocity-porosity and velocity-density relationship for shallow sediments in the Kerala-Konkan basin of western Indian margin. J Geol Soc India 84(2):187–191
Oloruntobi O, Butt S (2020) The Shear-wave velocity prediction for sedimentary rocks. J Nat Gas Sci Eng 76:103084
Onalo DO (2019) Dynamic data driven investigation of petrophysical and geomechanical properties for reservoir formation evaluation (Doctoral dissertation, Memorial University of Newfoundland)
Onalo D, Adedigba S, Khan F, James LA, Butt S (2018) Data driven model for sonic well log prediction. J Petrol Sci Eng 170:1022–1037
Onyia EC (1988). Relationships between formation strength, drilling strength, and electric log properties. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers
Pickett GR (1963) Acoustic character logs and their applications in formation evaluation. J Petrol Technol 15(06):659–667
Raaen AM, Hovem KA, Joranson H, Fjaer E (1996) FORMEL: a step forward in strength logging. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers
Rabbani E, Sharif F, Salooki MK, Moradzadeh A (2012) Application of neural network technique for prediction of uniaxial compressive strength using reservoir formation properties. Int J Rock Mech Min Sci 56:100–111
Rajabzadeh MA, Moosavinasab Z, Rakhshandehroo G (2012) Effects of rock classes and porosity on the relation between uniaxial compressive strength and some rock properties for carbonate rocks. Rock Mech Rock Eng 45(1):113–122
Ramcharitar K, Hosein R (2016) Rock mechanical properties of shallow unconsolidated sandstone formations. In: SPE trinidad and tobago section energy resources conference. Society of Petroleum Engineers
Rasouli V, Pallikathekathil ZJ, Mawuli E (2011) The influence of perturbed stresses near faults on drilling strategy: a case study in Blacktip field, North Australia. J Petrol Sci Eng 76(1–2):37–50
Rastegarnia A, Teshnizi ES, Hosseini S, Shamsi H, Etemadifar M (2018) Estimation of punch strength index and static properties of sedimentary rocks using neural networks in south west of Iran. Measurement 128:464–478
Razavi S, Tolson BA (2011) A new formulation for feedforward neural networks. IEEE Trans Neural Netw 22(10):1588–1598
Rostami S, Rashidi F, Safari H (2019) Prediction of oil-water relative permeability in sandstone and carbonate reservoir rocks using the CSA-LSSVM algorithm. J Pet Sci Eng 173:170–186
Samui P (2008) Support vector machine applied to settlement of shallow foundations on cohesionless soils. Comput Geotech 35(3):419–427
Sarda JP, Kessler N, Wicquart E, Hannaford K, Deflandre JP (1993) Use of porosity as a strength indicator for sand production evaluation. In: SPE annual technical conference and exhibition. SPE
Schlumberger (1998) Log interpretation principles/applications, 7th printing, Houston, p 235
Sebtosheikh MA, Motafakkerfard R, Riahi MA, Moradi S, Sabety N (2015) Support vector machine method, a new technique for lithology prediction in an Iranian heterogeneous carbonate reservoir using petrophysical well logs. Carbonates Evaporites 30(1):59–68
Sharma MR, O'Regan M, Baxter CDP, Moran K, Vaziri H, Narayanasamy R (2010) Empirical relationship between strength and geophysical properties for weakly cemented formations. J Petrol Sci Eng 72(1–2):134–142
Singh TN, Kainthola A, Venkatesh A (2012) Correlation between point load index and uniaxial compressive strength for different rock types. Rock Mech Rock Eng 45(2):259–264
Smola AJ, Sch B, lkopf (2004) A tutorial on support vector regression. Stat Comput 14:199–222
Sonmez H, Tuncay E, Gokceoglu C (2004) Models to predict the uniaxial compressive strength and the modulus of elasticity for Ankara Agglomerate. Int J Rock Mech Min Sci 41(5):717–729
Suykens J, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9:293. https://doi.org/10.1023/A:1018628609742
Suykens JAK, Van Gestel T, Brabanter J, De Moor B, Vandewalle J (2002) Least squares support vector machines. World Scientific, Singapore
Taheri-Garavand A, Ahmadi H, Omid M, Mohtasebi SS, Mollazade K, Smith AJR, Carlomagno GM (2015) An intelligent approach for cooling radiator fault diagnosis based on infrared thermal image processing technique. Appl Therm Eng 87:434–443
Tariq Z, Elkatatny S, Mahmoud M, Ali AZ, Abdulraheem A (2017) A new technique to develop rock strength correlation using artificial intelligence tools. In: SPE reservoir characterisation and simulation conference and exhibition. Society of Petroleum Engineers.
Torabi-Kaveh M, Naseri F, Saneie S, Sarshari B (2015) Application of artificial neural networks and multivariate statistics to predict UCS and E using physical properties of Asmari limestones. Arab J Geosci 8(5):2889–2897. https://doi.org/10.1007/s12517-014-1331-0
Vapnik V (1995) The nature of statistical learning theory. Springer, New York
Vernik L, Bruno M, Bovberg C (1993) Empirical relations between compressive strength and porosity of siliciclastic rocks. Int J Rock Mech Min Sci Geomech Abstracts 30(7):677–680
Weingarten JS, Perkins TK (1995) Prediction of sand production in gas wells: methods and Gulf of Mexico case studies. J Petrol Technol 47(07):596–600
Williams DM (1990) The acoustic log hydrocarbon indicator. In: SPWLA 31st annual logging symposium. Society of Petrophysicists and Well-Log Analysts
Xavier-de-Souza S, Suykens JA, Vandewalle J, Bollé D (2009) Coupled simulated annealing. IEEE Trans Syst Man Cybern Part B (Cybern) 40(2):320–335
Yagiz S, Sezer EA, Gokceoglu C (2012) Artificial neural networks and nonlinear regression techniques to assess the influence of slake durability cycles on the prediction of uniaxial compressive strength and modulus of elasticity for carbonate rocks. Int J Numer Anal Meth Geomech 36(14):1636–1650
Yang Y, Zhang Q (1997) A hierarchical analysis for rock engineering using artificial neural networks. Rock Mech Rock Eng 30(4):207–222
Yılmaz I, Yuksek AG (2008) An example of artificial neural network (ANN) application for indirect estimation of rock parameters. Rock Mech Rock Eng 41(5):781–795
Yilmaz I, Yuksek G (2009) Prediction of the strength and elastic modulus of gypsum using multiple regression, ANN, and ANFIS models. Int J Rock Mech Min Sci 46(4):803–810
Yılmaz NG (2013) The influence of testing procedures on uniaxial compressive strength prediction of carbonate rocks from Equotip hardness tester (EHT) and proposal of a new testing methodology: Hybrid dynamic hardness (HDH). Rock Mech Rock Eng 46(1):95–106
Zendehboudi S, Elkamel A, Chatzis I, Ahmadi MA, Bahadori A, Lohi A (2014) Estimation of breakthrough time for water coning in fractured systems: Experimental study and connectionist modeling. AIChE J 60(5):1905–1919
Zendehboudi S, Rezaei N, Lohi A (2018) Applications of hybrid models in chemical, petroleum, and energy systems: a systematic review. Appl Energy 228:2539–2566

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