



OPEN

# Comprehensive study on the Python-based regression machine learning models for prediction of uniaxial compressive strength using multiple parameters in Charnockite rocks

Sowmya Kochukrishnan<sup>1</sup>✉, Premalatha Krishnamurthy<sup>1</sup>, Yuvarajan D.<sup>2</sup> & Nandagopal Kaliappan<sup>3</sup>✉

The strength of rock under uniaxial compression, commonly known as Uniaxial Compressive Strength (UCS), plays a crucial role in various geomechanical applications such as designing foundations, mining projects, slopes in rocks, tunnel construction, and rock characterization. However, sampling and preparation can become challenging in some rocks, making it difficult to determine the UCS of the rocks directly. Therefore, indirect approaches are widely used for estimating UCS. This study presents two Machine Learning Models, Simple Linear Regression and Step-wise Regression, implemented in Python to calculate the UCS of Charnockite rocks. The models consider Ultrasonic Pulse Velocity (UPV), Schmidt Hammer Rebound Number (N), Brazilian Tensile Strength (BTS), and Point Load Index (PLI) as factors for forecasting the UCS of Charnockite samples. Three regression metrics, including Coefficient of Regression ( $R^2$ ), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), were used to evaluate and compare the performance of the models. The results indicate a high predictive capability of both models. Notably, the Step-wise model achieved a testing  $R^2$  of 0.99 and a training  $R^2$  of 0.988 for predicting Charnockite strength, making it the most accurate model. The analysis of the influential factors indicates that UPV plays a significant role in predicting the UCS of Charnockite.

**Keywords** Python model, Machine learning techniques, Regression, Uniaxial compressive strength, Indirect parameters

Rocks and their structures require careful planning to prevent loss of life and economic damage from human error. In civil engineering, mining, cave mining, tunneling applications, and other related fields, evaluating rock quality heavily relies on Uniaxial Compressive Strength (UCS), a crucial parameter. The primary goal of conducting the UCS test is to determine the strength and properties of the rock. UCS testing can be laborious and costly, necessitating the preparation of meticulously crafted samples, particularly in the case of soft and jointed rock formations. As a result, researchers have proposed empirical equations that relate UCS to other parameters. Past studies have investigated correlations between UCS and other rock properties, and various researchers have proposed empirical equations for metamorphic rocks, which are listed in Table 1. Since the chosen rock type for the present study is a metamorphic rock, only empirical equations for metamorphic rocks are considered.

<sup>1</sup>Department of Civil Engineering, Anna University, Chennai, Tamil Nadu, India. <sup>2</sup>Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Saveetha University, Chennai, Tamil Nadu, India. <sup>3</sup>Department of Mechanical Engineering, Haramaya Institute of Technology, Haramaya University, Dire Dawa, Ethiopia. ✉email: sowmya.civil@gmail.com; nandagopal.kaliappan@haramaya.edu.et

Source	Equation	Reference	Rock type
Basu and Kamran (2010)	$UCS = 11.103Is(50) + 37.659$	P1	Metamorphic
Kahraman and Gunaydin (2009)	$UCS = 18.45Is(50) + 13.63$	P2	Metamorphic
Singh and Singh (1993)	$UCS = 23.37Is(50)$	P3	Metamorphic
Singh et al. (2012)	$UCS = 22.8Is(50)$	P4	Metamorphic
Diamantis et al. (2009)	$UCS = 19.79Is(50)$	P5	Metamorphic
Tandon and Gupta (2015)	$UCS = 4.792Is(50) + 44.37$	P6	Metamorphic
Diamantis et al. (2009)	$UCS = 17.81Is(50)^{1.06}$	P7	Metamorphic
Diamantis et al. (2009)	$UCS = 21.54Is(50) - 6.02$	P8	Metamorphic
Fereidooni (2016)	$UCS = 24.36Is(50) - 2.14$	P9	Metamorphic
Fereidooni (2016)	$UCS = 10.03BTS + 55.19$	B1	Metamorphic
Yilmaz and Sendir (2002)	$UCS = \exp(0.818 + 0.059N)$	S6	Metamorphic
Gupta (2009)	$UCS = 0.64N + 37.5$	S7	Metamorphic
Fereidooni (2016)	$UCS = 0.02N^{2.28}$	S8	Metamorphic
Torabi et al. (2010)	$UCS = 0.0465N^2 - 0.1756N + 27.682$	S9	Metamorphic

**Table 1.** Equations for estimation of UCS using PLI, BTS, N

Several researchers have proposed empirical equations for predicting different rock types' unconfined compressive strength (UCS) using physio-mechanical parameters and non-destructive tests. For instance, Fakir et al.<sup>1</sup> developed equations for the granitoid rocks of South Africa, while Habib et al.<sup>2</sup> established excellent correlations for sedimentary rocks in Algeria. Jalali et al.<sup>3</sup> prepared linear regression equations for igneous and metamorphic rocks from different locations in Iran using block punch index, cylindrical punch index, and UCS tests. Similarly, Kurtulus et al.<sup>4</sup>, Son and Kim<sup>5</sup>, Aldeeky and Hattamleh<sup>6</sup>, Arman and Paramban<sup>7</sup>, Bolla and Paronuzzi<sup>8</sup>, and Chawre<sup>9</sup> found promising results in relating UCS to NDT tests such as Schmidt hammer, Ultrasonic pulse velocity (UPV), and total sound-signal energy for different rock types. Furthermore, Mishra et al.<sup>10</sup> conducted mechanical, physical, and petrological studies on rock types in India and classified them accordingly. Aladejare et al.<sup>11</sup> compiled a dataset of experimental correlations between uniaxial compressive strength (UCS) and various other rock properties based on published literature, which can help in selecting a suitable regression equation for estimating the UCS of a rock site.

Various studies have established predictive models for determining the UCS of rocks using soft computing techniques. Wang et al.<sup>12</sup> developed a model using a random forest algorithm that showed consistent results with laboratory tests. Dadhich et al.<sup>13</sup> analyzed the efficacy of an ML model based on various features and concluded that random forest regression was the optimal method. Wang et al.<sup>14</sup> applied two machine learning models using non-destructive and petrographic studies and showed that the extreme gradient boosting model outperformed the random forest model in predicting UCS. Tang et al.<sup>15</sup> established a predictive model using an improved least squares tree algorithm that demonstrated the model's usefulness in engineering applications. Fattahi<sup>16</sup> introduced a new relevance vector regression (RVR) method enhanced by two algorithms to forecast the UCS of weak rocks. They found that the RVR optimized by the harmony search algorithm outperformed the one optimized by the cuckoo search. Lei et al.<sup>17</sup> conducted a comparative study of six prediction models that were hybrid and based on the BP neural network, along with six optimization algorithms using swarm intelligence. They proved that FA-BP was the best model among others in predicting UCS.

Several studies have demonstrated the effectiveness of Artificial Neural Networks (ANN) and Machine Learning (ML) in predicting the UCS. Momeni et al.<sup>18</sup> showed that particle swarm optimization-based ANN predictive model outperformed conventional ANN techniques in predicting UCS through direct and indirect estimation. Abdelhedi et al.<sup>19</sup> demonstrated that the combination of multiple linear regression and ANN effectively indicates the UCS values of carbonate rocks and mortar by correlating porosity, density, and ultrasonic measurements with UCS. Ozdemir<sup>20</sup> utilized artificial intelligence-based age-layered population structure genetic programming (ALPS-GP) and an artificial neural network (ANN) to estimate the unconfined compressive strength (UCS) and found both methods to be influential. Azarafa et al.<sup>21</sup> proposed a DNN model and demonstrated its efficacy in obtaining the strength of marlstone. The model was also verified using classifiers like support vector machine, logistic regression, decision tree, loss function, MAE, MSE, RSME, R-square, etc. Wei et al.<sup>22</sup> used the artificial neural network (ANN) approach to estimate the unconfined compressive strength (UCS) of sedimentary rocks at the Thar Coalfields. Their findings indicated that the Brazilian tensile strength had the most significant influence on UCS estimation. Fang et al.<sup>23</sup> predicted equations based on various training algorithms and established the supremacy of the ANFIS model over the other models considered. Gupta and Nagarajan<sup>24</sup>, Hassan and Arman<sup>25</sup>, Liu et al.<sup>26</sup>, Shahani et al.<sup>27</sup>, and Qiu et al.<sup>28</sup> studied the performance of different machine learning models in predicting UCS. They suggested the best model based on factors such as absolute error, root mean square error, coefficient of determination, etc.

Over time, multiple techniques have been developed to predict the rocks' strength accurately. These techniques have proven effective for different types of rocks. However, this study aims to introduce two simple machine-learning methods, the linear regression model and the step-wise regression model, implemented in Python to estimate the strength of Charnockite rocks. The models analyze the parameters BTS, PLI, N, and UPV to predict

the strength of Charnockite rocks. The models also provide a matching procedure to assess the strength of rocks in Chennai City, Tamil Nadu, India.

## Materials and methods

### Location of the study site

The Perungudi region, situated in the southern part of Chennai, is a bustling commercial hub with numerous construction projects. The area is home to Charnockite and Granite rocks, widely found throughout Chennai. This article presents an alternative method for estimating the UCS of Charnockite rocks in the Perungudi region, which can also be applied to other parts of the city. Figure 1 displays the location of the Perungudi region.

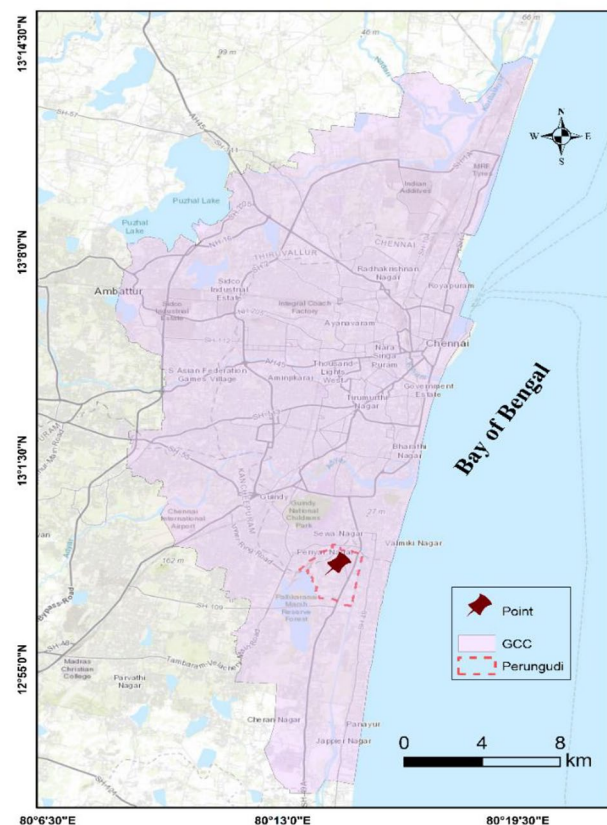
### Rock material testing and database

#### *Petrographic analysis*

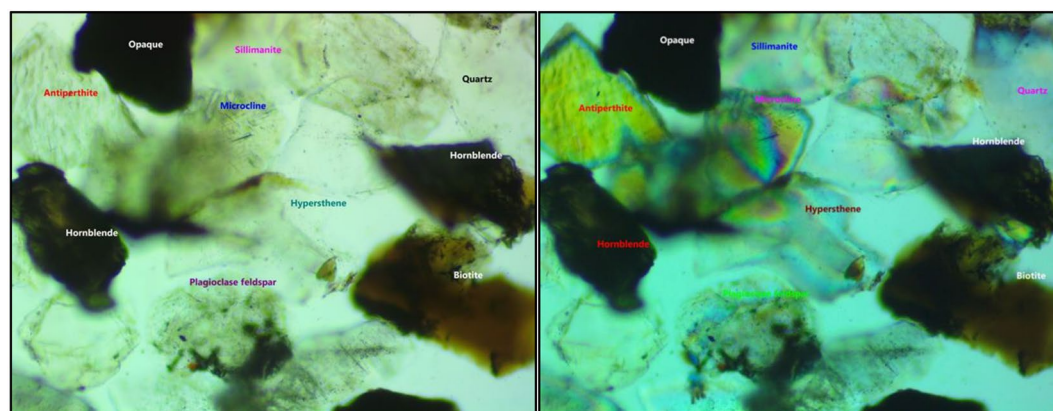
Macroscopic and microscopic observations were conducted to determine the type of rock. The rocks appeared dark grey to black in their fresh state, with a fine-to-medium-grained texture and an equigranular granoblastic homogeneous fabric without layering. It was difficult to distinguish the dark grey plagioclase from the mafic portion of the rock, and the presence of hornblende gave it a black hue. Thin sections of basic granulites exhibited granulitic texture, with minerals interfering with each other's growth. The common minerals constituting the basic granulites were labradorite, hypersthene, and augite, with a constant association of secondary hornblende, sometimes prepondering over the pyroxenes. Black opaque was present in negligible amounts, along with biotite and apatite. The development of faint gneissose structure due to the sublinear arrangement of mafic constituents was observed in some of the slides. The dark color of Charnockite is caused by thin greenish or yellowish-brown veins and stringers throughout the rock, particularly in the feldspars but also in quartz and other minerals. Images of rocks under a petrographic microscope are displayed in Fig. 2. The primary and minor minerals observed from thin section analysis are shown in Table 2.

### Laboratory testing methods

Various properties and parameters of rocks, including index properties, physical parameters, and destructive and non-destructive parameters, have been found to correlate with the UCS of different types of rocks. Equations have been established with high levels of accuracy by Bagherpour et al.<sup>29</sup> and Daoud et al.<sup>30</sup>. However, Aladejare<sup>31</sup> suggests using these equations only for the same rock types they were developed for. To investigate the UCS of rocks in the Perungudi region in Chennai, 84 specimens were collected from different locations. These specimens were then transported to the laboratory for various tests, including the UCS test, BTS test, PLI test, SHN, and



**Figure 1.** Map showing the location of Perungudi in Chennai.



Under Plain Polarizer Light/Parallel Nicols

Under Crossed Nicols

**Figure 2.** Images of rock under a petrographic microscope.

Mineral	Percentage
Plagioclase feldspar	32
Quartz	35
Hypersthene and Hornblende	30
Stained Quartz	<0.5
Mica—Biotite	<0.5
Mica—Muscovite	<0.2
Secondary Mineral—Sillimenite	<0.5
Rock type—Charnockite	

**Table 2.** Findings from the analysis of rocks under a thin section.

UPV, following ASTM standards. The specimens were prepared and tested by ASTM D4543-19<sup>32</sup> and ASTM D7012<sup>33</sup> standards. Procedures for determining pulse velocities, rock hardness by rebound hammer method, Brazilian tensile strength, and point load index were suggested by ASTM D2845-08<sup>34</sup>, ASTM D5873-95<sup>35</sup>, ASTM D3967-95a<sup>36</sup>, and ASTM D5731-02<sup>37</sup>, respectively. All tests reported in this paper adhered to the standards (Fig. 3). The results of the various tests performed in the laboratory are shown in Table 3. The table provides statistical information on the properties of the rocks, which was used as a database for the study. Figure 4 depicts a histogram plot showing the variation of the properties.

### Regression ML models using Python

In a recent study, Liu et al.<sup>26</sup> examined the UCS prediction capabilities of three different boosting machine models: adaptive boosting, category gradient boosting, and extreme gradient boosting. They compared the models' performance using five regression metrics. Another study by Xu et al.<sup>38</sup> introduced a novel prediction model called the SSA-XGBoost model. This model was more effective in predicting the UCS of rocks than six other models evaluated using RMSE, correlation coefficient, MAE, and variance interpretation. The paper presents two machine learning-based regression models, namely, Linear Regression and Step-wise Regression, that can predict the UCS of Charnockite rocks. The models are constructed using 84 sample data from 4 different parameters (UPV, N, BTS, and PLI) to find the UCS value of the Charnockite rock samples. In both models, 80% of the data are trained with the help of Python supported machine learning techniques like (scipy, numpy, pandas, seaborn, and matplotlib) and with those trained data, the model is fitted in such a way to predict the coefficient and intercept of the data. These findings are necessary to form the linear regression equation, and the r value is obtained from that. The performance of the models was evaluated, and a superior UCS estimation model was reported.

### Linear regression ML model

Linear regression is the most widely used machine learning model. The Scikit-Learn module is used in Python to build, train, and test the linear regression model. To estimate the unconfined compressive strength (UCS) of Charnockite rocks, we utilize the following rock properties: UPV, N, BTS, and PLI. These properties are extracted from a dataset of 84 Charnockite rock samples and fed into the Jupyter Notebook for prediction. To understand the dataset better, a pair plot can be generated. The dataset has been split into two distinct sets: one for training





a. UPV test on a rock specimen



b. BTS test on a rock specimen



c. Point load test on a rock specimen



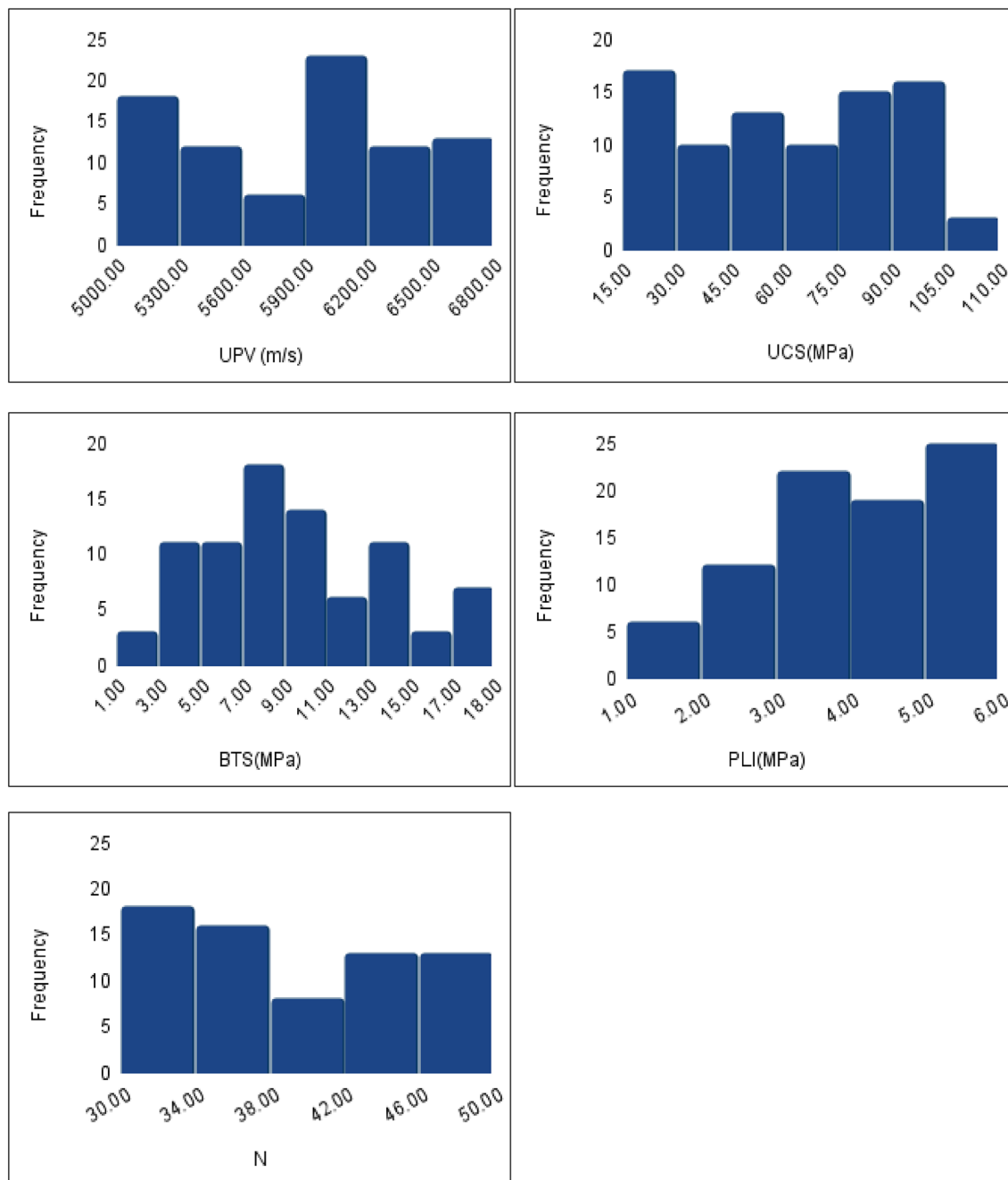
d. UCS test on a rock specimen

**Figure 3.** Laboratory tests on a rock specimen.

Property	Range	Mean	Standard deviation
PLI (MPa)	1.8–5.98	4.05	1.248
BTS (MPa)	2.25–17.88	9.49	4.17
Rebound number	24–50	37.76	7.44
UPV (m/s)	5008–6721	5858.49	525.78
UCS (MPa)	16.16–109.16	62.17	28.03

**Table 3.** Engineering properties of Charnockite rocks.

purposes and another for testing. This study used a split of 80:20, with 80% of the dataset allocated for training and the remaining 20% reserved for testing. A linear regression machine learning model was then created and trained using the segregated data. To simplify the process, sci-kit-learn was utilized. Afterward, the machine learning model was applied to the test data set to generate predictions. Scatter plots were created to compare the predicted values with the actual values. To visually evaluate the performance of the model, the residuals were plotted. Python code has been developed considering all four parameters. All 4 data have been considered for both test and train datasets. After training the model with the train data, the test data of variables (UPV, N, BTS, and PLI) are introduced to the trained model to determine the predicted UCS. Then, the model calculates the correlation coefficient by comparing the predicted UCS and the test value of the observed UCS. To represent the findings visually, Python’s graphical options present them in chart and plot formats.



**Figure 4.** Plot of the properties of Charnockite rocks.

*Python code for linear regression model*

```

Import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import scipy as sp
import seaborn as sns
plt.style.use('plot')
raw_data = pd.read_csv('R2model.csv')
x = raw_data[['UPV']]
y = raw_data['UCS']
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.9)
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(x_train, y_train)
print(model.coef_)
print(model.intercept_)
line = f'UCS = {model.coef_}UPV + ({model.intercept_})'
print(line)
#UPV = x_train
#UCS = y_train
#fig, ax = plt.subplots()
#ax.plot(UPV, UCS, linewidth=0, marker='s', label='Training Data points')
#ax.plot(UPV, model.intercept_ + model.coef_ * UPV, label=line)
#ax.set_xlabel('UPV')
#ax.set_ylabel('UCS')
#ax.legend(facecolor='white')
plt.show()
m = pd.DataFrame(model.coef_, x.columns, columns=['Coeff'])
print(m)
predictions = model.predict(x_test)
r = sp.stats.linregress(y_test, predictions)
print(r)
slope, intercept, r, p, stderr = sp.stats.linregress(y_test, predictions)
line = f'r = {r:.2f}'
fig, ax = plt.subplots()
ax.plot(y_test, predictions, linewidth=0, marker='s', label='Test Data points(UPV)')
ax.plot(y_test, intercept + slope * y_test, label=line)
ax.set_xlabel('Observed UCS')
ax.set_ylabel('Predicted UCS')
ax.legend(facecolor='white')
plt.show()
#plt.scatter(y_test, predictions)
#plt.show()
plt.hist(y_test - predictions)
plt.show()
from sklearn import metrics
a = metrics.mean_absolute_error(y_test, predictions)
line = f'MAE = {a}'
print(line)
b = metrics.mean_squared_error(y_test, predictions)
line = f'MSE = {b}'
print(line)
c = np.sqrt(metrics.mean_squared_error(y_test, predictions))
line = f'SQRT(MSE) = {c}'
print(line)

```

**Step-wise regression ML model**

Step-wise regression is a method that involves step-by-step inclusion or exclusion of variables to create a regression model that precisely explains the data with the minimum number of essential variables. This approach automatically selects the most significant variables and excludes the insignificant ones, making it superior to many other regression techniques. Initially, all four variables are considered, and at each step, the most insignificant variable is eliminated to provide a better result and to determine the significance of the data.

The data array must be defined and converted to a data frame using NumPy and pandas packages to perform step-wise regression. The Sequential Feature Selector function from the mixed package is used to conduct step-wise regression, and the chosen features are specified using the `k_features` parameter. Once the step-wise regression is complete, the desired features are used to examine the characteristics of the data. A data frame consisting solely of the selected components and the `k_feature_names_` property is then created. The dataset is divided into training and testing sets using the `train_test_split` method provided by the scikit-learn library. Then, a logistic regression model is fitted based on the chosen features. Finally, the `accuracy_score` function from the sklearn library is used to assess the model's performance. The sklearn module of Python will handle the computational complexity, so even in real-time predictions or large-scale application usage, it will be feasible and effective since this model is used for fitting regression models with predictive models. It is carried out automatically. Each step adds or subtracts the variable from the set of explanatory variables. The approaches for step-wise regression are forward selection, backward elimination, and bidirectional elimination.

In this method, Ultrasonic pulse velocity (UPV) and Schmidt Hammer Rebound Number (N) have been eliminated in each step, and the model predicted Brazilian Tensile Strength (BTS) and Point Load Index (PLI) are the significant variables to determine the result.

#### *Python code for step-wise regression model*

```

Import numpy as np
import statsmodels.api as sm
import pandas as pd
import scipy as sp
import matplotlib.pyplot as plt
plt.style.use('ggplot')
from sklearn.linear_model import LinearRegression
data = pd.read_csv('R2model.csv')
x_columns = data[['BTS', 'PLI']]
y = data['UCS']
def get_stats():
    results = sm.OLS(y, x_columns).fit()
    print(results.summary())
    get_stats()
linear_model = LinearRegression()
linear_model.fit(x_columns, y)
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x_columns, y, test_size = 0.2)
y_pred = linear_model.predict(x_test)
print("Prediction for UCS is ", y_pred)
from sklearn import metrics
a = metrics.mean_absolute_error(y_test, y_pred)
line = f'MAE = {a}'
print(line)
b = metrics.mean_squared_error(y_test, y_pred)
line = f'MSE = {b}'
print(line)
c = np.sqrt(metrics.mean_squared_error(y_test, y_pred))
line = f'SQRT(MSE) = {c}'
print(line)
r = sp.stats.linregress(y_test, y_pred)
print(r)
slope, intercept, r, p, stderr = sp.stats.linregress(y_test, y_pred)
line = f'r = {r:.2f}'
fig, ax = plt.subplots()
ax.plot(y_test, y_pred, linewidth = 0, marker = 's', label = 'Test Data points')
ax.plot(y_test, intercept + slope * y_test, label = line)
ax.set_xlabel('Observed UCS')
ax.set_ylabel('Predicted UCS')
ax.legend(facecolor = 'white')
plt.show()

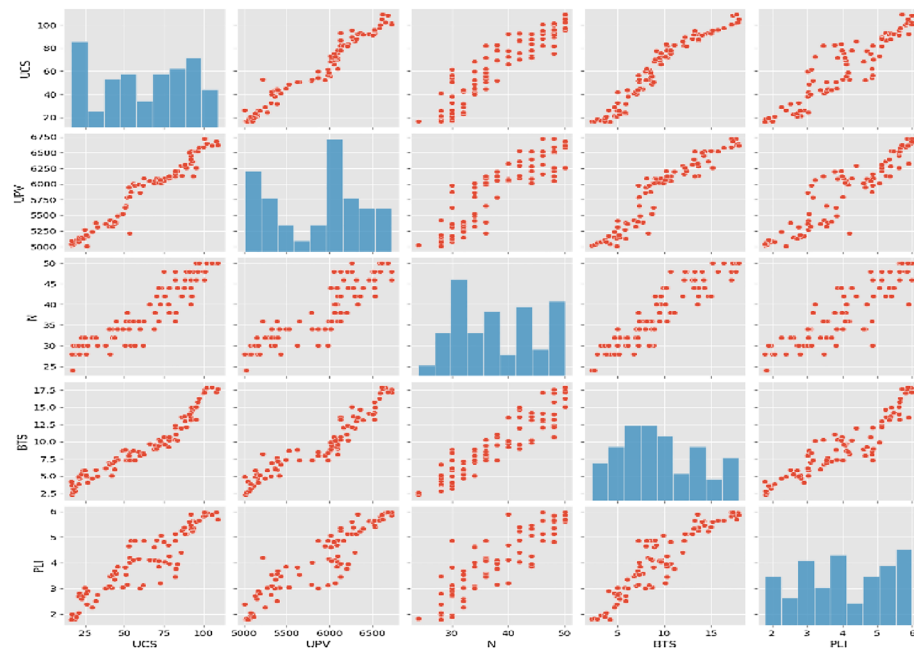
```

## Results and discussion

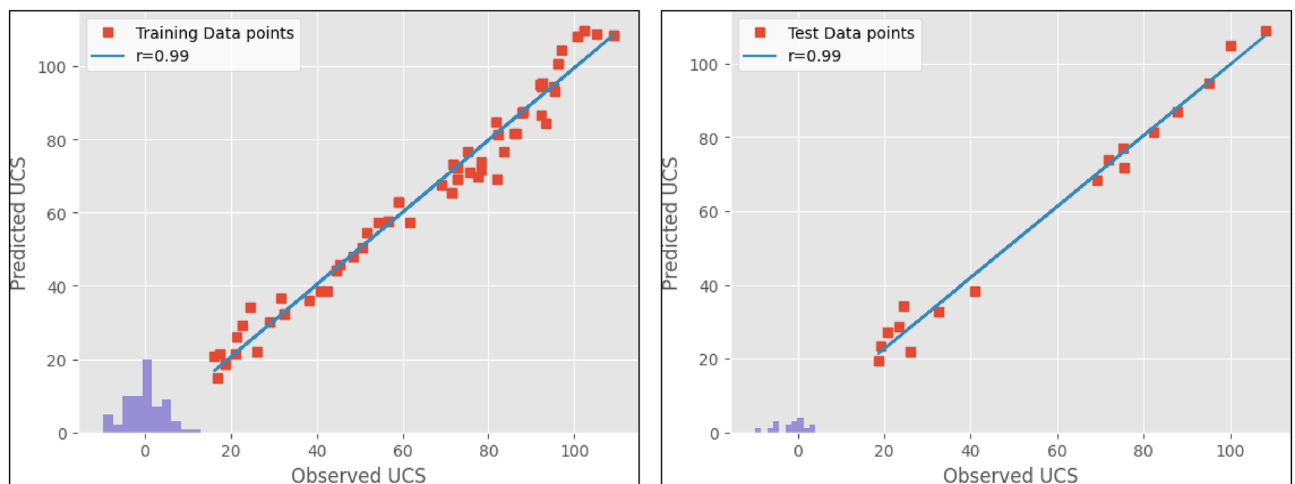
Two machine learning models, Linear Regression and Step-wise Regression, were developed using the majority (80%) of the dataset for training, and the remaining 20% was utilized for testing the model. The model will initially train the data of all four variables to predict the linear regression equation using the Python functionalities (scipy, numpy, pandas) to find the  $r$  value. After that, the sklearn functionality will handle the multicollinearity issue with the trained model and predict the UCS with the help of test data.

Regression metrics, including  $R^2$ , MAE, and RMSE, were used to evaluate the accuracy of the models. The database was provided by sampling and testing the rocks obtained from the Perungudi region in Chennai. The following datasets were used for training and testing to validate the models. These data are representative of the





**Figure 5.** Pair plot of the database.



**a.** Regression line for Training data points

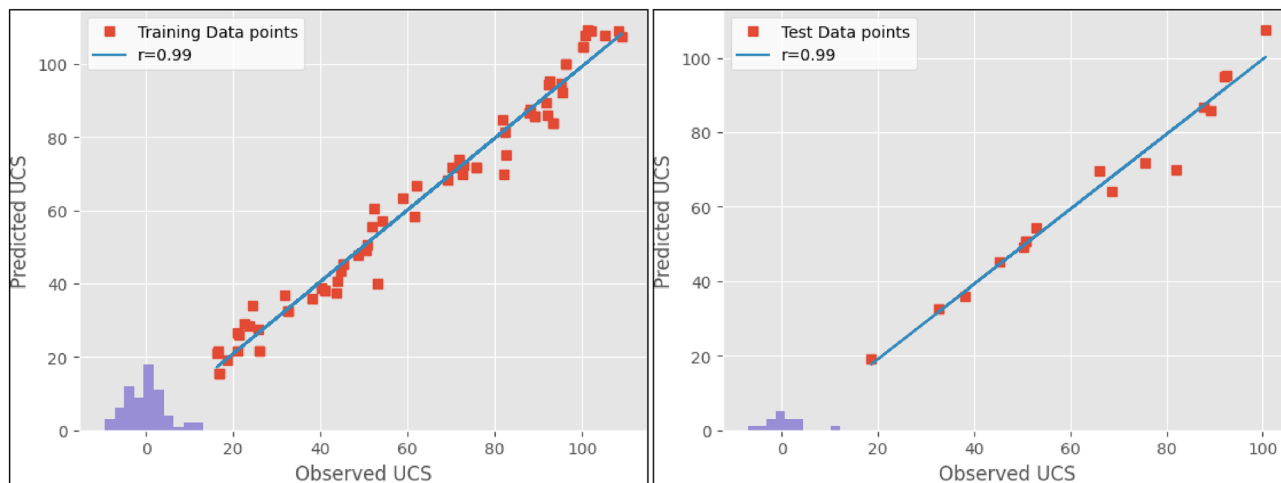
**b.** The regression line for Test data points

**Figure 6.** Regression line for ML linear regression model.

Charnockite population, as seen in the various literature datasets. UPV range is 5016–6721 m/s, UCS range is 16.72–109.16 MPa, SHN range is 24–50, PLI range is 1.81–5.98 MPa and BTS range is 2.25–17.88 MPa.

The results obtained from the experiments are presented in Figs. 5, 6 and 7. Both models proved to be highly accurate methods for predicting the UCS of the rock. Figure 5 shows the scatter distributions between the variables in the linear regression model. Figure 6 shows the predicted values of the UCS against the observed values during the training and testing of the linear regression model. Figure 7 shows the predicted values of the UCS against the experimental values during training and testing of the step-wise regression model. The experiments were carried out by the ASTM standards on 84 Charnockite samples that were recovered from various parts of the Perungudi region. The samples were tested in the laboratory to assess properties such as UCS, UPV, N, BTS, and PLI. The results from the experiments were used to evaluate the performance of the models. The predictive values that the linear and step-wise regression models prepared were compared with the observed values by regression. The results showed that both models correlated well with the observed values.

Table 4 provides the empirical equations for estimating UCS with individual parameters. The analysis showed a good correlation between all parameters and the UCS, with UPV having the highest correlation and N having the lowest. The estimated  $R^2$  for the linear regression model was 0.98 and 0.986 for training and test datasets, indicating high accuracy of data overlap. Similarly, for the step-wise regression model, the estimated  $R^2$  was 0.988



a. Regression line for Training data points      b. The regression line for Test data points

**Figure 7.** Regression line for step-wise regression.

Parameter involved	Linear regression	Correlation coefficient (R <sup>2</sup> )
UPV	$UCS = 0.052 \cdot UPV - 244.46$	0.95
BTS	$UCS = 6.54 \cdot BTS - 1.58$	0.92
N	$UCS = 3.14 \cdot N - 54.03$	0.84
PLI	$UCS = 21.06 \cdot PLI - 22.29$	0.85

**Table 4.** Equation for UCS estimation from various parameters.

Model	Training data			Test data		
	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>	MAE	RMSE
Linear regression	0.98	3.41	4.41	0.986	2.90	3.83
Step-wise regression	0.988	3.63	4.66	0.99	2.71	3.60

**Table 5.** Regression metrics table for the Models.

and 0.99 for training and test datasets, indicating high accuracy. The straight line in both scatter plots suggests that the machine-learning models have accurately predicted the UCS values. The residuals from the model's histogram show that they are typically distributed, demonstrating its efficiency.

Table 5 presents the regression metrics of the linear regression model and step-wise regression model. The indices show the variations between the values that were predicted and the ones that were observed. According to the analysis, for the training dataset in the linear regression model, the estimated values for MAE and RMSE were 3.41 and 4.41, respectively. For the test dataset, the values were 2.90 and 3.83, respectively. Similarly, for the step-wise regression model, the estimated MAE and RMSE for the training dataset were 3.63 and 4.66, respectively, and for the test dataset, the values were 2.71 and 3.60, respectively. The MAE and RMSE values were lower in the step-wise regression model than in the linear regression model. The reduction in the MAE and RMSE values indicated the high accuracy of the prediction capacity of the step-wise regression model.

## Conclusions

The strength of rock, along with other properties, plays a crucial role in civil projects. Many properties can be conveniently tested in both the laboratory and the field. In this research, we performed tests to measure the rock properties in the laboratory. We determined properties such as UPV, N, BTS, and PLI for Charnockite samples obtained from the Perungudi region in Chennai. We used Python to implement the Linear Regression ML model and Step-wise Regression to predict the Uniaxial Compressive Strength (UCS). Petrographic studies confirmed the rock as Charnockite rocks, displaying high percentages of Quartz, Feldspar, Hypersthene, and Hornblende, with slight traces of mica and Sillimanite. The statistical analysis showed UPV had the most significant effect on the UCS. We evaluated the criteria of the models (R<sup>2</sup>, MAE, RMSE), which showed high accuracy for estimating the UCS using these methods. Among the models, the Step-wise regression model showed the best results for

forecasting the UCS. Comparing the results of the two models, the Step-wise regression model, with  $R^2 = 0.99$ ,  $MAE = 2.71$ , and  $RMSE = 3.60$ , showed the best performance for estimating the UCS of Charnockite rocks.

## Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 8 November 2023; Accepted: 25 March 2024

Published online: 28 March 2024

## References

- Fakir, M., Ferentinou, M. & Misra, S. An investigation into the rock properties influencing the strength in some Granitoid Rocks of KwaZulu-Natal, South Africa. *Geotech. Geol. Eng.* **35**(3), 1119–1140. <https://doi.org/10.1007/s10706-017-0168-1> (2017).
- Habib, R., Belhai, D. & Alloul, B. Estimation of uniaxial compressive strength of North Algeria sedimentary rocks using density, porosity, and schmidt hardness. *Arab. J. Geosci.* **10**(17), 383. <https://doi.org/10.1007/s12517-017-3144-4> (2017).
- Jalali, S. H., Heidari, M., Zarrinshoja, M. & Mohseni, N. Predicting of uniaxial compressive strength of some igneous and metamorphic rocks by block punch index and cylindrical punch index tests. *Int. J. Rock Mech. Mining Sci.* **119**, 72–80. <https://doi.org/10.1016/j.ijrmms.2019.04.013> (2019).
- Kurtulus, C., Sertcelik, F. & Sertcelik, I. Estimation of unconfined uniaxial compressive strength using Schmidt hardness and ultrasonic pulse velocity. *Tehnicki Vjesnik.* **25**(5), 1569–1574 (2018).
- Son, M. & Kim, M. Estimation of the compressive strength of intact rock using non-destructive testing method based on total sound-signal energy. *Geotech. Testing J.* **40**(4), 643–657. <https://doi.org/10.1520/GTJ20160164> (2017).
- Aldeeky, H. & Hattamleh, O. A. Prediction of engineering properties of basalt rock in Jordan using ultrasonic pulse velocity test. *Geotech. Geol. Eng.* **36**(3), 3511–3525. <https://doi.org/10.1007/s10706-018-0551-6> (2018).
- Arman, H. & Paramban, S. Correlating natural, dry, and saturated ultrasonic pulse velocities with the mechanical properties of rock for various sample diameters. *Appl. Sci.* **10**(24), 9134. <https://doi.org/10.3390/app10249134> (2020).
- Bolla, A. & Paronuzzi, P. UCS field estimation of intact rock using the Schmidt hammer: A new empirical approach. *IOP Conf. Series Earth Environ. Sci.* **833**, 012014. <https://doi.org/10.1088/1755-1315/833/1/012014> (2021).
- Chawre, B. Correlations between ultrasonic P-wave velocities and rock properties of quartz-mica schist. *J. Rock Mech. Geotech. Eng.* **10**(3), 594–602. <https://doi.org/10.1016/j.jrmge.2018.01.006> (2018).
- Mishra, S., Khetwal, A. & Chakraborty, T. Physio-mechanical characterization of rocks. *J. Testing Evaluat.* **49**(3), 1976–1998. <https://doi.org/10.1520/JTE20180955> (2021).
- Aladejare, A. E. *et al.* Empirical estimation of uniaxial compressive strength of rock: Database of simple, multiple, and artificial intelligence-based regressions. *Geotech. Geol. Eng.* **39**(4), 4427–4455. <https://doi.org/10.1007/s10706-021-01772-5> (2021).
- Wang, M., Wan, W. & Zhao, Y. Prediction of the uniaxial compressive strength of rocks from simple index tests using a random forest predictive model. *Comptes Rendus Mecanique* **348**(1), 3–32. <https://doi.org/10.5802/crmeca.3> (2020).
- Dadhich, S., Sharma, J. K. & Madhira, M. Prediction of uniaxial compressive strength of rock using machine learning. *J. Inst. Eng. India Series A.* **103**(4), 1209–1224. <https://doi.org/10.1007/s40030-022-00688-4> (2022).
- Wang, Y., Hasanipanah, M., Rashid, A. S. A., Le, B. N. & Ulrikh, D. V. Advanced tree-based techniques for predicting unconfined compressive strength of rock material employing non-destructive and petrographic tests. *Materials* **16**, 3731. <https://doi.org/10.3390/ma16103731> (2023).
- Tang, Z., Li, S., Huang, S., Huang, F. & Wan, F. Indirect estimation of rock uniaxial compressive strength from simple index tests: Review and improved least squares regression tree predictive model. *Geotech. Geol. Eng.* **39**, 3843–3862. <https://doi.org/10.1007/s10706-021-01731-0> (2021).
- Fattahi, H. A new method for forecasting uniaxial compressive strength of weak rocks. *J. Mining Environ.* **11**(2), 505–515. <https://doi.org/10.22044/jme.2020.9328.1835> (2020).
- Lei, Y., Zhou, S., Luo, X., Niu, S. & Jiang, N. A comparative study of six hybrid prediction models for uniaxial compressive strength of rock based on swarm intelligence optimization algorithms. *Front. Earth Sci.* **10**, 930130. <https://doi.org/10.3389/feart.2022.930130> (2022).
- Momeni, E., Armaghani, D. J., Hajihassani, M. & Amin, M. F. M. 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> (2014).
- Abdelhedi, M., Jabbar, R., Mnif, T. & Abbes, C. Prediction of uniaxial compressive strength of carbonate rocks and cement mortar using artificial neural network and multiple linear regressions. *Acta Geodynamica et Geomaterialia.* **17**(3), 367–377. <https://doi.org/10.13168/AGG.2020.0027> (2020).
- Ozdemir, E. A new predictive model for uniaxial compressive strength of rock using machine learning method: artificial intelligence-based age-layered population structure genetic programming (ALPS-GP). *Arab. J. Sci. Eng.* **47**, 629–639. <https://doi.org/10.1007/s13369-021-05761-x> (2022).
- Azarafza, M., Bonab, M. H. & Derakhshani, R. A deep learning method for the prediction of the index mechanical properties and strength parameters of marlstone. *Materials* **15**(19), 6899. <https://doi.org/10.3390/ma15196899> (2022).
- Wei, X., Shahani, N. M. & Zheng, X. Predictive modeling of the uniaxial compressive strength of rocks using an artificial neural network approach. *Mathematics* **11**(7), 1650. <https://doi.org/10.3390/math11071650> (2023).
- Fang, Z. *et al.* Application of non-destructive test results to estimate rock mechanical characteristics—Case study. *Minerals* **13**(4), 472. <https://doi.org/10.3390/min13040472> (2023).
- Shyamala, G., Hemalatha, B., Devarajan, Y., Lakshmi, C., Munuswamy, D. B., & Kaliappan, N. Experimental investigation on the effect of nano-silica on reinforced concrete Beam-column connection subjected to Cyclic Loading. *Sci. Rep.* **13**(1). <https://doi.org/10.1038/s41598-023-43882-5> (2023).
- Hassan, M. Y. & Arman, H. Several machine learning techniques comparison for the prediction of the uniaxial compressive strength of carbonate rocks. *Sci. Rep.* **12**(1), 20969. <https://doi.org/10.1038/s41598-022-25633-0> (2022).
- Liu, Z. *et al.* Rock strength estimation using several tree-based ML techniques. *Comput. Model. Eng.* **133**(3), 1–26. <https://doi.org/10.32604/cmes.2022.021165> (2022).
- Shahani, N. M., Kamran, M., Zheng, X., Liu, C. & Guo, X. Application of gradient boosting machine learning algorithms to predict uniaxial compressive strength of soft sedimentary rocks at Thar coalfield. *Adv. Civ. Eng.* **2021**, 1–19. <https://doi.org/10.1155/2021/2565488> (2021).
- Qiu, J., Yin, X., Pan, Y., Wang, X. & Zhang, M. Prediction of uniaxial compressive strength in rocks based on extreme learning machine improved with metaheuristic algorithm. *Mathematics* **10**(19), 3940. <https://doi.org/10.3390/math10193490> (2022).

29. Bagherpour, R., Aghababaei, M. & Behina, M. Combination of the physical and ultrasonic tests in estimating the uniaxial compressive strength and Young's modulus of intact limestone rocks. *Geotech. Geol. Eng.* **35**(2), 1–9. <https://doi.org/10.1007/s10706-017-0281-1> (2017).
30. Daoud, H. S., Alshkane, Y. M. & Rashed, K. A. Prediction of uniaxial compressive strength and modulus of elasticity for some sedimentary rocks in Kurdistan Region-Iraq using Schmidt Hammer. *Kirkuk Univ. J. Sci. Stud.* **13**(1), 52–67. <https://doi.org/10.32894/kujss.2018.142383> (2018).
31. Aladejare, A. E. Evaluation of empirical estimation of uniaxial compressive strength of rock using measurements from the index and physical tests. *J. Rock Mech. Geotech. Eng.* **12**, 256–268. <https://doi.org/10.1016/j.jrmge.2019.08.001> (2020).
32. ASTM D 4543-19 Standard practice for preparing Rock Core specimens and determining dimensional and shape tolerances; ASTM International: West Conshohocken, PA, USA
33. ASTM D 7012; Standard Test Methods for Compressive Strength and Elastic Moduli of Intact Rock Core Specimens under Varying Stress and Temperatures. ASTM International: West Conshohocken, PA, USA, 2014
34. ASTM D 2845-08 Standard Test Method for Laboratory Determination of Pulse Velocities and Ultrasonic Elastic Constants of Rock; ASTM International: West Conshohocken, PA, USA
35. ASTM D 5873-95 Standard Test Method for Determination of Rock Hardness by Rebound Hammer Method; ASTM International: West Conshohocken, PA, USA
36. ASTM D 3967-95a Standard Test Method for Splitting Tensile Strength of Intact Rock Core Specimens. ASTM International: West Conshohocken, PA, USA. 2016
37. ASTM. ASTM D 5731-02, Standard Test Method for Determination of the Point Load Strength Index of Intact Rock; ASTM International: West Conshohocken, PA, USA, 2002
38. Xu, B. *et al.* Study on the prediction of the uniaxial compressive strength of rock based on the SSA-XGBoost model. *Sustainability* **15**, 5201. <https://doi.org/10.3390/su15065201> (2023).

### Author contributions

All authors have equally contributed to the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.K. or N.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024