Journal of Rock Mechanics and Geotechnical Engineering 17 (2025) 2619–2632

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Smart prediction of rock crack opening displacement from noisy data recorded by distributed fiber optic sensing

Shuai Zhao [a, b], Shao-Qun Lin [b], Dao-Yuan Tan [a, *], Hong-Hu Zhu [a, c], Zhen-Yu Yin [b, d], Jian-Hua Yin [b]

[a] School of Earth Sciences and Engineering, Nanjing University, Nanjing, 210023, China
[b] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, 999077, China
[c] Engineering Research Centre for Earth Sensing and Disaster Control of Jiangsu Province, Nanjing, 210023, China
[d] Research Centre for Resources Engineering Towards Carbon Neutrality (RCRE), The Hong Kong Polytechnic University, Hong Kong, 999077, China

ARTICLE INFO

Article history:
Received 3 February 2024
Received in revised form
22 July 2024
Accepted 1 September 2024
Available online 10 September 2024

Keywords:
Rock microcrack
Crack opening displacement
Bayesian optimization-based random forest
Anti-noise robustness
Fiber optic sensing data

ABSTRACT

The commonly used method for estimating crack opening displacement (COD) is based on analytical models derived from strain transferring. However, when large background noise exists in distributed fiber optic sensing (DFOS) data, estimating COD through an analytical model is very difficult even if the DFOS data have been denoised. To address this challenge, this study proposes a machine learning (ML)-based methodology to complete rock's COD estimation from establishment of a dataset with one-to-one correspondence between strain sequence and COD to the optimization of ML models. The Bayesian optimization is used via the Hyperopt Python library to determine the appropriate hyper-parameters of four ML models. To ensure that the best hyper-parameters will not be missing, the configuration space in Hyperopt is specified by probability distribution. The four models are trained using DFOS data with minimal noise while being examined on datasets with different noise levels to test their anti-noise robustness. The proposed models are compared each other in terms of goodness of fit and mean squared error. The results show that the Bayesian optimization-based random forest is promising to estimate the COD of rock using noisy DFOS data.

1. Introduction

Crack is a visual indicator of rock failure patterns in the field of rock engineering. Crack opening displacement (COD) is often used as one of the important parameters that reflect the severity of cracks. Understanding the change in COD is beneficial for study of crack growth, which is closely related to energy exploitation, such as exploitation of shale gas (Cong et al., 2022; Ma et al., 2024a) and oil (Wang et al., 2022; Martyushev et al., 2023). Controlling COD within a certain threshold is of vital importance for deep rock structures, such as radioactive waste disposal tunnels (Bernier et al., 2017), wellbore stability in deep gas-bearing formations (Ma et al., 2024b), and rock tunnels for mining (Zhao et al., 2017). Therefore, it is necessary to precisely measure or estimate the COD of rock cracks.

Various monitoring devices and/or techniques, such as fiber Bragg grating (FBG)-based sensors (Babanajad and Ansari, 2017; Morgese et al., 2022), digital image correlation (DIC) (Munoz and Taheri, 2017; Aliabadian et al., 2019), Michelson interferometer (Feng et al., 2013), and linear variable differential transformer (LVDT) sensors (Bassil et al., 2020), have been developed for COD measurement. However, the above-mentioned methods either only measure COD at a limited number of measure points or have cumbersome procedures to use. For instance, the complicated post-processing procedures are needed when using DIC method to measure COD. To address the aforementioned limitations, distributed fiber optic sensing (DFOS) techniques are used by numerous researchers to monitor the strain of structures and then estimate the COD using the strain data (e.g. Li et al., 2022; Li et al., 2023; Liu and Bao, 2023), because the DFOS techniques demonstrate unique advantages over the aforementioned monitoring techniques in terms of high-spatial-resolution distributed sensing, large sensing range, and immunity to electromagnetic and electrical interference. In addition, the strain accuracy of DFOS techniques (3 με under fiber spatial resolution of 1 mm) (Zhang et al., 2024) is higher than that

* Corresponding author.
E-mail address: dytan@nju.edu.cn (D.-Y. Tan).
Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

https://doi.org/10.1016/j.jrmge.2024.09.009

of StereoDIC (50 με) (Abdulqader and Rizos, 2020). For instance, Feng et al. (2013) established the relationship between the COD and DFOS data by introducing the COD as an additional local discontinuity of the host material into the strain transfer equations, which enabled estimation of the COD using measured strain data. Based on the results in Feng et al. (2013)'s study, several analytical models were further developed to quantify the COD from the DFOS strains (Babanajad and Ansari, 2017; Bassil et al., 2020; Morgese et al., 2022). These analytical models can achieve a satisfactory result on the COD estimation when minimal noise is associated with the DFOS data. However, when the DFOS data have large background noise, estimating COD through these analytical models is very difficult even though the data are processed using noise reduction techniques.

In recent years, deep learning (DL) and machine learning (ML) have found an increasingly wide utilization in the field of civil engineering (Zhao et al., 2023b, 2024; Chen et al., 2024; Guo et al., 2024; Li et al., 2024). ML has a relatively strong ability to map the nonlinear relationship between inputs and outputs, and this characteristic makes it be applied to microcrack identification using DFOS data (Song et al., 2020, 2021; Zhao et al., 2023a). Zhao et al. (2023a) developed a hybrid attention convolutional neural network (HACNN) for rock microcrack identification. The HACNN demonstrated anti-noise robustness and achieved relatively satisfactory results among different DFOS datasets with different signal-to-noise ratios (SNRs). Now that a well-designed ML model can be robust to noise when identifying rock microcracks, it may also possess anti-noise robustness in prediction of the COD of rock. Hence, it would be meaningful if ML models could be developed to predict the rock COD from the DFOS data with strong noises, given that the analytical models are challenging to do this.

An examination of the existing literature on application of the ML regression models in geotechnical engineering (e.g. Matin et al., 2018; Zhang et al., 2021a, 2021b; Zhu et al., 2021; Hou et al., 2022, 2023) reveals that the random forest (RF) (Breiman, 2001) and extreme gradient boosting (XGBoost) (Chen and Guestrin, 2016) become a relatively good option for predicting geotechnical parameters. For example, Zhang et al. (2021b) proposed a reasonably practical strategy of using XGBoost and RF models to predict undrained shear strength of soft sensitive clays using five basic soil parameters, which provides strong support for building surrogate models for estimation of soil parameters relevant to design. The reason that RF is so popular is because RF is unexcelled in accuracy among current algorithms owing to its integrating results of many learners. Moreover, an RF with an appropriate number of trees also runs efficiently on large databases because learners of it can be operated in parallel (Breiman, 2001). In addition, RF can perform random sampling on a given dataset and randomly select features from data, which makes RF has stable performance and is robust to data noise. Motivated by these characteristics of RF, especially its robustness against noise, RF-based models are developed for predicting COD of rock microcracks in this study. As is known, hyper-parameters generally have a significant effect on the performance of ML models. A poorly-configured ML may perform no better than chance. For example, in a RF model, it is not that more trees are better, because more tress may affect the training speed. Hence, the Bayesian optimization is used to determine hyper-parameters of ML models in this study. Compared to the two commonly used parameter tuning methods (i.e. grid search methods and random search methods), Bayesian optimization has fast speed and will not miss important points in the search space (Bergstra et al., 2015), which enables the best parameters for a ML model.

In this study, a machine learning (ML)-based methodology is proposed to predict COD of rock from DFOS data. The idea of this study is to train the Bayesian optimization-based ML models using DFOS data with minimal noise from a laboratory test while testing the trained ML models using data with different SNRs. The idea arises from the fact that different levels of noise (varying widely depending on the environment) are associated with the field fiber optic sensing data, and not all the data can be obtained from different noisy environment and made into labelled training samples. Hence, it is very meaningful for COD prediction from the noisy field fiber optic sensing data if the used Bayesian optimization-based ML models are robust against data noise.

2. Analytical model for quantification of crack opening displacement

To prevent the brittle fiber core from damage during structure monitoring, the silica bare fiber is normally coated with a protective coating layer. Furthermore, the optical fiber is commonly attached to the surface of a structure through different types of adhesives (Fig. 1). However, the protective coatings and adhesives would result in imperfect strain transfer between the bare fiber and structures due to their low modulus compared to that of the silica bare fiber. Based on assumptions, strain transfer between the bare fiber and host material (i.e. material being monitored) was investigated (Ansari and Libo, 1998; Li et al., 2006).

Feng et al. (2013) found the optical fiber sensor would bridge crack appearing during monitoring when subjected to localized strain discontinuity at the crack location. To consider the effect of crack-induced optical fiber strain localization on the strain transfer, they introduced the COD into the strain transferring model (STM) and derived an equation to describe the relationship between COD and strain distribution in the optical fiber. Based on the Feng et al. (2013)'s approach, a new analytical model was developed by Bassil et al. (2020) to consider the effects of imperfectly bonded multi-layers on strain transfer between the optical fiber and host material. The proposed analytical model (denoted as Bassil's model) enabled the precise estimation of COD through the measured strains at the crack location. Considering that multiple cracks would occur on the surface of a cylindrical granite specimen during the uniaxial compressive strength (UCS) test, Lin et al. (2021) introduced Bassil's model to calculate the COD of each microcrack on the surface of the granite specimen. In this study, the granite, epoxy adhesive layer, and optical fiber form a typical three-layer system, as presented in Fig. 1. The relationship between the measured optical fiber strain (ε) and COD of each microcrack can be expressed as follows (Lin et al., 2021):

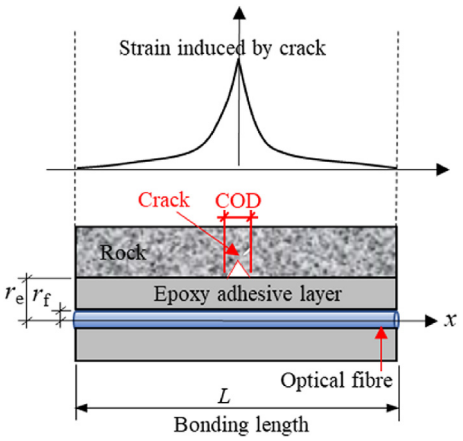


Fig. 1. Schematic of strain transfer on the fractured rock surface (adapted from Lin et al. (2021)).

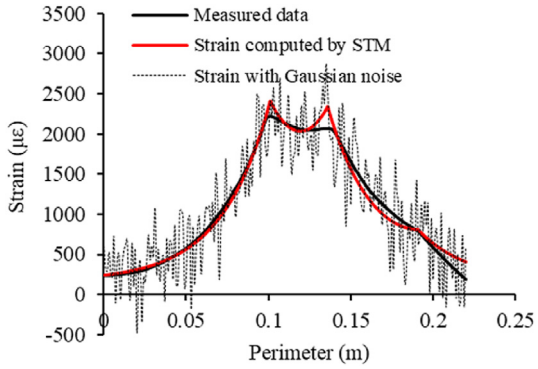


Fig. 2. Strain computed by STM and strain with Gaussian noise.

$$\varepsilon = \varepsilon_0 + \sum_{i=1}^{n} \lambda \frac{COD_i}{2} e^{-\lambda|x-x_i|} \tag{1}$$

$$\lambda^2 = \frac{2G}{r_f^2 E_f \ln\left(\frac{r_e}{r_f}\right)} \tag{2}$$

where ε_0 is the basal strain on the surface of granite (μm/m), λ is the strain lag parameter (1/m), COD_i is the crack opening displacement of i-th microcrack (μm), x_i is the location (coordinate) of i-th microcrack along the direction of optical fiber (m), n is the total numbers of microcracks, G is the shear modulus of the epoxy adhesive material (MPa), E_f is the elastic modulus of fiber core (MPa), r_e and r_f are the radii of the epoxy adhesive layer (m) and optical fiber (m), respectively.

The noiseless crack-induced optical fiber strain and the noiseless optical fiber stain in the vicinity of the crack locations (black solid line in Fig. 2) can be fitted using Eq. (1), and the fitting curve is the red solid line, as presented in Fig. 2. As the crack locations (x_i), the optical fiber strain (ε), and the strain lag parameter (λ) is known, the COD at each location can be computed. However, when the measured strain has a certain level of noise (e.g. grey dotted line in Fig. 2), it is challenging to accurately fit the strain using Eq. (1). Therefore, the COD at each location cannot be computed using Eq. (1).

ML is good at mapping the nonlinear relationship between inputs and outputs, and this characteristic makes the precise prediction of COD possible. To predict the COD using ML approaches, several consecutive strain values collected in the vicinity of crack locations and its corresponding COD can be used as input and output of ML algorithms, respectively. As the computation of COD using the noiseless monitoring strain and STM is easy, the data with minimal noise can be used to train a ML model to predict the COD. However, the obtained ML model trained using noiseless data usually performs poorly among the data with a certain level of noise. The idea of this study is to make a ML model perform well among the new collected data with different SNRs, even though the ML model is trained using noiseless data.

3. Methodology for prediction of crack opening displacement of rock microcracks

A ML-based methodology is proposed for COD prediction of rock microcracks, and the framework of the proposed methodology is illustrated in Fig. 3. In the proposed methodology, a dataset with one-to-one correspondence between strain sequence and COD is first established. Then, four ML models will be optimized using Bayesian optimization method, and the obtained Bayesian optimization-based ML models will be tested on the established testing dataset. Different levels of Gaussian noises are then added to the testing dataset to test the anti-noise robustness of the Bayesian optimization-based ML models.

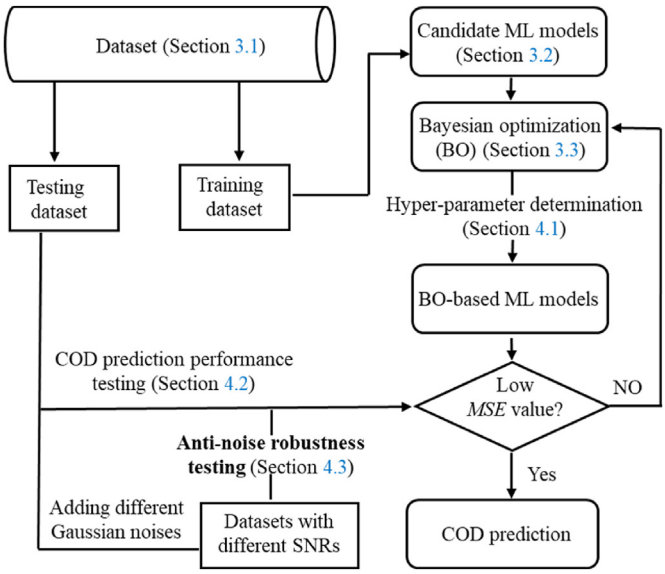


Fig. 3. Overall framework of the proposed methodology.

3.1. Dataset establishment for machine learning models

The dataset should be established prior to the development of ML models. This study aims to predict the COD of rock, and the data are collected from a previous UCS test (Lin et al., 2021). In that UCS test, a cylindrical specimen of granite with a height of 140.3 mm and diameter of 69.2 mm was loaded on the uniaxial compression tester (UCT). To collect the strain during the loading process, an optical fiber was attached to the surface of the granite specimen forming a spiral fiber (marked as 'S') and five hoop fibers (marked as 'H_1 - H_5'), as presented in Fig. 4. More details about this UCS test can be found in Lin et al. (2021)'s work. The fiber was connected to a Rayleigh optical frequency domain reflectometry (OFDR) to acquire strain data. The following characteristics of the Rayleigh OFDR ensure that the monitoring strain signal has minimal noises. First, the Rayleigh OFDR is good at demodulating the signal monitored within 100 m. The Rayleigh scattering-based system has a spatial resolution of 1 mm and acquires strain data every 5 s. The fiber attached to the surface of the granite specimen is less than 6 m; thus, when the total fiber loss is small, the fiber loss-induced noise is small. Therefore, the monitoring signal contains minimal noises. Second, the slow loading speed of the granite sample does not cause significant strain changes inside the optical fiber, which also makes the Rayleigh backscattered light suffer less interference. The strain curve plotted in Fig. 4 uses the collected strain data by OFDR. It can be seen from Fig. 4 that the collected data have few or no noises because the curve is basically free from noise-induced fluctuations.

Since the collected strain data have minimal noises, they can be fitted using the STM (i.e. Eq. (1)), and then the COD can be computed. It should be noted that the strain used in Eq. (1) to compute the COD refers to as the crack-induced optical fiber strain at the centre of crack location. The strain in the vicinity of the crack locations influences COD, given that microcracks have a certain

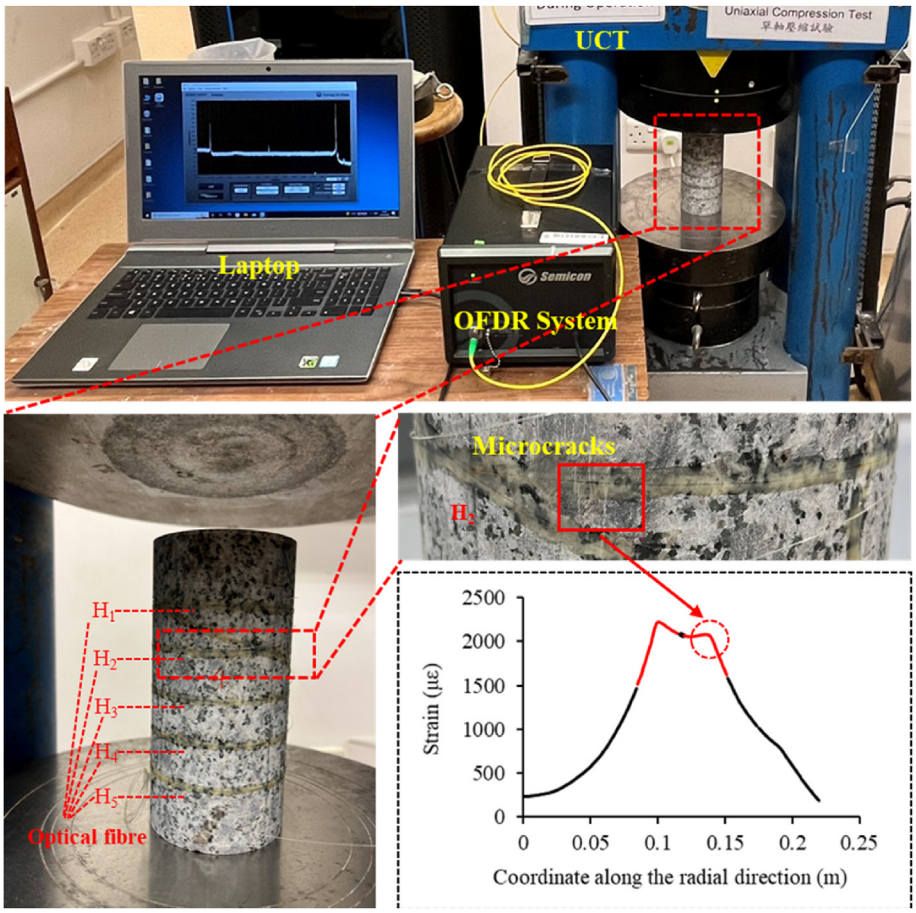


Fig. 4. Uniaxial compressive strength test.

width. Therefore, the idea in this study is to use strain sequence with a certain length to map a COD through ML algorithms. To establish a dataset with one-to-one correspondence between strain sequence and COD, the following two steps are needed:

(1) Compute the COD using Eq. (1) and the measured strain data. During the calculation of COD, the strain lag parameter (λ) of 34 is used for the optical fiber with acrylate cables, as suggested by Bassil et al. (2020). Fig. 5 shows an example of the computation of COD using the STM. It shows that three microcracks present in H_2 region at the location of 0.101m (region II), 0.136 m (region III), and 0.190 m (region IV), respectively, and three local peak points of the measured strain (black solid line) correspond to the three cracks (Fig. 5). The strain computed by the STM is plotted with a red curve in Fig. 5. The residuals between the measured strain and computed strains are −178 με, −262 με, −8 με at regions II, III, and IV, respectively. The maximum residual is only 11.2% of the measured strain. Therefore, the error of computing COD via Eq. (1) is not larger than 11.2%. It is feasible to compute the COD using Eq. (1) because of the low error.

(2) Select a strain sequence with a certain length for computing a COD. In this study, the values from 32 consecutive measurement points with the peak point as center are sampled as a strain sequence to map a COD, as presented in Fig. 6. The reason for selecting 32 points as a strain sequence is that the length ensures that the points for two neighboring 32-point strain sequences do not overlap. For example, the center (peak point) corresponding to crack I and crack II are 0.101 m and 0.136 m, respectively, as shown in Fig. 6. If taking 32 strain points with the peak point as center, the right point corresponding to crack I and the left point corresponding to crack II are at the location of 0.116 m and 0.120 m, respectively. There is a gap of three points, and the points for the two neighboring 32-point strain sequences do not overlap. Besides, the value of 32 is an integer power of 2, which facilitates operations.

Through the two steps, a dataset with one-to-one correspondence between strain sequence and COD can be established. The length of a strain sequence is 32, and the strain sequence can be regarded as a 32×1 matrix. A total of 6120 strain sequences corresponding 6120 COD values are obtained from the UCS test. A total of 5508 samples are randomly chosen for training set, while 612 samples are set aside for testing.

3.2. Machine learning algorithms

3.2.1. Support vector regression (SVR)

SVR is a supervised ML approach to handle a regression problem that allows for a real-valued function estimation. SVR handles nonlinear data through a kernel function that transforms the input data to a higher-dimensional kernel space. The detailed introduction about SVR can be found in Awad et al. (2015). To train the SVR, setting of kernel parameter (γ) and penalty parameter (C) is

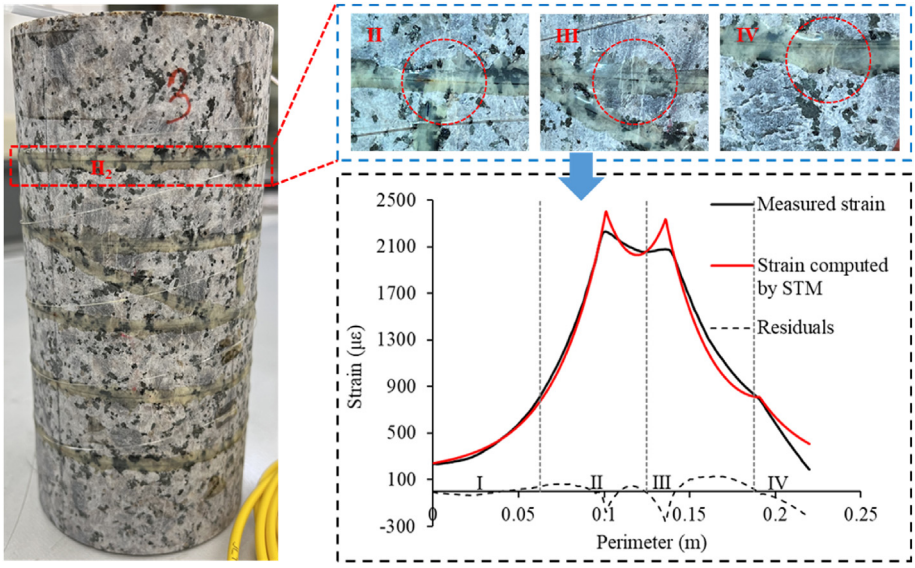


Fig. 5. The testing specimen and measured strain as well as the strain fitted by STM.

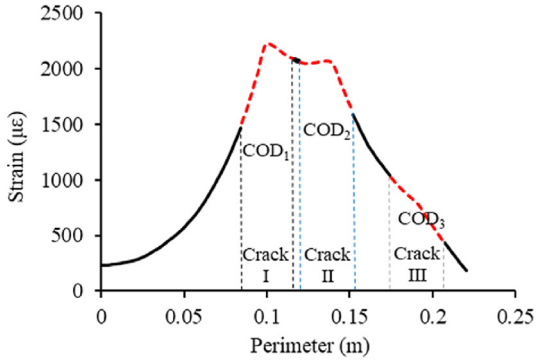


Fig. 6. Schematic of strain sequences for COD of rock microcracks.

important. Following a previous study (Zhao et al., 2023a), using 'scale' mode to automatically calculate kernel parameter (γ) can help an SVR to achieve good performance. The best penalty parameter (C) can be determined through optimization. A total of 5508 matrices (each having a length of 32×1) are used to train the SVR, and 612 matrices not used as training samples are employed to examine (i.e. test) the trained SVR model.

3.2.2. Multi-layer perceptron (MLP)

MLP consists of an input layer, one or more hidden layers, and an output layer. The number of hidden layers in an MLP and the number of nodes in each layer can vary for a given problem. The power of an MLP comes precisely from nonlinear activation functions. The initial learning rate influences the training speed and performance of an MLP. More details of MLP can be found in Murtagh (1991). In this study, the MLP with 3 hidden layers is used, and these 3 hidden layers have 24, 16, and 8 neurons. The initial learning rate and training batch size are important parameters, and they can be determined through optimization. The data that are used to train and test SVR model are also used to train and test MLP, respectively.

3.2.3. Extreme gradient boosting (XGBoost)

XGBoost is a boosting tree model that sequentially integrates the outputs of many base decision trees. During the growth of XGBoost, each newly added tree learns from its former trees and updates the residuals. A greedy algorithm is used to establish new decision trees through iteratively adding branches starting from a single leaf node. The information needed to build the decision tree, namely the weight assigned to each leaf node, the information gain after splitting the node, and the eigenvalue importance ranking function, can be obtained from the objective function. More details of XGBoost can be found in Chen and Guestrin (2016). The number of trees of XGBoost (denoted as n_estimators) and the maximum depth of the tree (denoted as max_depth) are main parameters to determine the XGBoost regression trees (RTs). The optimization of these two parameters and other parameters will be explained in Section 4.1. The data that are used to train and test SVR model are also used to train and test XGBoost, respectively.

3.2.4. Random forest

RF is an ensemble algorithm that integrates multiple trees through the idea of Bagging. The basic unit of RF is the classification tree or RT. The RT uses the heuristic method to divide the sample space, and selects the j-th variable and its value (s) as the splitting variable and splitting point. Therefore, the following two sample subspaces can be defined as

$$R_1(j,s) = \left\{ x \middle| x^{(j)} \le s \right\} \tag{3}$$

$$R_2(j,s) = \left\{ x \middle| x^{(j)} > s \right\} \tag{4}$$

The mean square error (MSE) of the two sample subspaces is computed as

$$MSE_s = \min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (x_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (x_i - c_2)^2 \right] \tag{5}$$

$$c_m = \frac{1}{N_m} \sum_{x_i \in R_m(j,s)} x_i \quad (m = 1, 2) \tag{6}$$

The RT can automatically find the j and s to obtain the minimum

MSE_s, and it will divide the sample following MSE_s until the node where the sample is located is a leaf node, as shown in Fig. 7. More details on RT can be found in Loh (2011). The RF algorithm performs random sampling with replacement (bootstrap sampling) for the established training set in Section 3.1. It will randomly select N training samples from the training set to form a sub-training set, and the sub-training set is allowed to contain duplicate samples. Therefore, the sub-training set for each tree is different. For each selection of N training samples to establish the k-th RT, approximately one-third of the samples will not participate in the generation of the k-th tree. This one-third of the samples is the out-of-bag (OOB) samples, and they will be used to compute the OOB error to evaluate the performance of k-th RT (Breiman, 2001). The growth process of RF is presented in Fig. 8. A number of n RTs will be grown to predict n values of COD, and the average value of the n COD values will be calculated as the final COD value.

In the growth process of RF, the number of trees of RF (denoted as n_estimators), the maximum depth of the tree (denoted as max_depth), and the number of features to consider when searching for the best segmentation (denoted as max_features) have significant effect on the performance of RF. These three parameters can be determined through the Bayesian optimization process. The training and testing of the RF also uses the data that are used to train and test SVR model, respectively.

3.3. Bayesian optimization

There are two commonly used parameter tuning methods for ML algorithms, namely grid search and random search. However, the speed of grid search methods is slow, and random search methods may miss important points in the search space. Fortunately, a third option exists: Bayesian optimization. Bayesian optimization enables the identification of best parameters for all models, and therefore the best models can be selected through comparison. The general process of Bayesian optimization can be summarized as follows:

(1) A surrogate function [$s(x)$] is selected as the approximation of function of the objective function [$f(x)$], and the prior probability distribution [$p(y)$] of the surrogate function can be easily initialized;
(2) An acquisition function [$\alpha(x)$] is selected, and the x value that maximizes the value of $\alpha(x)$ is determined to obtain the next value (i.e. x_{n+1}); third, the value of $f(x_{n+1})$ is then computed and stored; fourth, the $(x_{n+1}, f(x_{n+1}))$ is used to update $s(x)$ and obtain the posterior probability distribution [$p(y|x, D_n)$], which will be used as the prior probability distribution of the next iteration; finally, repeat steps 2–4 until the maximum number of iterations is reached. This process is summarized in Table 1. The key to the Bayesian optimization is to establish surrogate function and acquisition function. One of the commonly used surrogate function is the tree-structure Parzen estimator (TPE), which uses the Expected Improvement (EI) as acquisition function. More details about TPE can be referred to the authors' previous work (Zhou et al., 2021).

In this study, Hyperopt, a Python library, is used to perform the Bayesian optimization of ML algorithms. Hyperopt can be used to efficiently optimize the ML algorithm in which many hyperparameters need to be set (Bergstra et al., 2015). Information about all the points evaluated during the optimization process can be accessed and visualized via Hyperopt. Moreover, Hyperopt formalizes the practice of model optimization, so that benchmarking experiments can be reproduced. There are three steps to use Hyperopt and they are summarized in Fig. 9.

3.4. Evaluation indicators

It is crucial to evaluate the ML algorithms using appropriate performance indices. For this purpose, five indices, namely the goodness of fit (R^2), mean squared error (MSE), root-mean-square error (RMSE), standard deviation (STD), and correlation coefficient (r) are selected to evaluate the prediction performance of ML algorithms. R^2 value close to 1 indicates a better predictive ability of a ML model. The R^2 value is computed as

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{7}$$

where SST is the total sum of squares and SSE is the error sum of squares:

$$SST = \sum_{i=1}^{N} (y_i - \overline{y})^2 \tag{8}$$

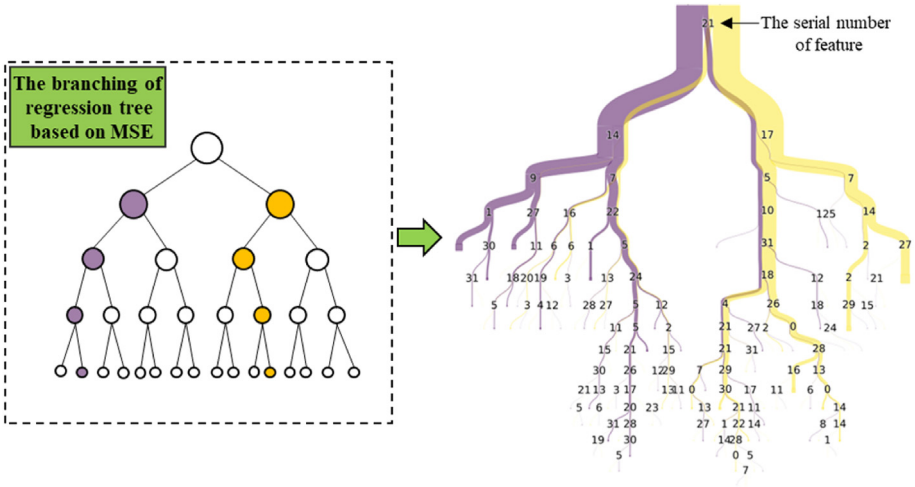


Fig. 7. Schematic of regression tree (RT).

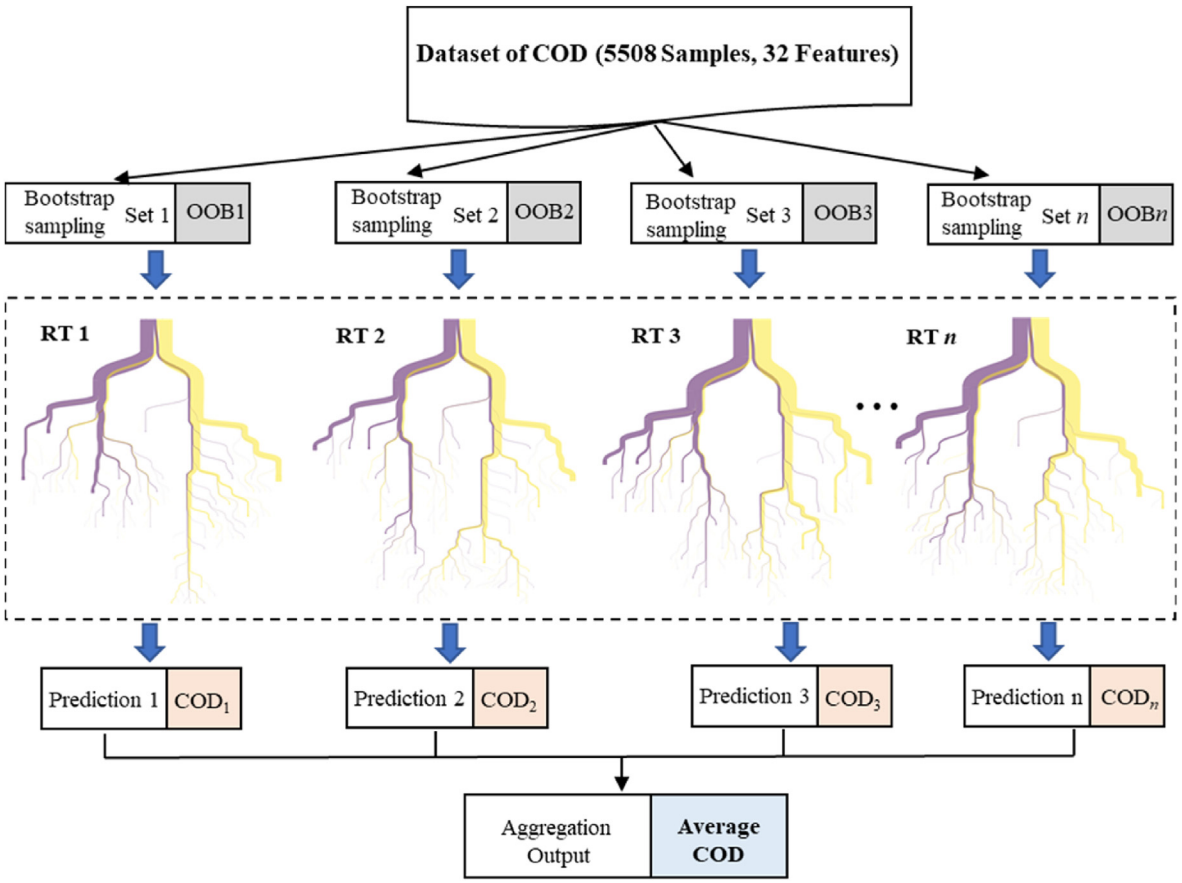


Fig. 8. Schematic of random forest.

Table 1
Process of Bayesian optimization.

Algorithm Bayesian optimization
1: for $n = 1, 2, \ldots$do
2: Select new x_{n+1} by optimizing acquisition function α
$x_{n+1} = \arg[\max_x \alpha(x; D_n)]$
3: Query objective function to obtain $f(x_{n+1})$
4: Augment data $D_{n+1} = \{D_n, (x_{n+1}, f(x_{n+1}))\}$
5: Update surrogate function [$s(x)$]
6: end for

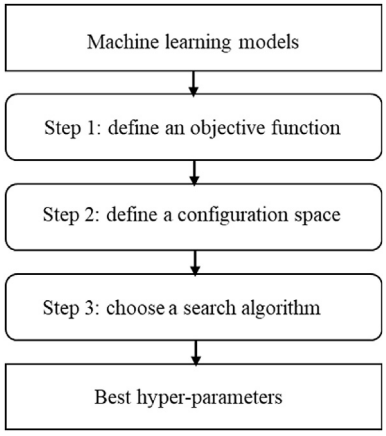


Fig. 9. The steps of using Hyperopt.

$$SSE = \sum_{i=1}^{N} (f_i - y_i)^2 \tag{9}$$

where f_i is the predicted value of ML algorithms, y_i is the label (actual value), and $\bar{y}$ is the average value of all the y_i. *SSR* is regression sum of squares ($SSR = SST - SSE$).

The MSE and RMSE values close to 0, the STD value close to the observation value (i.e. SST), and r value close to 1 indicate a better predictive ability of a ML model. The *MSE*, *RMSE*, *STD*, and r values are computed by

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (f_i - y_i)^2 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (f_i - y_i)^2} \tag{11}$$

$$STD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (f_i - \bar{f})^2} \tag{12}$$

$$r=\frac{\sum_{i=1}^{N}(f_i-\bar{f})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(f_i-\bar{f})^2}\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}} \tag{13}$$

where $\bar{f}$ is the average value of all the f_i.

4. Analysis of COD prediction results of ML models

4.1. Hyper-parameter optimization

The SVR, MLP, XGBoost, and RF are optimized using Bayesian optimization method. Through Bayesian optimization, the hyper-parameters of these four ML algorithms are summarized in Table 2. All the hyper-parameters that are to be estimated in searching space is specified by uniform distribution (Table 2). Hyperopt enables obtaining information about all the points evaluated during the optimization process. Taking BO-RF as an example, the number of trees (n_estimators), the maximum depth of the tree (max_depth), and the number of features to consider when searching for the best segmentation (max_features) are visualized, and the results are presented in Fig. 10. The lighter point in Fig. 10 indicates that RF obtains a better result. It can be known in Fig. 10 that the RF can achieve higher cross validation accuracy (i.e. cross validation R^2) when the values of n_estimators, max_features, and max_depth are set to16, 6, and 21, respectively.

The effect of the number of trees on the performance of the BO-RF is also investigated, and the results are plotted in Fig. 11. It shows in Fig. 11, the performance of the model hardly increases when the number of trees is higher than 16.

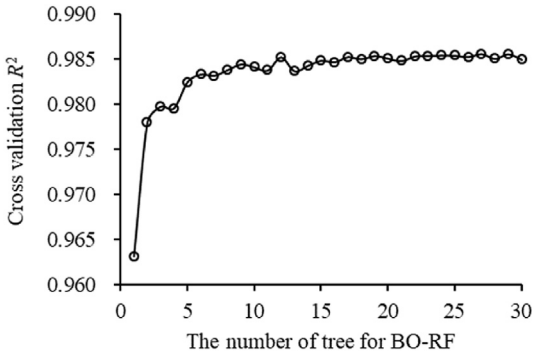


Fig. 11. The effect of number of trees on the performance of the BO-RF.

4.2. Results of ML models on training and testing datasets

The above mentioned four ML models are executed using the hyper-parameters introduced in Table 2 on a desktop equipped with an Intel Core i9-12900K CPU, an NVIDIA RTX 3090 GPU, and 64 GB of RAM. These models are trained and tested using the established dataset in Section 3.1. It should be noted that the data in the dataset have minimal noises.

The 10-fold cross-validation is used to investigate the relationship between training samples (sizes) and the goodness of fit (R^2). As indicated by R^2 during training process (Fig. 12), the learning ability of BO-SVR and BO-MLP is poorer than that of BO-XGBoost and BO-RF when the training samples are less than 3000. The learning ability of BO-RF is slightly higher than that of BO-XGBoost. The testing R^2 score of these four ML models increases with increase of the training samples. These ML models, except for BO-

Table 2
The best hyper-parameters of the four used ML models.

Approach	Hyper-parameters	Searching space	Optimal value	Tuning time (s)
SVR	C	[0, 100]	60	170
MLP	batch_size	[5, 20]	10	195
	learning_rate_init	[0.0005, 0.005]	0.001	
XGBoost	n_estimators	[100, 1000]	456	262
	max_depth	[2, 20]	12	
	learning_rate	[0.001, 0.01]	0.01	
	gamma	[1, 9]	6.4873	
	min_child_weight	[0, 10]	7	
	subsample	[0.8, 1]	0.9569	
	colsample_bytree	[0.1, 1]	0.3448	
	reg_alpha	[0.1, 1]	0.97	
	reg_lambda	[0.1, 1]	0.1618	
RF	n_estimators	[1, 25]	16	61
	max_features	[1, 8]	6	
	max_depth	[1, 25]	21	

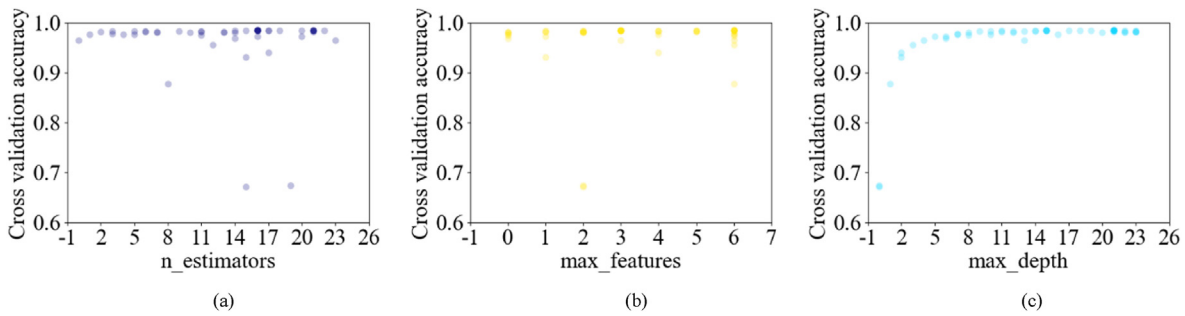


Fig. 10. Determination of hyper-parameters.

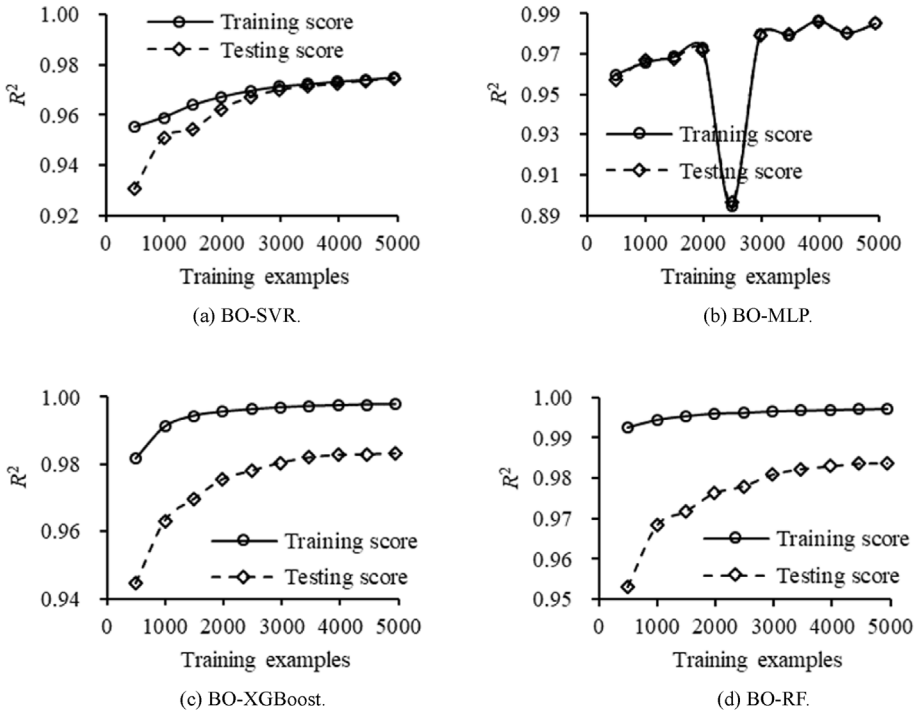


Fig. 12. The effect of training samples on the performance of the four ML models.

MLP, almost maintain a stable R^2 score when the training samples are large than 3000. It shows that the division of the amount of data for the training and testing dataset are appropriate.

Fig. 13 shows the Taylor diagram of the four ML models in predicting COD on the testing dataset. The Taylor diagram includes three metrics: STD, *r*, and RMSE. For an ideal ML model to predict COD, the value of *r* and RMSE would be close to 1 and 0, respectively, while the STD value would be close to the observation value. Herein, the observation value refers to as STD of the label (measured value). Therefore, the closer the point to the observation point in Taylor diagram, the better the performance of the ML model. For the testing dataset, it can be seen in Fig. 13 that the capability of the BO-MLP, the BO-XGBoost, and the BO-RF in predicting COD are very similar, and these three ML models are superior to the BO-SVR model in terms of COD prediction.

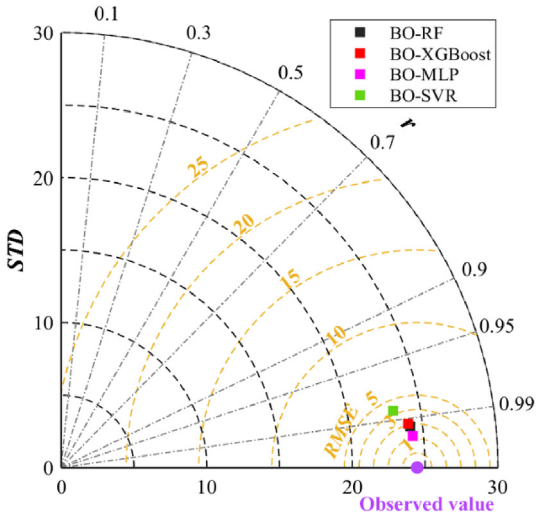


Fig. 13. Taylor diagrams of different ML models for the testing dataset.

In this study, the labels of COD (i.e. the COD computed by STM) in the established dataset are denoted as the measured COD. Fig. 14 shows the distribution of measured COD and predicted COD along the datum line $y = x$. If the predicted COD is equal to the measured COD predicted by the four ML models, the points will fall on the datum line $y = x$. It can be seen in Fig. 14 that all points are concentrated near the datum line $y = x$. The discreteness of the distribution of predicted COD for the BO-MLP, BO-XGBoost, and BO-RF is relatively similar, which means the predictive ability of these three models for COD is similar. The discreteness of the distribution of predicted COD for the BO-SVR model is slightly larger than that for the other three models; however, the predicted results of COD are also satisfactory.

From the above analysis of 10-fold cross-validation results, the Taylor diagram, and the prediction results of COD, it shows that all the SVR, MLP, XGBoost, and RF can achieve satisfactory performance on the testing data with minimal noises after the Bayesian optimization of hyper-parameters, and their predictive ability are very similar. However, in engineering practice, the field data collected from sensors or optical fibers typically contain various types of noises, which vary widely depending on the environment. Therefore, it is necessary to investigate the predictive ability of the models that are trained using the data with minimal noises, when facing noisy testing data.

4.3. Performance comparisons between different approaches under noisy environment

Training ML models on clean data (or data with minimal noise) and evaluating their performance on noisy data are an effective means to examine a model's robustness against data noises. The additive white Gaussian noise has a clear analytical expression and is often used to study the noise resistance of DL or ML models (Zhang et al., 2018; Zhao et al., 2019). Therefore, in this study, only the additive white Gaussian noises with different SNRs are added to the established testing dataset to create composite noisy datasets.

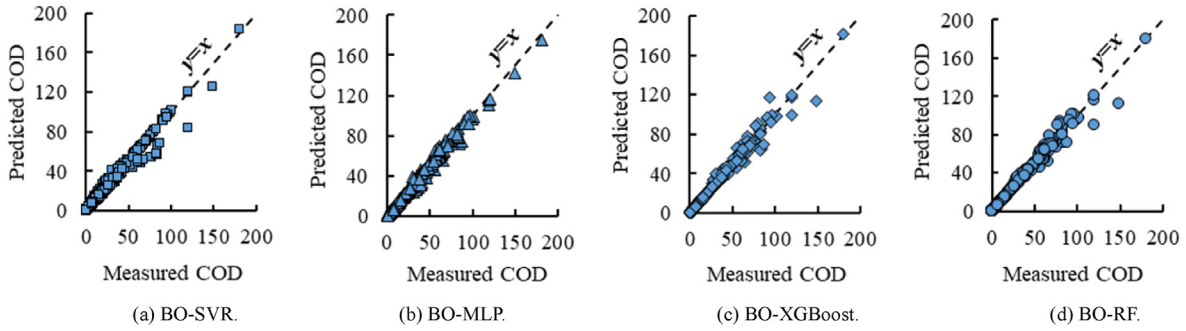


Fig. 14. Comparison of the measured COD and COD predicted by different models on the testing dataset.

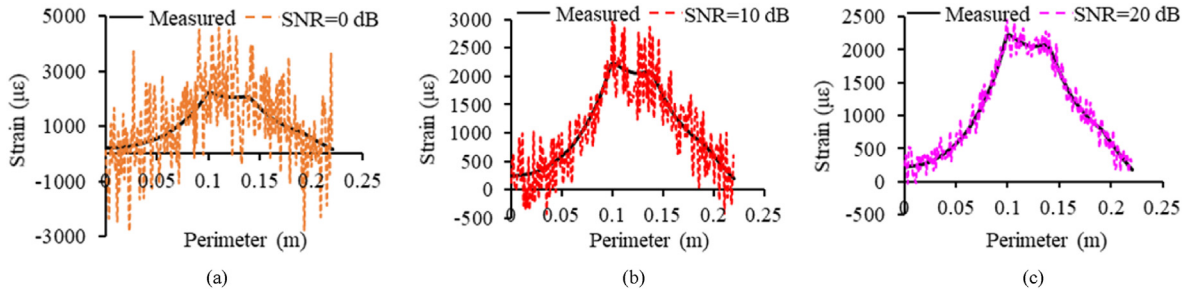


Fig. 15. A monitored strain sample that is added Gaussian noises with different SNR values.

The larger the value of SNR, the smaller the noise intensity. The *SNR* (unit: dB) is computed as

$$SNR = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{14}$$

where P_{noise} and P_{signal} are the power of noises and signals, respectively.

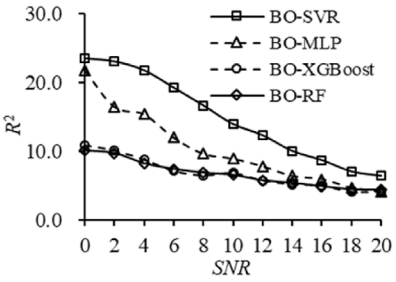
The composite noisy datasets with an SNR value ranging from 0 to 20 dB are used to investigate the anti-noise robustness of BO-SVR, BO-MLP, BO-XGBoost, and BO-RF model. Fig. 15 presents an original measured data fragment and its corresponding composite noisy data fragment with SNR values of 0, 10 dB, and 20 dB. The composite signals with SNR value of 0 and 10 dB fluctuate violently, as displayed in Fig. 15.

The goodness of fit (R^2) and MSE of the four ML models on the composite noisy datasets with different SNRs are summarized in Table 3 and Fig. 16. It can be seen in Fig. 16a that the R^2 values of the BO-XGBoost model and the BO-RF model are far better than those of the BO-MLP model and the BO-SVR model when the SNR value is lower than 6. Even when the SNR value is 0, the R^2 value of the BO-XGBoost model and the BO-RF model are above 0.8, whereas the R^2 values of the BO-MLP model and the BO-SVR model are below 0.25. As displayed in Fig. 16b, the MSE of the BO-XGBoost model and the BO-RF model is much smaller than that of the BO-MLP model and the BO-SVR model when the SNR value is lower than 6. Moreover, the MSE of the BO-RF model is smaller than that of the BO-XGBoost model. The Taylor diagram of the four ML models in predicting COD on the composite dataset with SNR of 0 is presented in Fig. 17. Compared with the points representing the BO-SVR and the BO-MLP model, the points representing the BO-XGBoost and the BO-RF model are closer to the observation point. Conclusion can be drawn from above analysis that the BO-XGBoost model and the BO-RF model are more robust for Gaussian noises, in comparison with the BO-MLP model and the BO-SVR model. The anti-noise robustness of the BO-RF model is slightly superior to that of the BO-XGBoost model.

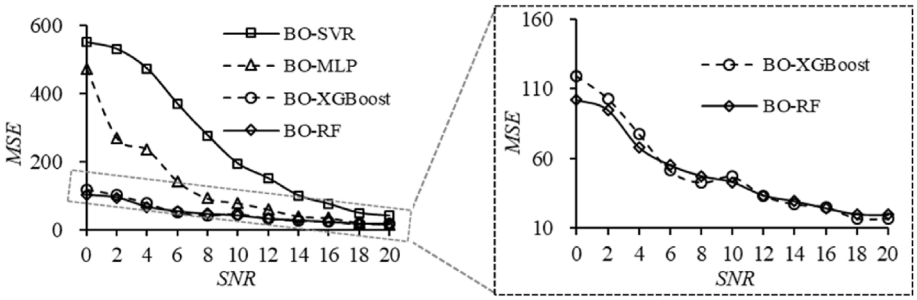
Fig. 18 shows an example of the COD prediction results from the four ML models on the composite dataset with $SNR = 2$ dB. It shows that the discreteness of the distribution of predicted COD for the BO-SVR and the BO-MLP model is much larger than that for the BO-XGBoost and the BO-RF model. Moreover, the COD values predicted

Table 3
The R^2 and *MSE* of different ML models on composite datasets with different SNR values.

Model	Metric	Testing dataset	*SNR* (dB)					
			20	16	12	8	4	0
BO-SVR	R^2	0.9708	0.9300	0.8722	0.7452	0.5354	0.2080	0.0774
	MSE	17.4358	41.7974	76.3385	152.2228	277.5179	473.1061	551.1096
BO-MLP	R^2	0.9922	0.9717	0.9403	0.8975	0.8443	0.6033	0.2047
	MSE	4.6632	16.8875	35.6695	61.2371	93.0263	236.9491	475.0907
BO-XGBoost	R^2	0.9839	0.9715	0.9577	0.9439	0.9282	0.8694	0.8002
	MSE	9.6039	17.0108	25.2705	33.4845	42.9031	78.0430	119.3357
BO-RF	R^2	0.9863	0.9672	0.9591	0.9438	0.9204	0.8858	0.8293
	MSE	8.1780	19.5967	24.4326	33.5848	47.5460	68.1789	101.9942



(a) Relationship between R^2 and *SNR*.



(b) Relationship between *MSE* and *SNR*.

Fig. 16. The R^2 and *MSE* of different ML models among composite noisy datasets.

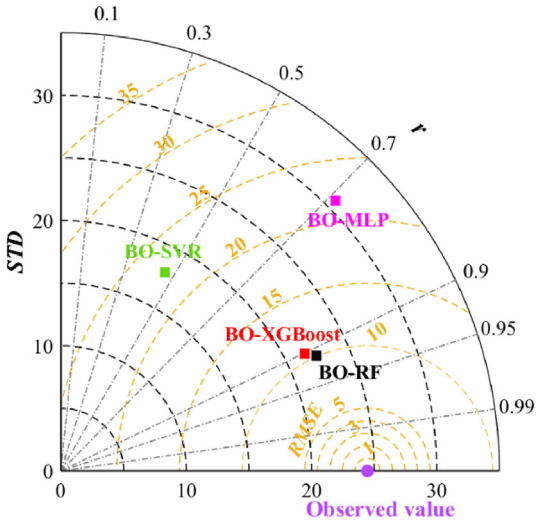


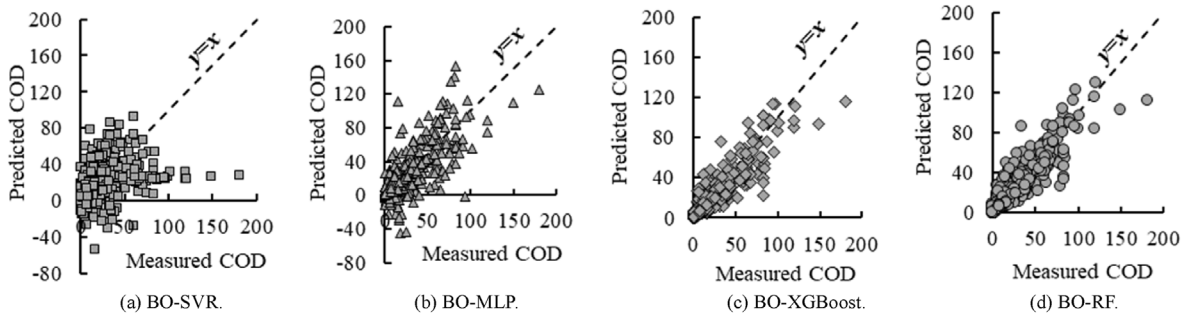
Fig. 17. Taylor diagrams of different ML models for the composite dataset with *SNR* = 0 dB.

by the BO-SVR and BO-MLP models have negative values, indicating that the COD predictive ability of the BO-SVR and BO-MLP model is poor when the SNR is low (i.e. the noise intensity is high).

The COD values predicted by the BO-RF model on composite datasets with different SNRs are plotted in Fig. 19. As shown in Fig. 19, the predicted COD values become increasingly accurate with the increase of SNR value (i.e. the decrease of noise intensity). Even when the SNR value is low (e.g. ranging from 0 to 8), the COD values predicted by the BO-RF model can also be distributed relatively evenly on both sides of the datum line $y = x$. The BO-RF model demonstrates a good anti-noise robustness even though the noise intensity is high. It should be noted that only additive white Gaussian noise is applied to test the noise resistance of the used ML models in this study, so the BO-RF model is more useful for data with additive white Gaussian noise.

5. Discussion

The reasons that the BO-RF model achieves the best performance among the used ML models will be discussed by analyzing the differences in COD prediction performance. As shown in Table 3, the BO-SVR and the BO-MLP can achieve good performance with



(a) BO-SVR. (b) BO-MLP. (c) BO-XGBoost. (d) BO-RF.

Fig. 18. Comparison of different models in predicting COD on the composite dataset with *SNR* = 2 dB.

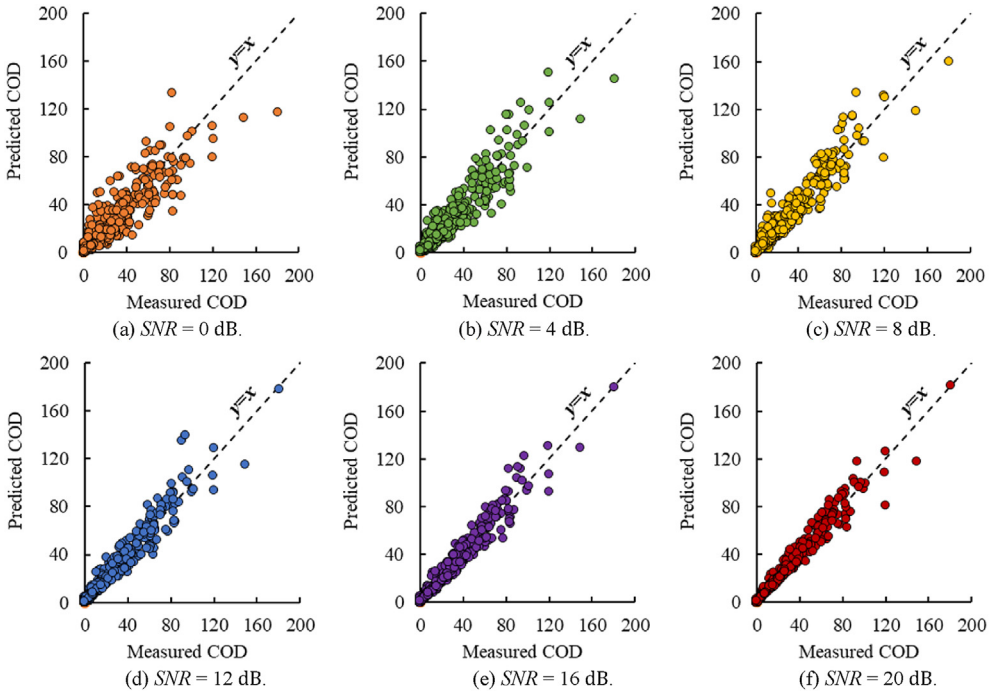


Fig. 19. The COD predicted by BO-RF among composite datasets with different SNR values.

high R^2 and low MSE values on the testing dataset, but their performance on the datasets with low SNR value is poor. The BO-SVR and the BO-MLP are single algorithms, and they can learn crack features well from a given dataset through optimizing hyperparameters. However, after adding Gaussian noises to the testing dataset, a lot of new features, as well as features similar to those on the training dataset, appear in the composite noisy datasets. The obtained BO-SVR and BO-MLP model cannot provide a good fit to the data from the noisy datasets that are not used to training process, because features on noisy datasets similar to those on the training dataset interfere with accurate prediction of COD.

The BO-XGBoost and BO-RF model are ensemble learning models, and they integrate results from many learners. During the growth of BO-XGBoost, the residual of the previous RT is learned to obtain a current RT. Each subsequent RT aims to reduce the residual of the previous RT. This iterative process continues until the residual no longer decreases. The BO-XGBoost finally integrates the results of each RT. The underlying principle of the BO-XGBoost (i.e. Boosting) is that a proper synthesis of the results of multiple learners will produce a result superior to the result of any individual learner. The prediction result of the BO-XGBoost on the datasets with strong noise are also a synthesis of the results of multiple learners. Therefore, the BO-XGBoost is robust to noise, and thus its COD prediction performance is better than that for single algorithms (i.e. BO-SVR and BO-MLP).

The BO-RF model will perform random sampling on a given dataset. This sampling method guarantees that different training datasets can be obtained, and thus the individual learners trained on these different datasets are independent and have large differences. Therefore, the ensemble learning model integrated by these individual learners have strong generalization ability. Moreover, the features among a dataset are randomly selected during the training of the ensembled BO-RF model. The importance of features obtained from the training of the BO-RF and the BO-XGBoost is shown in Fig. 20. Since there are 32 points in a strain sequence (see Section 3.1), the sequence can be regarded as having 32 features

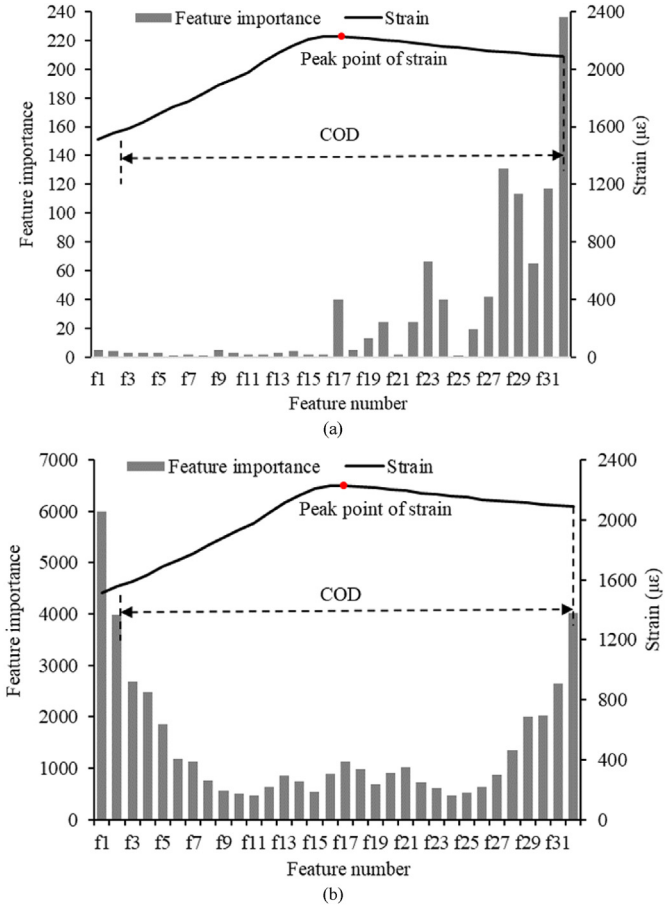


Fig. 20. Feature importance: (a) BO-RF, and (b) BO-XGBoost.

(denoted as f1 to f32). The points representing features of f1 - f16 are at the left of the peak point (representing features of f17), and the points representing features of f18 - f32 are at the right of the peak point. It can be known from Fig. 20 that the BO-RF model randomly select features, some feature points on one side of the peak point (including it) are considered very important by the BO-RF model. Whereas the feature selection of the BO-XGBoost has certain rules, the feature points of the boundary are considered most important, followed by feature points in the vicinity of peak point. It is the random sampling, random selection of features, and characteristics of result integration that make the BO-RF robust to data noise. Therefore, the anti-noise robustness of the BO-RF is better than that of BO-XGBoost on the DFOS with different SNR values.

6. Conclusions

This study proposes a ML-based methodology to achieve the crack opening displacement (COD) prediction using distributed fiber optic sensing (DFOS) data. Through the proposed methodology, the prediction of COD using noisy DFOS data with different SNR values is achieved.

(1) In the proposed methodology, a dataset with one-to-one correspondence between strain sequence and COD is first established using the DFOS data. The Bayesian optimization method is then used to select the more important parameters associated with the used ML models. Hyperopt, a Python library, is used to perform the Bayesian optimization, since Hyperopt can be used to simultaneously optimize the multiple hyper-parameters of a ML algorithm in an efficient way.
(2) A total of four ML models, namely the BO-SVR, the BO-MLP, the BO-XGBoost, and the BO-RF, are optimized using Bayesian optimization method. For the COD prediction, all four models achieve an R^2 value above 97% on the original testing dataset. To test the anti-noise ability of the BO-RF model, the original testing dataset is added into additive white Gaussian noises with SNRs ranging from 0 to 20 dB. The BO-RF model demonstrate a relatively optimal anti-noise ability, and it performs random sampling with replacement on a given dataset and randomly selects features for training. This random property increases the anti-noise ability of the BO-RF model and makes it superior to the other used ML models in terms of anti-noise robustness.

This paper provides an alternative approach to predict the microcrack's COD. For ongoing and future research, the BO-RF model will be combined with convolutional neural networks (CNNs) to achieve the more better COD prediction performance from the noisy DFOS data. The fully connected layer of a CNN can be replaced with the BO-RF to further improve the model's anti-noise robustness. It should be noted that a CNN should be delicately designed in order to prevent big errors of the CNN's extracted features from being introduced into the BO-RF model.

CRediT authorship contribution statement

Shuai Zhao: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Shao-Qun Lin:** Investigation, Data curation, Conceptualization. **Dao-Yuan Tan:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Hong-Hu Zhu:** Writing – review & editing, Supervision. **Zhen-Yu Yin:** Writing – review & editing, Supervision. **Jian-Hua Yin:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 42407250), the Fund from Research Centre for Resources Engineering towards Carbon Neutrality (RCRE) of The Hong Kong Polytechnic University (Grant No. No. 1-BBEM), and the Fund from Natural Science Foundation of Jiangsu Province (Grant No. BK20241211) are gratefully acknowledged.

References

Abdulqader, A., Rizos, D.C., 2020. Advantages of using digital image correlation techniques in uniaxial compression tests. Results Eng 6, 100109.
Aliabadian, Z., Zhao, G.-F., Russell, A.R., 2019. Crack development in transversely isotropic sandstone discs subjected to Brazilian tests observed using digital image correlation. Int. J. Rock Mech. Min. Sci. 119, 211–221.
Ansari, F., Libo, Y., 1998. Mechanics of bond and interface shear transfer in optical fiber sensors. J. Eng. Mech. 124 (4), 385–394.
Awad, M., Khanna, R., Awad, M., Khanna, R., 2015. Support vector regression. In: Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Apress, pp. 67–80.
Babanajad, S.K., Ansari, F., 2017. Mechanistic quantification of microcracks from dynamic distributed sensing of strains. J. Eng. Mech. 143 (8), 04017041.
Bassil, A., Chapeleau, X., Leduc, D., Abraham, O., 2020. Concrete crack monitoring using a novel strain transfer model for distributed fiber optics sensors. Sensors 20 (8), 2220.
Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., Cox, D.D., 2015. Hyperopt: a python library for model selection and hyperparameter optimization. Comput. Sci. Discov. 8 (1), 014008.
Bernier, F., Lemy, F., De Cannière, P., Detilleux, V., 2017. Implications of safety requirements for the treatment of THMC processes in geological disposal systems for radioactive waste. J. Rock Mech. Geotech. Eng. 9 (3), 428–434.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Chen, B., Mao, W., Lin, Y., Ma, W., Hu, N., 2024. Manufacturing-induced stochastic constitutive behaviors of additive manufactured specimens: testing, data-driven modeling, and optimization. Rapid Prototyp. J. 30, 662–676.
Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, pp. 785–794.
Cong, Z., Li, Y., Tang, J., Martyushev, D.A., Yang, F., 2022. Numerical simulation of hydraulic fracture height layer-through propagation based on three-dimensional lattice method. Eng. Fract. Mech. 264, 108331.
Feng, X., Zhou, J., Sun, C., Zhang, X., Ansari, F., 2013. Theoretical and experimental investigations into crack detection with BOTDR-distributed fiber optic sensors. J. Eng. Mech. 139 (12), 1797–1807.
Guo, W., Chen, B., Yang, Y., Xia, Y., Xiao, Q., Liu, S., Wang, H., 2024. Effect of curing regimes and fiber contents on flexural behaviors of milling steel fiber-reinforced ultrahigh-performance concrete: experimental and data-driven Studies. J. Mater. Civ. Eng. 36, 04024152.
Hou, S., Liu, Y., Yang, Q., 2022. Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. J. Rock Mech. Geotech. Eng. 14 (1), 123–143.
Hou, S., Liu, Y., Zhuang, W., Zhang, K., Zhang, R., Yang, Q., 2023. Prediction of shield jamming risk for double-shield TBM tunnels based on numerical samples and random forest classifier. Acta. Geotech. 18 (1), 495–517.
Li, D., Li, H., Ren, L., Song, G., 2006. Strain transferring analysis of fiber Bragg grating sensors. Opt. Eng. 45 (2), 24402–24408.
Li, H.-J., Zhang, C.-X., Wu, H.-Y., Zhu, H.-H., Reddy, N.G., Garg, A., 2022. Monitoring flexure behavior of compacted clay beam using high-resolution distributed fiber optic strain sensors. Geotech. Test J. 45 (3), 627–643.
Li, J., Zhu, H.-H., Wu, B., Hu, L.-L., Liu, X.-F., Shi, B., 2023. Study on actively heated fiber Bragg grating sensing technology for expansive soil moisture considering the influence of cracks. Measurement 218, 113087.
Li, R.-D., Zhang, P., Yin, Z.-Y., Sheil, B., 2024. Enhanced hybrid algorithms for segmentation and reconstruction of granular grains from X-ray micro computed-tomography images. Int. J. Numer. Anal. Methods GeoMech. https://doi.org/10.1002/nag.3832.
Lin, S.-Q., Tan, D.-Y., Yin, J.-H., Li, H., 2021. A Novel Approach to surface strain measurement for cylindrical rock specimens under uniaxial compression using distributed fibre optic sensor technology. Rock Mech. Rock Eng. 54, 6605–6619.
Liu, Y., Bao, Y., 2023. Automatic interpretation of strain distributions measured from distributed fiber optic sensors for crack monitoring. Measurement 211, 112629.
Loh, W.Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev.-Data Mining Knowl. Discov. 1 (1), 14–23.

Ma, T., Liu, K., Su, X., Chen, P., Ranjith, P., Martyushev, D.A., 2024a. Investigation on the anisotropy of meso-mechanical properties of shale rock using micro-indentation. Bull. Eng. Geol. Environ. 83 (1), 29. https://doi.org/10.1007/s10064-023-03510-y.
Ma, T., Liu, J., Fu, J., Qiu, Y., Fan, X., Martyushev, D.A., 2024b. Fully coupled thermo-hydro-mechanical model for wellbore stability analysis in deep gas-bearing unsaturated formations based on thermodynamics. Rock Mech. Rock Eng. 1–32.
Martyushev, D.A., Yang, Y., Kazemzadeh, Y., Wang, D., Li, Y., 2023. Understanding the mechanism of hydraulic fracturing in naturally fractured carbonate reservoirs: microseismic monitoring and well testing. Arabian J. Sci. Eng. 1–14.
Matin, S., Farahzadi, L., Makaremi, S., Chelgani, S.C., Sattari, G., 2018. Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. Appl. Soft Comput. 70, 980–987.
Morgese, M., Ying, Y., Taylor, T., Ansari, F., 2022. Method and theory for conversion of distributed fiber-optic strains to crack opening displacements. J. Eng. Mech. 148 (12), 04022072.
Murtagh, F., 1991. Multilayer perceptrons for classification and regression. Neurocomputing 2 (5–6), 183–197.
Munoz, H., Taheri, A., 2017. Specimen aspect ratio and progressive field strain development of sandstone under uniaxial compression by three-dimensional digital image correlation. J. Rock Mech. Geotech. Eng. 9 (4), 599–610.
Song, Q., Yan, G., Tang, G., Ansari, F., 2021. Robust principal component analysis and support vector machine for detection of microcracks with distributed optical fiber sensors. Mech. Syst. Signal Process. 146, 107019.
Song, Q., Zhang, C., Tang, G., Ansari, F., 2020. Deep learning method for detection of structural microcracks by Brillouin scattering based distributed optical fiber sensors. Smart. Mater. Struct. 29 (7), 075008.
Wang, D.-B., Zhou, F.-J., Li, Y.-P., Yu, B., Martyushev, D., Liu, X.-F., Wang, M., He, C.-M., Han, D.-X., Sun, D.-L., 2022. Numerical simulation of fracture propagation in Russia carbonate reservoirs during refracturing. Petrol. Sci. 19 (6), 2781–2795.
Zhang, W., Li, C., Peng, G., Chen, Y., Zhang, Z., 2018. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. Mech. Syst. Signal Process. 100, 439–453.
Zhang, W., Wu, C., Li, Y., Wang, L., Samui, P., 2021a. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. Georisk 15 (1), 27–40.
Zhang, W., Wu, C., Zhong, H., Li, Y., Wang, L., 2021b. Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. Geosci. Front. 12 (1), 469–477.
Zhang, X., Zhu, H., Jiang, X., Broere, W., 2024. Distributed fiber optic sensors for tunnel monitoring: a state-of-the-art review. J. Rock Mech. Geotech. Eng. https://doi.org/10.1016/j.jrmge.2024.01.008.
Zhao, M., Zhong, S., Fu, X., Tang, B., Pecht, M., 2019. Deep residual shrinkage networks for fault diagnosis. IEEE Trans. Ind. Inf. 16 (7), 4681–4690.
Zhao, S., Tan, D., Lin, S., Yin, Z., Yin, J., 2023a. A deep learning-based approach with anti-noise ability for identification of rock microcracks using distributed fibre optic sensing data. Int. J. Rock Mech. Min. Sci. 170, 105525.
Zhao, S., Wang, F.-Y., Tan, D.-Y., Yang, A.-W., 2024. A deep learning informed-mesoscale cohesive numerical model for investigating the mechanical behavior of shield tunnels with crack damage. Structures 66, 106902.
Zhao, S., Wu, S., Yang, L., Wang, H., 2017. Analysis of secondary roof structure of the working face in Shendong mining area. Geotech. Geol. Eng. 35 (1), 195–202.
Zhao, S., Zhang, G., Zhang, D., Tan, D., Huang, H., 2023b. A hybrid attention deep learning network for refined segmentation of cracks from shield tunnel lining images. J. Rock Mech. Geotech. Eng. 15 (12), 3105–3117.
Zhou, M., Chen, J., Huang, H., Zhang, D., Zhao, S., Shadabfar, M., 2021. Multi-source data driven method for assessing the rock mass quality of a NATM tunnel face via hybrid ensemble learning models. Int. J. Rock Mech. Min. Sci. 147, 104914.
Zhu, X., Chu, J., Wang, K., Wu, S., Yan, W., Chiam, K., 2021. Prediction of rockhead using a hybrid N-XGBoost machine learning framework. J. Rock Mech. Geotech. Eng. 13 (6), 1231–1245.

Dao-Yuan Tan is an Associate Professor in School of Earth Sciences and Engineering, Nanjing University, China, and the awardee of the Excellent Young Scientists Fund (Overseas) by the National Natural Science Foundation of China. He obtained his Ph.D. in Geotechnical Engineering from the Hong Kong Polytechnic University, China, in 2019. His areas of research interest include geohazards management, intelligent monitoring of geotechnical structures and development of smart city infrastructure. He has been the Principal Investigator leading one General Research Fund by the Hong Kong Research Grants Council. He has published more than 30 research papers in reputable SCI journals. He currently serves as the nominated committee member of TC220, Field Monitoring in Geomechanics, ISSMGE. He was awarded the Ringo Yu Prize for Best PhD Thesis in Geotechnical Studies from The Hong Kong Institution of Engineers (HKIE) in 2020 and Fugro Prize (1st Runner-up) in 2023.