Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.081660

ARTICLE

Machine Learning-Based Prediction of Rock Fracture under Uniaxial Loading Using Infrared Radiation

Naseer Muhammad Khan[1,2], Liqiang Ma[3,*], Majid Khan[4], Sajjad Hussain[5], Waleed Inqiad[6], Tariq Feroze[2] and Danial Jahed Armaghani[7,*]

[1]Department of Mining and Explosives Engineering, Missouri University of Science and Technology, 324 McNutt Hall, 1400N. Bishop Ave., Rolla, MO, USA
[2]Sustainable Advanced Geomechanical Engineering, National University of Sciences and Technology, Risalpur, Pakistan
[3]Xinjiang Key Laboratory of Coal-Bearing Resources Exploration and Exploitation, Xinjiang Institute of Engineering, Urumqi, China
[4]School of Civil and Resources Engineering, University of Science and Technology Beijing, China
[5]Department of Mining Engineering, University of Engineering & Technology, Peshawar, Pakistan
[6]Department of Civil Engineering, College of Engineering & Physical Sciences, Aston University, Aston St, Birmingham, UK
[7]School of Civil and Environmental Engineering, University of Technology Sydney (UTS), Ultimo, NSW, Australia

*Corresponding Authors: Liqiang Ma. Email: ckma@cumt.edu.cn;
Danial Jahed Armaghani. Email: danial.jahedarmaghani@uts.edu.au

Received: 06 March 2026; Accepted: 27 May 2026; Published: 30 June 2026

ABSTRACT: Rock fracture behavior under stress is vital for risk evaluation in underground engineering excavation because the presence of water can significantly increase the extent of cracks and fractures in rock, leading to structural damage. This can result in catastrophic failures, including rock bursts, coal bursts, and water inrush. Hence, reliable prediction of rock damage and fracture processes is still lacking, which, in turn, enables the safe and efficient conduct of engineering projects in rock-mass environments. Thus, this study examines both dry and saturated sandstone samples under loading using Infrared Radiation (IR), Acoustic Emission (AE) monitoring, and Particle Flow Computation (PFC) techniques to effectively evaluate the fracture process in rocks under loading. Additionally, seven different artificial intelligence techniques, such as Gene Expression Programming (GEP), Gradient Boost Regression (GBR), Extreme Gradient Boosting (XGB), Adaptive Boosting (AdaBoost), Light Gradient Boosting Machine (LGBM), Categorical Boosting (CatBoost), were employed along with Explainable Machine Learning (XML) to predict the rock damage and fracture process. These models helped in the development of early warning signals to prevent catastrophic accidents. Both the experimental and simulation results have shown that the fracture density measured in terms of PFC and AE cumulative energy is significant in the saturated conditions compared to the dry conditions. Also, stress levels of 0.72 and 0.75 were found to be the warning signs in both dry and saturated conditions, based on the IR index (Average Infrared Radiation Temperature, AIRT) and AE characteristics. The comparison showed that the prediction accuracy of the XGB algorithm was the highest, followed by GBR, CatBoost, LGBM, GEP, and AdaBoost. However, GEP expressed its output in the form of an empirical equation owing to its grey-box nature, and thus, the law of fracture estimation in the form of an empirical equation was developed. The XML methods were added in order to enhance the interpretability of the high-performing, but black-box, XGB model. Such methods, along with a user-friendly Graphical User Interface (GUI), improved the model transparency and facilitated the integration of data-driven decision-making. XML and GUI tools may be instrumental in improving the safety measures adopted in coal mines and tunnels by reducing the risks and increasing operational safety.

KEYWORDS: Rock fracture prediction; acoustic emission; infrared radiation; particle flow computation; water contents

1 Introduction

Rock failure is an important and complex phenomenon in rock engineering; its understanding is important to avoid catastrophic failure and ensure safety and stability in underground engineering projects, including tunnels, mines, and caverns. Primarily, rock failure is caused by fracture initiation and propagation, which affects the geological integrity and causes significant damage to the rock mass under loading conditions. The rock fracture mechanism is affected by various factors, including rock composition, geological conditions, external loads, and environmental factors (water and dry). Moreover, the extent of rock fractures increases due to the presence of water, which raises pore water pressure, deteriorates rock integrity (by weakening minerals bond through chemical reaction), and accelerates crack propagation due to hydraulic, freeze, and thaw effects. This promotes the rock failure process and potentially affects the safety and efficiency of engineering projects in deep rock mass environments [1–4]. Therefore, evaluating rock fracture processes in water-bearing rock is crucial for improving both safety and execution of projects.

Researchers have comprehensively investigated the water effects on rock fracture processes and highlighted the significant findings related to rocks in both water-saturated and dry conditions. Khan et al. [5] found that the mechanical properties of rock reduce with increasing moisture. Gong et al. [6] evaluated the behavior of saturated coal's dynamic fracture. During the crack propagation process, saturated coal samples were found to be more vulnerable to fracture initiation at early stages, whereas cracked coal samples were resistant to crack propagation at later stages. Zhou et al. [7] utilized dynamic notch semicircular bending tests to evaluate saturated and dry specimens and reported a significant reduction in the fracture propagation rate in saturated samples compared with dry samples. However, there was a greater dependent increase in the fracture propagation rate for the saturated samples than for the dry samples. Hao et al. [8] findings show that moisture has a significant effect on subcritical crack development and the durability of fractures during the double torsion test of rock. Water promotes crack initiation and reduces fracture toughness, significantly impacting the propagation of subcritical cracks and the resistance to fracture. Ma et al. [9] found that clay minerals present in water-saturated rocks undergo dissolution, resulting in the expansion of pore structures and an increase in the number of interconnected pores. Furthermore, the presence of water within the rock degrades the integrity of the surrounding structure near the primary fissure, causing splitting fractures to readily propagate over the joint surface. Ma et al. [9] stated that when water-bearing rock is subjected to uniaxial stress, fracture propagation becomes more irregular. This results in the formation of additional secondary cracks within the rock samples and a notable reduction in the strength of the rock. In the context of rock mechanics, the energy theory poses that both dissipative and elastic strain energies accumulate during the fracture process. Furthermore, the rock failure can be caused by changes in these energies subjected to loading of rock in dry and wet conditions, while it can be determined by the critical stress during sudden structural failure [10–12]. The above research indicates that identifying rock failure from an energy perspective is feasible and necessary. Dissipative strain energy primarily drives the initiation and propagation of microcracks within rock masses, while the accumulation of elastic strain energy is implicated in the catastrophic failure of rock engineering structures within these masses [13–16]. During the loading of rock, the conversion of mechanical and strain energy can lead to internal movement and molecular collisions, generating heat within the rock mass. This uneven heat, as it transfers from the interior to the surface, causes temperature variations across the rock's surface. Therefore, a significant relationship exists between changes in surface temperature and the damage characteristics of the rock mass. Current monitoring techniques in mining operations utilize acoustic emission (AE) and infrared radiation (IR) temperature monitoring to assess rock mass internal and external damage.

Researchers have explored the application of AE and IR technologies to predict rock failure precursors, aiming to provide early warnings and improve the monitoring of engineering projects under various load and water conditions. Cai et al. [17] analyzed fracture development during rock loading using AE. Du et al. [18] compared AE waveform signals in shear and tensile tests, finding that shear has longer wavelengths, while lower signal frequencies characterize tensile. Yang et al. [19] investigated the deformation and strength properties of red sandstone, noting shear failure as the predominant mode in triaxial compression tests. AE technology has demonstrated precision in identifying fracture propagation. Through AE detection, the gradual process of rock failure, encompassing crack initiation and growth along with energy release, can be effectively studied [20]. The loading of rock at deforming and failure conditions can be carefully monitored, and evaluations of the indices can be done using infrared (IR) radiation. Specifically, the average infrared radiation temperature index (AIRT) has been developed to predict the development of the rock fracture processes.

This concept was first suggested by Wu and Wang [21]; later, Ma et al. [22], and Liu et al. [23] developed the methodology and proposed the infrared radiation variance (IRV) and the variance of differential infrared image temperature (VDIIT) as measures of crack damage and rock-failure precursors. Further, Shen et al. [24] suggested the top infrared radiation temperature (MIRT), whereas Cao et al. [25] studied the loading and unloading response ratio depending on the IR properties, such as the AIRT and the infrared radiation temperature rate (IRTR) during rock-sample experiments. Of all the indices, AIRT has proved to be the most used index in assessing the process of rock failures as well as the development of precursors. In this regard, Lin et al. [26] obtained a correlation between the AIRT attributes of a sandstone and the porosity and strength of the sandstone.

According to He et al. [16], the AIRT properties of various rocks during freeze-thaw weather cycles were measured, and it was observed that the AIRT of sandstone, marble, and granite decreases with an increase in the rate of applied stress. Simultaneously, Cai et al. [27] reported that the IRR temperature with axial stress on dry rock is lower in comparison with the temperature in saturated rock, and also found that loading rates and water saturation increase the number of fissures that occur at the rock surface. Yuan et al. [28] studied sandstone through triaxial rock-burst experiments, which were supplemented with IR technology; they reported a rapid increase in the mean infrared temperature difference and the mean infrared standard deviation before the occurrence of rock-burst events. Moreover, Tang et al. [29] found a strong effect of water on the IR properties of compression experiments on coal and rock, and lastly [30–33] suggested a predictive approach of damage in the water-bearing sandstone rock using the AIRT index. Research shows that combining IR and acoustic emission (AE) can be a better choice to significantly predict rock failure by monitoring the energy evolution and cracking process. The IR predicts the surface characteristics while AE observes the elastic wave during the internal cracks of the rock underloading. Therefore, IR and AE can effectively monitor the stages of rock failure and give strong signals for the prediction of rock failure precursors.

The literature demonstrates extensive application of IR thermography and AE techniques in rock engineering for monitoring failure processes and identifying precursors. However, their reliability remains a subject of debate, particularly under varying hydro-mechanical conditions, as groundwater presence significantly alters IR radiation characteristics and AE signal propagation due to changes in saturation and mineralogical interactions. The Artificial Intelligence (AI) techniques have proved their reliability in the prediction of various components of the engineering projects, such as rock mass classification, landslide, rock strength, etc. However, AI techniques applied to robust failure prediction and precursor development under coupled environmental conditions remain insufficiently explored. Thus, this study addresses these limitations through a multifaceted framework that integrates IR, AE, and advanced AI approaches within

a physically informed context. Sandstone specimens were systematically tested under dry and saturated conditions to explicitly quantify the influence of water on precursor evolution. A comprehensive suite of AI algorithms, including Gene Expression Programming (GEP), Gradient Boost Regression (GBR), Extreme Gradient Boosting (XGB), Adaptive Boosting (AdaBoost) Regression, Light Gradient Boosting Machine (LGBM), Categorical Boosting (CatBoost), and Explainable Machine Learning (XML), was implemented and rigorously compared using a unified dataset. Notably, GEP was employed to derive an explicit predictive formulation, enhancing model transparency and interpretability. Furthermore, PFC simulations were incorporated to establish a quantitative link between data-driven predictions and mesoscale fracture mechanisms. The novelty of this work is therefore multi-fold, arising from the integration of hydro-condition sensitivity, multi-model benchmarking, interpretable AI-based formulation, and physics-based numerical validation, collectively advancing the reliability of rock failure prediction in complex engineering environments.

2 Materials and Methods

2.1 Samples Preparation

The sandstone samples used in this study were obtained in the form of boulders from an active working face of a coal mine in Shandong, China. The representative specimens were made into cubic geometries of each size 50 mm × 50 mm × 50 mm and tested according to ASTM guidelines. The selected sandstone primarily comprises clay minerals, dolomite, calcite, siderite, feldspar, quartz, and plagioclase, as illustrated in Fig. 1. This research involved measuring the ultrasonic wave velocity in 20 rock samples using a U510 nonmetallic ultrasonic detector. An ultrasonic probe was affixed to each sandstone sample with the aid of a coupling medium. Samples that exhibited visible macro damage, such as surface microcracks, and demonstrated an ultrasonic wave velocity variation exceeding 10% were excluded from further analysis (this limit was taken as the main homogeneity parameter to guarantee that the sandstone samples chosen had the same internal features and density prior to testing). Consequently, 12 rock samples were selected for additional studies, which included AE and IR detection.

In order to test the samples, 12 rock samples were placed in a furnace and dried at 105°C for 48 h to remove all moisture. The 12 dry samples were then split into two individual groups, which were now termed as Group K and Group L. The six samples of Group K were packed and kept under plastic covers to maintain the dry form, and those of Group L were put in water tanks to be saturated over a period of one month (under constant water temperature of 25°C and under 101.3 kPa atmospheric pressure) to ensure uniform moisture distribution across the specimens. The weight of the specimens was measured with special care, and noted prior to and after immersion; the obtained results are shown in Table 1. The table is a summary of the statistical data on the weight and water content of the rock specimens of Group K and Group L before and after immersion.

2.2 Experimental Equipment

2.2.1 Loading Units

Fig. 2 illustrates the uniaxial compression tests conducted on the specimens. The loading equipment consisted of a universal testing machine (model MTS C64.106) with an electrohydraulic servo system, which was capable of supporting a maximum load of 1000 kN. This system offers a high level of accuracy, with load and displacement measurements reported to be precise within ±0.5% of the actual values. The system of data acquisition was set to sample at 10 Hz, thus providing a granular and accurate recording of experimental conditions. All the experiments were conducted in a controlled laboratory with an ambient temperature of (22 ± 1°C) and a relative humidity of (45 ± 5%), minimizing the effects of environmental variations on

infrared radiation (IR) observations. A displacement-controlled loading protocol was used to measure all measurements, with a uniform loading speed of 0.1 mm/min.

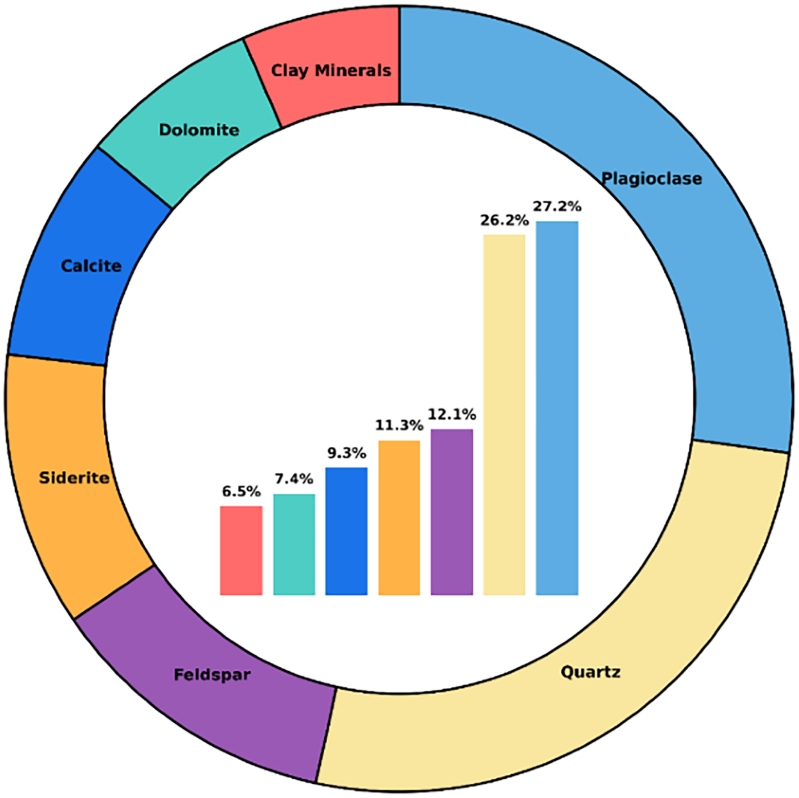

Figure 1: Mineral composition of sandstone.

Table 1: Rock samples weight and moisture content before and after saturation.

Sample No.	W_d (g)	W_S (g)	$W_S - W_d$ (g)	M_c (%)
K_1	540.6	–	540.6	0
K_2	555.4	–	555.4	0
K_3	558.6	–	558.6	0
K_4	563.4	–	563.4	0
K_5	543.2	–	543.2	0
K_6	538.9	–	538.9	0
L_1	536.7	556.9	20.2	3.77
L_2	541.6	559.7	18.1	3.34
L_3	551.2	569.8	18.6	3.38
L_4	547.3	568.6	21.3	3.89
L_5	545.6	565.8	20.2	3.70
L_6	556.7	575.8	19.1	3.43

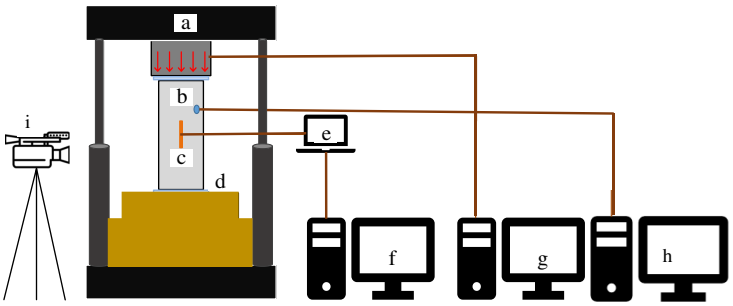

Figure 2: Experimental scheme for fracture prediction; (**a**) loading unit; (**b**) sandstone sample; (**c**) strain gauge; (**d**) plastic film; (**e**) strain gauge; (**f**) data acquisition system; (**g**) load control system; (**h**) AE monitoring system; and (**i**) FLIR A615 infrared thermal imager.

2.2.2 AE Setup

The AE monitoring system that was utilized in the research was designed by PAC and was referred to as the PCI-2 model. As the rock damage was being recorded in real-time, the associated acoustic-emission events were also represented in real-time. The PCI-2 system has a number of technical features: its acoustic emission channels have low noise, a dynamic range of over 85 dB, and its bandwidth reaches 3 MHz to 1 kHz. It also has an integrated 10-channel filter configuration with four high-pass filters and six low-pass filters having adjustable ranges, an 18-bit ADC with 10 kHz sampling rate, and a HIT processor with digital signal processing circuitry that makes it very precise and reliable. The sound emission positioning parameters were set by opening the acoustic emission positioning setting interface and setting the positioning type to 3D positioning; testing the wave speed with an AST (automatic sensor test) and inputting the wave rate into the corresponding position; configuring the event definition value to 128, the event blocking value to 256, and the over positioning value to 12.8; leaving the remaining parameters at their default values; and opening the positioning display to set the sample size and the arrangement of the acoustic emission sensors.

2.2.3 IR Camera

The infrared radiation monitoring device used in experiments had a VarioCAM HD head 880, an uncooled thermal imaging camera produced by the German company InfraTec. This advanced instrument contains a state-of-the-art 1240 × 768 pixel HD-type micro-thermal focal-plane detector with a thermal subcutaneous of less than 0.02°C. The device has a high-quality infrared optical lens (f/1.0) which works within a 7.5 to 14 m spectral range. Detector calibration is normally carried out with temperatures between −40°C and +1200°C, and this can be further extended to 2000°C; the camera provides a spatial resolution of about 17 μm. In turn, the VarioCAM HD head 880 transfers a number of significant benefits: the user-friendly interface does not limit the location of the observation, which makes it the perfect choice in terms of obtaining the infrared signature of rock surfaces in real-time. In addition, it produces very precise quantitative information and vivid images and protects the substrate by non-contact working.

2.2.4 High-Speed Camera

The Manta G-507B, a German high-speed monitoring system with a resolution of 1920 × 1200 pixels and a frame rate of 1300 fps, was used for this study.

3 Research Methodology

In this study, AI-Performance evolution-based prediction models for rock fracture under uniaxial loading were developed and proposed. The following methodology steps were adopted (Fig. 3):

- *Experimental data acquisition:* The stress-strain, AE, and IR data are acquired from experiments.
- *Fracture data:* This data is acquired from the PFC simulation.
- *Data wrangling:* The data is clean from outlier, transformed into a usable form that is used for ML analysis.
- *Splitting data:* The data is divided into a training set of 70% of the observations, and the rest of the 30% is a test set.
- *ML algorithm:* ML algorithms were used, including GEP, XGB, LGBM, and AdaBoost, etc.
- *Performance evaluation:* Use different performance indicators for the best algorithm selection in terms of accuracy.
- *Feature impact assessment:* To inform the effect of the input variables on the modeled outcomes and to facilitate evidence-based decision making in the industrial sector, we adopted explainable machine-learning tools, including Shapley additive explanations (SHAP) and individual conditional expectation (ICE) plots.
- *Graphical user interface (GUI):* The GUI is an interactive interface that has been developed, enabling stakeholders to perform predictive studies of rock fracture phenomena under uniaxial loading conditions in an efficient and streamlined way.

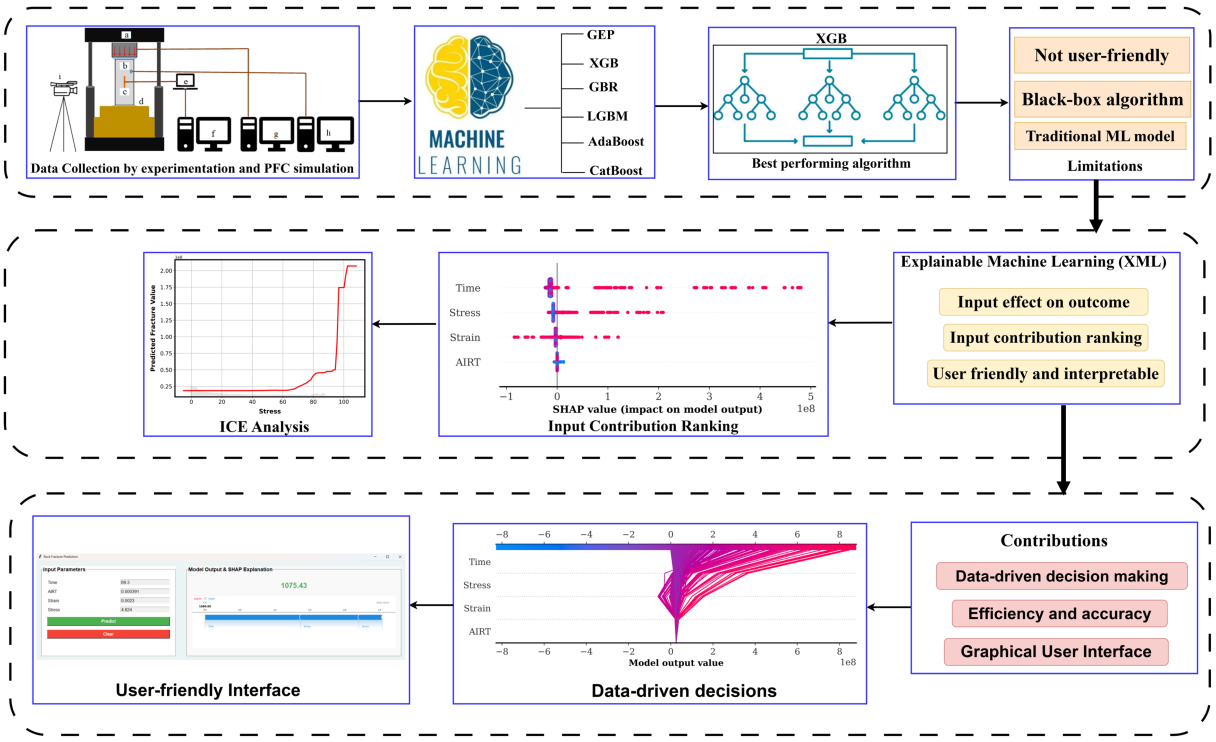

Figure 3: Research methodology followed for developing the ML models.

3.1 Evolutionary Algorithms

Genetic programming (GP) was introduced in 1994 by Poli and Koza [34]. The basic concept of GP is the combination of the theory of natural selection and the concept of genetics [35]. The GP uses the

predetermined lengths of the binary chains analyzed by predetermined lengths, and this feature makes it a useful tool for modeling and mathematical problems. GP develops mathematical models by combining natural selection and regression techniques, which helps the predictive model to converge to solve the smallest population accuracy by effectively removing the sets of the least accuracy of the population [36].

Gene Expression Programming (GEP)

This is a generalized and advanced implementation of the paradigm of Genetic Programming based upon solid principles of genetic population algorithms [37]. It goes beyond the classical GP model by making use of chromosomes with a fixed length. GEP uses a library of genes that are specifically designed to implement a set of nonlinear encodings of individuals (in the form of linear strings of various sizes and structures) [38]. The expression trees consist of variables, constants, and arithmetic operators, and each ET is a subpart of the whole code. GEP combines an assembly of ET configurations to synthesize a solution to solve a given problem. Fig. 4 shows a schematic representation of the ET structure, genetic operations of mutation and crossover, and a flowchart representing the GEP methodology. The mutation and crossover operations are performed over the ETs to produce new individuals, as shown in Fig. 4. Crossover refers to the process that combines two parents to derive offspring, while mutation changes an existing parent. The evolution of a program starts with generating a population of random individuals called chromosomes [39]. Each chromosome is a potential solution and is measured for its accuracy according to a given fitness function. Chromosomes whose accuracy value is at an acceptable threshold are selected to compose the next generation, and those that have too low an accuracy value are removed. This selection process allows the overall accuracy of the algorithm to improve over generations [36]. Those chromosomes with acceptable accuracy form the parents, which enter into the offspring generation through mutation, crossover, and so on. Mutation introduces random changes, whereas recombination involves exchanging parts between parent chromosomes to create new solutions. This cycle of selection, mutation, and recombination is repeated over many generations, resulting in successive populations with progressively better accuracy [40].

3.2 Ensemble Algorithms

3.2.1 Gradient Boosting Regression (GBR)

GBR boosts weak learners in a systematic manner using repetitive learning to build one strong learner, as described in reference [41]. In contrast to bagging techniques, which build base frameworks in a static way, boosting algorithms build base frameworks sequentially. This process of refinement can produce several models, which focus on the analysis of complex learning signals, and thus boost the strength of the predictive system. The preliminary boosting system is based on the training data; the improvements of the model are made by further iterations. The structure uses the leftovers of the previous tree to inform remedial changes. Boosting dates back to the reply of Schapire to the conceptual question of Kearns [42]: "Can the group of weak learners substitute the dichotomy strong and weak learners?". The resulting frameworks in this regard have performance that exceeds random approximations. In his work, Schapire explained that with numerous weak learners, a single strong system can be formed in order to attain higher performance [43]. In boosting, the training process focuses on those observations that were misclassified or erroneously assessed during previous rounds and puts greater weight on them. This plan makes sure that every new model corrects the errors of the former. On the other hand, in bagging, models are only trained on randomly chosen data, and the main focus is to reduce variance and not to directly correct the previous mistakes [44].

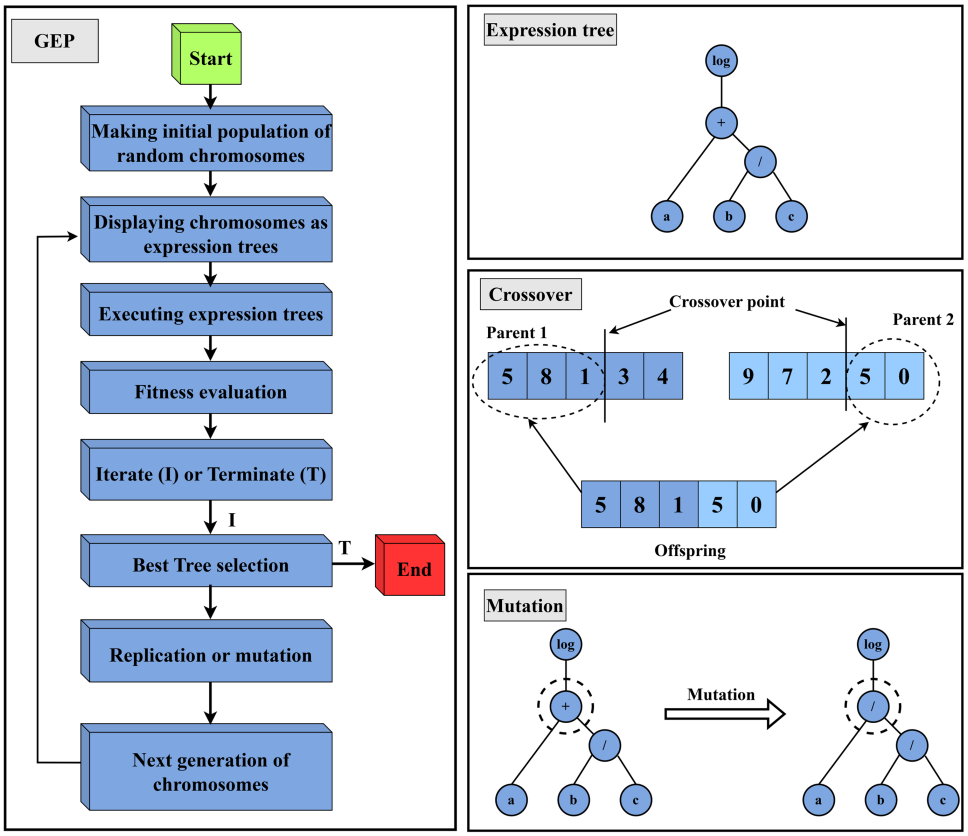

Figure 4: Flowchart of the expression tree, the underlying genetic processes, and the technique of gene expression programming [39].

3.2.2 Extreme Gradient Boosting (XGB)

Tianqi Chen and Carlos Guestrin built the XGB library, an efficient ML framework that is both portable and scalable. XGB is an improved variant of the traditional gradient-boosted decision tree (GBDT) algorithm [45]. In XGB, the decision tree serves as the core component of the model. The decision trees used in XGB are referred to as "weak learners," which play a crucial role in the algorithm's effectiveness. The XGB algorithm consists of a sequence of decision trees arranged in a sequential manner. During the construction of each decision tree, the errors from the previous tree's predictions (as reflected in the objective function) are considered along with the prediction results of the current tree [5,45]. Each decision tree is generated via the complete dataset. The XGB weak learners are integrated with strong learners to enhance performance. However, XGB improves predictive power by incorporating additional regularization terms to mitigate the overfitting and the complexity of the model [46].

3.2.3 AdaBoost

In 1995, Freund and Schapire proposed an ensemble machine learning algorithm called adaptive boosting (AdaBoost) [47]. In comparison to other machine learning algorithms, AdaBoost is highly efficient in learning, improves model performance, reduces overfitting, and provides effective learning through continuous training [48]. Consequently, it is frequently used to combine weak learners, including decision trees, ANNs, random forests, support vector machines, and deep learning models, creating a more robust

model and improving prediction accuracy. The detailed flowchart of the AdaBoost algorithm is available in [48].

3.2.4 Light Gradient Boosting Machine (LGBM)

A novel methodology has been developed to tackle the scalability and efficiency challenges associated with the model by employing LGBM, an advanced variant of gradient boosted decision trees (GBDTs) [48]. Traditional GBDT methods necessitate the evaluation of all data samples to ascertain the information gain for potential split points within a decision tree. In contrast, LGBM introduces two innovative strategies: gradient-based one-sided sampling (GOSS) and exclusive feature bundling (EFB) to enhance the conventional GBDT approach. In calculating the information gain for identifying split points, GOSS only considers those samples with large gradients. In addition, this EFB technology resolves the NP-hard difficulty in the integration of the complementary characteristics, which play little role in deciding the split point. EFB does this through clustering mutually exclusive features, i.e., features that rarely take nonzero values at the same time, which improves model efficiency [48].

3.2.5 CatBoost

It was introduced by Yandex in 2017 and is based on gradient boosting, which is specifically designed to handle categorical data. CatBoost converts categorical data into a numerical value to prevent overfitting and uses random permutations to effectively manage categorical features [49]. CatBoost offers an unbiased estimate of gradients and enhances model resilience by training different base learners with numerous permutations, thereby decreasing gradient bias [49]. The construction process of CatBoost shows oblivious trees as base learners, where the criteria for evaluating each node at each layer remain the same. So far, oblivious trees are relatively simple and effective, improving prediction accuracy during model fitting. CatBoost is well known for its reduced number of hyperparameters, enhanced resilience, and a user-friendly interface [50,51].

4 Analysis of the Results

4.1 Mechanical Properties

Table 2 identifies the detailed statistical results on the rock mechanical properties at dry and saturated conditions. The sandstone samples, as shown in Fig. 5, had different properties depending on their moisture content. The mean peak stress of 107.57 MPa was impressive in these samples in the dry state. On the other hand, at saturation, this value reduced to 83.25 MPa, which translates to a significant decrease of 23.00 per cent. Similarly, the samples of the dry sandstones had an average elastic modulus of 6.19 GPa, which is evidence of their intrinsic resilience. At saturated conditions, the modulus decreased to 5.30 GPa, which is a considerable reduction of 15% change. The above empirical data highlight a strong erosion in the strength of saturated sandstone compared to the dry counterpart. Moreover, they support the earlier studies, which strengthens the longstanding negative impact of water on the integrity of structures and stability of rock structures [52–56].

Table 2: Mechanical properties of representative samples under different water contents.

Rock Index	Water Content (%)	E (GPa)	σ (MPa)
K_1	0	6.55	125.65
K_2	0	6.29	113.23
K_3	0	6.27	110.09
K_4	0	6.19	104.50
K_5	0	5.95	102.80
K_6	0	5.88	89.16
L_1	3.10	5.69	88.96
L_2	2.96	5.68	87.44
L_3	3.11	5.37	83.33
L_4	3.01	4.97	77.63
L_5	3.13	5.12	81.85
L_6	3.12	5.01	80.29

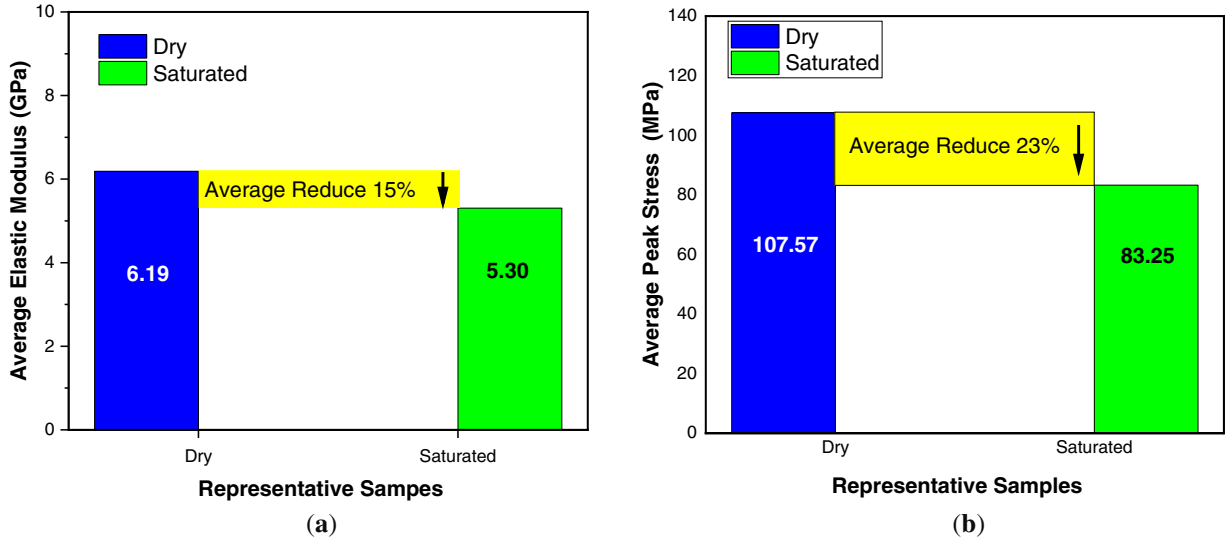

Figure 5: Strength properties of sandstone representative samples under dry and saturated conditions; (**a**) peak stress; (**b**) elastic modulus.

4.2 AE Characteristics

Fig. 6 shows the AE energy and cumulative AE energy of dry and wet rock throughout the loading process over time. During the preloading step, the sample remained intact internally, and the AE energy was minimal. During the fracture propagation phase, the AE energy exhibited a sudden and significant increase, reaching its highest point at the time of failure. The fluctuations in the AE energy are intricately associated with the mechanisms of fracture, damage, and energy dissipation within the sample. Furthermore, the greater the brittleness of rock, such as dry sandstone, the more abrupt the increase in the AE energy at collapse, and the less apparent the precursor. For rocks that are less prone to breaking easily, the significant loss of energy during the period when cracks start to spread before reaching their maximum size results in an increase in the AE energy. This increase in AE energy also provides more detailed information on the events leading to

failure. Fig. 7 presents the normalized progression of the cumulative AE energy values for this specific set of rock samples. When the rocks become less brittle (saturated samples), the increase in AE energy clearly occurs more rapidly during the latter phases of loading.

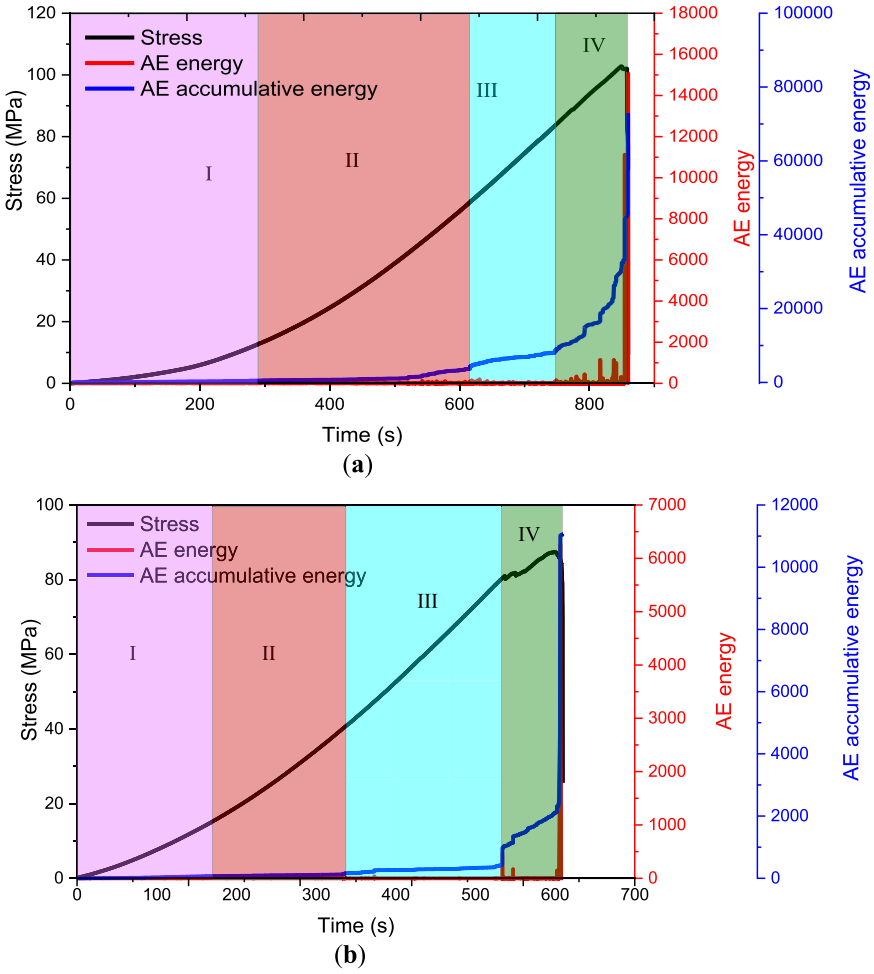

Figure 6: Sandstone AE characteristics under different water conditions: (**a**) dry and (**b**) saturated.

4.3 IR Observation

During uniaxial compressive testing, the machine applies mechanical energy in the form of loading to the representative sandstone sample. This energy is transformed into internal strain energy in the samples when it subjected to loading and deformation. The part of energy dissipated in the form of heats and, which propagated to the rock's surface and is released as infrared radiation. Infrared radiation (IR), extends from 0.76 to 1000 μm, appearing beyond the range of the eye and consequently is invisible. The FLIR A615 infrared thermal imaging system, however, is constructed to detect and observe this otherwise undetectable infrared energy released from the sandstone during the loading process. The FLIR A615 sensory technology employs a combination of signal processing and photoelectric conversion techniques to detect infrared radiation and transform it into quantifiable data. With some post-processing, it outputs a thermal image on infrared that reflects the temperature of the rock sample surface. The thermal image, as given in Fig. 8, shows the distinct temperatures across the sandstone sample, with different colors corresponding to different temperatures. Moreover, the thermal image is converted into pixels, each representing a specific temperature value. Based

on these temperature frames, different indices have been proposed by researchers, as already discussed in Section 1. In this study, the average infrared radiation temperature index (IR) will be used to analyze fracture evolution.

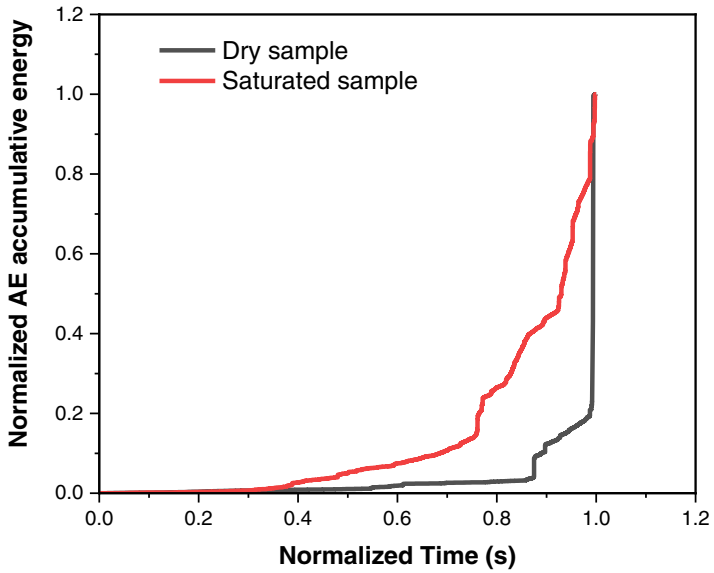

Figure 7: Normalized cumulative energy of AE in different water conditions.

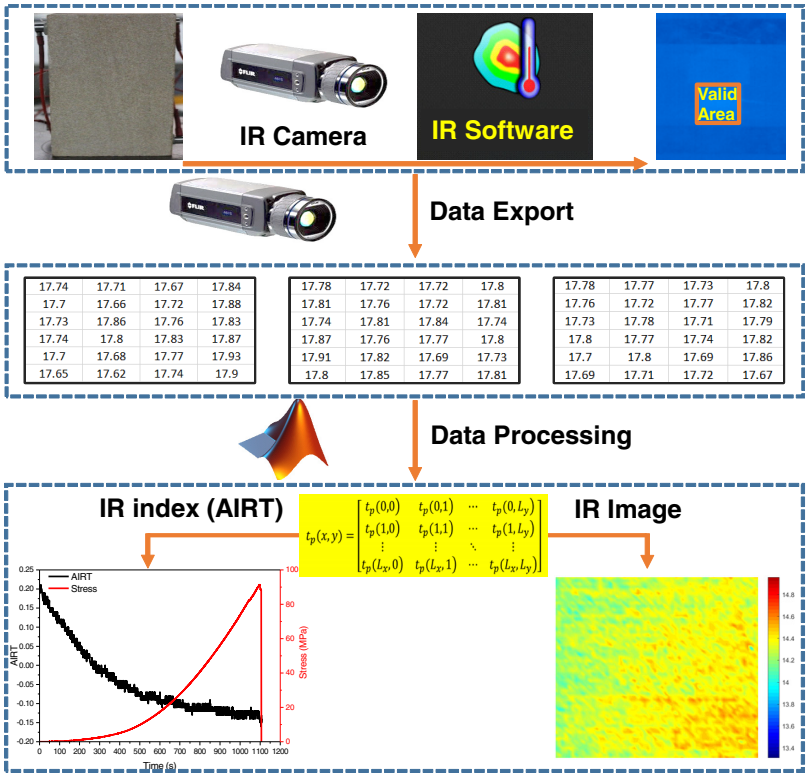

Figure 8: Infrared radiation feature processing process.

The changing sandstone surface temperature through uniaxial compression can be monitored. For this, the infrared geometric sensor that monitors infrared radiation emitted from the sample records the shape of the sandstone sample at different loads to determine the temperature changes at specific locations in the sample during loading. The data processing software embedded in the thermal imager captures valid infrared radiation in the defined rectangular contour. Outputting these data as a time series with the labels T_0, T_1, ..., T_n. Data are analyzed in MATLAB software and converted into a temperature matrix in two dimensions. The relation between sandstone surface temperature t in a point and a given frame number P is expressed in Eq. (1).

$$t_p(x,y) = \begin{bmatrix} t_p(0,0) & t_p(0,1) & \cdots & t_p(0,L_y) \\ t_p(1,0) & t_p(1,1) & \cdots & t_p(1,L_y) \\ \vdots & \vdots & \ddots & \vdots \\ t_p(L_x,0) & t_p(L_x,1) & \cdots & t_p(L_x,L_y) \end{bmatrix} \tag{1}$$

where x represents the horizontal coordinate, and y denotes the vertical coordinate of the matrix profile corresponding to the sandstone sample. The matrix contains values that reflect the temperature readings of the rock at specific locations. These temperature data can be visualized using MATLAB software, where varying colors represent different magnitudes of temperature readings. This visualization enables the creation of a spatial distribution map that illustrates the infrared radiation temperature of the rock sample at a particular moment. For example, Fig. 9 presents infrared thermal images captured at four distinct time points: 200, 400, 600, and 860 s, utilizing the sandstone samples. These images effectively demonstrate the evolution of temperature distribution over time as the samples undergo loading.

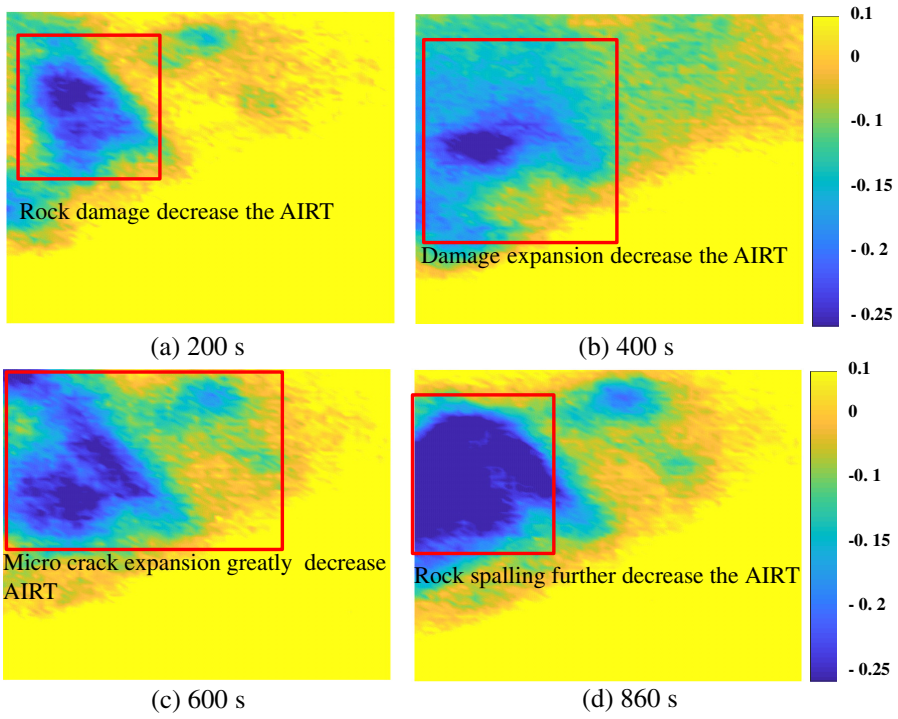

Figure 9: Loading rock IR thermography at different times.

Uniaxial compression was applied to the sandstone samples, and the infrared thermal images showed a significant difference in temperatures on the surfaces of the samples. The average value of the infrared radiation temperature decreased step by step with the rise of the water content in the rock sample; however, the water was not uniformly distributed in the rock sample. Such a reduction in AIRT occurs with increasing damage to the rock, where the heat energy released from the fractures does not propagate outward but instead into dissipative energy as thermal energy is dissipated into the water content. With the increased loading time, the high-temperature regions developed with an outward progress from the sample boundary to the center, and reduced the relatively low-temperature area in the center gradually. So, it was indicative that during the application of uniaxial stress, the surface temperature of the sandstone thickness range increased starting at the edges toward the center.

The sandstone samples also showed a reproducible pattern of fracture growth and failure with applied loads, with cracks initiating at the edges and growing toward the center, as shown in Fig. 9. It was suggested that the statistical dispersion of infrared radiation temperature on the surface of the rock sample would be able to serve as a criterion for the location of internal fractures when it comes to sandstone. In this paper, AIRT was employed to hurdle the statistical variance of surface temperature measurements. We can use Eq. (2) to calculate the AIRT for a specific frame number. For studying the infrared radiation characteristics of sandstone samples at various stages of the loading process, a method exhibited in Fig. 10 was developed.

$$\mathrm{AIRT}\,(p) = \frac{1}{L_x}\frac{1}{L_y}\sum_{x=0}^{L_x}\sum_{y=0}^{L_y} t_p\,(x, y) \tag{2}$$

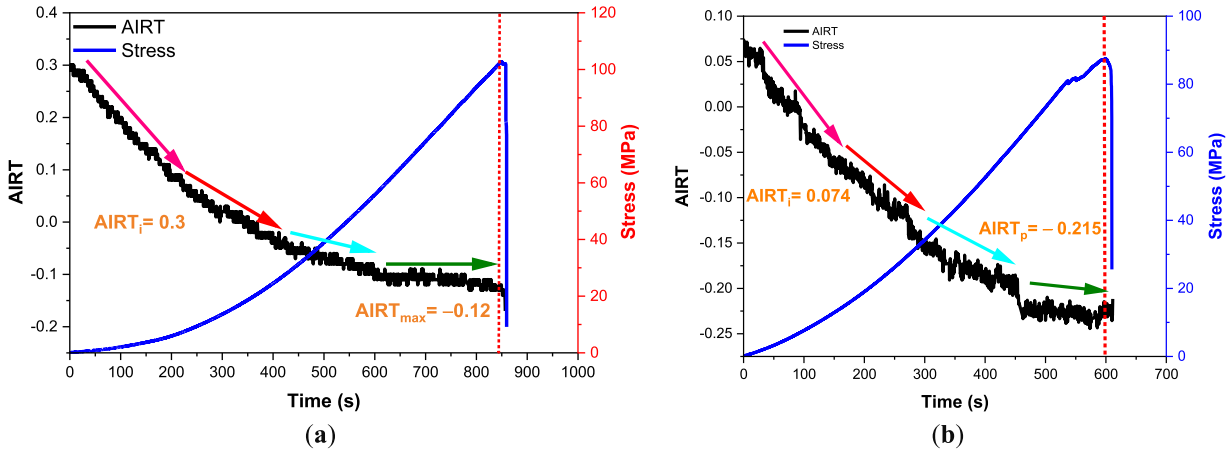

Figure 10: Sandstone AIRT and stress behavior under loading under different conditions: (**a**) dry and (**b**) saturated samples.

The processed data of the IR (AIRT index) for the dry and saturated samples may exhibit significant discrepancies compared to those of other samples within the same group. This may be the consequence of experimental procedures or equipment faults. As a consequence, future research and analysis will focus primarily on dry and saturated samples.

4.4 Average Infrared Radiation Temperature

Literature shows that the specific heat capacity of sandstone and that of water are 800 and 4200 J/kg/°C, respectively [57]. Thus, water is able to hold more heat energy in it before it rises in temperature due to its high heat content as compared to the rock. When more water fills these interstitial spaces, the heating behavior of the rock will vary, especially when it is subjected to loading cases where stress and deformation magnify, and consequently, the emission of Anomalous Infrared Radiation Temperature (AIRT) by the rock surface is reduced. Saturated sandstone had a moisture content 3.055 per cent greater than that of dry sandstone at the same geometric conditions. Since the specific heat capacity of the saturated sandstone is an isotopic measure, it would be concluded that the measure of specific heat capacity is greater than that of dry sandstone. Referring to the definition of specific heat capacity, molecules that have high heat capacity need a lot of energy to increase their temperature by 1°C in a given mass. The infrared thermal image in Fig. 10 shows that the surface of the sandstone sample exhibits a spatial temperature distribution in which both flanks of the sample surface show a tendency to decrease in temperature as the centers move towards the middle of the sample. The saturated sandstone is able to take in significantly more energy at both ends of the specimen due to its high specific heat capacity. This means that the temperature at the center of the specimen will be found to have a sharp fall compared to the dry case. The authors based on thermodynamic considerations the postulation of the larger statistical dispersion of the surface temperature of the saturated sandstone subject to the uniaxial load compared to the dry sandstone. Accordingly, the AIRT of the saturated sandstone surface is higher than that of the dry sandstone [5]. To support this hypothesis, the changes in stress and air permeability of the sandstone samples in dry and saturated conditions over time are shown in Fig. 10.

The AIRT in the example of dry and saturated rock is shown to have a progressive decay process from the point of mechanical loading commencement till the time of catastrophic failure, as shown in Fig. 10. However, the presence of the intrinsic voids in the dry specimens means that most of the external pressure is used up in compressing the intrinsic voids during the initial compaction phase. A resultant consequence of this is that the surface thermal behaviour of the dry rock is relatively inert, which also translates to a slight attenuation in the measured AIRT. On the other hand, AIRT in the saturated rock samples exhibits a steady decrease during the whole duration of the loading. This tendency is arguably explained by the fact that the saturated sandstone contains water-filled pores, which are explained by the prolonged submergence of this rock. The mechanical energy provided by the compression device has a direct effect on the evolution of surface temperature of the sandstone specimen in the compaction interval, which is characterized by a rapid change in AIRT. During the subsequently elastic deformation regime, the velocity of AIRT reduction is relatively small in comparison to the compaction stage, which is typified by stronger gradient lines. This rate reduces further in the following stages of plastic deformation, but the outcome of this is an even steeper slope. In plastic deformation stages, this rate further decreases, resulting in steeper slopes. Just before failure, the rate of AIRT change is very low, displaying a moderately steep slope, which intensifies at the point of failure, indicating rock failure. Notably, the rate of AIRT change in saturated rock is greater than in dry rock. Stress significantly influences infrared radiation, contributing to a reduction in the maximum strength of saturated rocks and shifting the AIRT of the rock from the onset of loading until failure. The variables $AIRT_i$ and $AIRT_p$ represent the AIRT at the initial stress and peak strength, respectively, for the sandstone samples. The processed data show that for sample A1, the $AIRT_i$ is 0.3 and the $AIRT_p$ is −0.12. For sample A_2, the $AIRT_i$ is 0.74, and the $AIRT_p$ is −0.215. In comparison, dry sandstone samples exhibit a markedly greater AIRT than saturated sandstone samples.

The AIRT values of 12 rock samples were compared between the saturated and dry groups (Table 3). The mean $AIRT_m$ for dry samples is 0.22266, whereas the mean $AIRT_m$ for saturated samples is −0.28683, which is 2.28 times lower than its counterpart. Therefore, it is evident that the AIRT features on the surface of

saturated sandstone are more prominent under uniaxial stress than those of dry sandstone. This observation aligns with the findings of thermodynamic theory.

Table 3: AIRT of sandstone samples at their lower and peak strengths.

Samples	$AIRT_i$	$AIRT_m$	Samples	$AIRT_i$	$AIRT_m$
K1	1.205	0.861	L1	0.053	−0.284
K2	0.865	0.281	L2	0.068	−0.221
K3	0.675	0.123	L3	0.150	−0.11
K4	0.754	0.321	L4	0.189	−0.271
K5	0.300	−0.120	L5	0.237	−0.398
K6	0.028	−0.13	L6	0.289	−0.437

4.5 Microcrack Structures and Simulation

Numerical Simulation via PFC

The findings from AIRT and AE reveal important differences in the microstructures of both dry and saturated sandstone samples. This difference is an effect of the inconsistencies between the two types of samples. There is compelling evidence that these variations in microstructure impact external features as well as mechanical properties like IR and AE. However, the application of microelectron microscopy for observation is constrained by notable limitations, since it is only able to capture one, unchanging point. Also, for the purpose of ascertaining the correlation between various IR characteristics and the advancement of fractures in the sandstone samples, it is crucial to assess the dynamic development of fractures in the samples at each stage of loading and deformation. Currently, PFC methodology and the PFC software program have undergone consistent enhancements, establishing themselves as very influential approaches for studying the concepts of rock mechanics. In addition, the grain flow model in PFC enables the transformation of rocks into particle aggregates and simulates the spatiotemporal mechanical properties of rocks by examining particle size and how they interact with different particle sizes. PFC can be used to precisely replicate the production of fractures in rocks at the microscopic level.

The simulation utilized a fictitious object consisting of components K4, K5, L3, and L4. Four simulation samples, labeled X4, X5, X3, and X4, were used to create the model. The dimensions of the numerical model were 50 mm × 50 mm, and the particle sizes varied from 0.80 mm (largest) to 0.40 mm (smallest). The porosity ratio was 0.12, and the particle density was 2.64 g/cm^3. Initially, the numerical model experienced acceleration on all four borders (top, bottom, left, and right) to compress and consolidate the particles into a suitable starting state. This step was essential for improving computational efficiency and reducing the instability and errors that are sometimes present in numerical simulations. After the preloading step, the particles inside the model came into contact with each other. The model's microscopic contact parameters were modified by including mechanical characteristics obtained from uniaxial testing via iterative trial and error techniques [58]. Table 4 presents the specific microscopic features relevant to the interactions between particles. To replicate the uniaxial loading procedure, the model's left and right walls were removed. The top and bottom walls were subsequently subjected to constant velocities in the downward and upward directions, respectively. Fig. 11 shows the particle contact and loading characteristics of the simulated sample.

Following the simulation of sample failure, stress–strain data for the simulated sample were acquired via the FISH function. Fig. 12 presents a comparative analysis of the stress–strain curves of the simulated and experimental samples.

Table 4: Calibrated microparameters of the numerical model.

Sample	Linear Group			Parallel-Bond Group		
	Fric	emod/Gpa	Kratio	pb_ten/MPa	pb_coh/MPa	pb_fa/°
Dry	0.7	11.70	2.5	58.08	45.77	27
Saturated	0.8	11.11	2.5	51.00	40.90	27

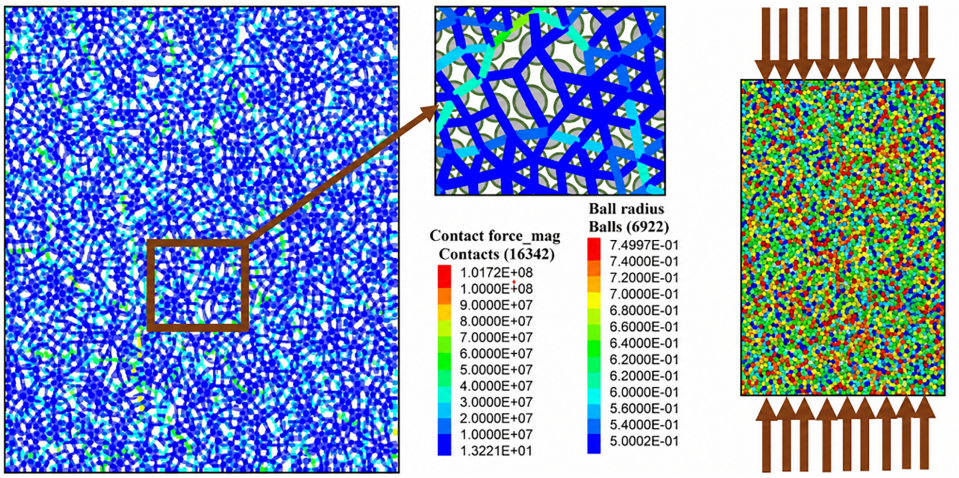

Figure 11: Numerical model particle contact diagram.

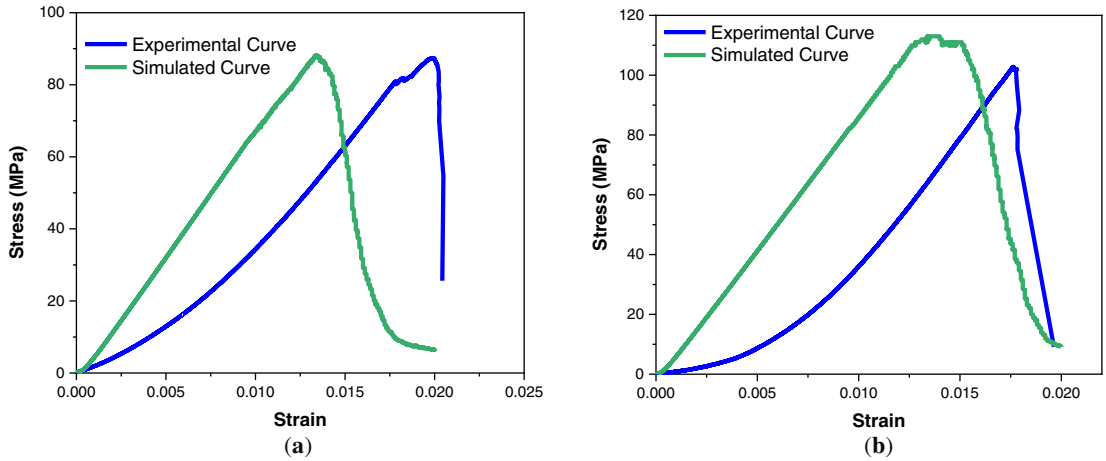

Figure 12: Sandstone characteristics in the experiments and simulations: (**a**) wet and (**b**) dry samples.

As shown by the findings in Fig. 12, the PFC simulation program successfully reproduces the behavior of uniaxial compression in the samples under a range of moisture conditions. The values of peak stress along with the elastic modulus obtained from the numerical model closely correspond to the experimental results. However, there is a constraint in adequately representing the early compaction of particles, which results in the deviation of the simulated sample's compaction from the actual data. The PFC simulation software uses the built-in function dfn.fracture.num to determine the quantity of fractures inside the simulated sample.

This function allows for the examination of how fractures form internally when a sample is subjected to different stress conditions prior to the onset of failure under uniaxial compression.

As shown in Fig. 13, internal fractures in the sample progress in a pattern that is comparable to that of its stress–strain curve. This progression can be broken down into three distinct stages: the no-slope stage, which indicates that the fracture curve is almost horizontal; the steep-slope stage, which indicates that fracture development and propagation begin to occur gradually; and the extremely steep-slope stage, which indicates that the number of fractures increases significantly, and that fracture propagation accelerates more rapidly than the steep-slope stage does. Research has suggested that when internal fractures in sandstone reach a certain threshold before maximum strength, its internal structural stability is compromised. Hence, this aspect is pivotal in anticipating sandstone failure [58].

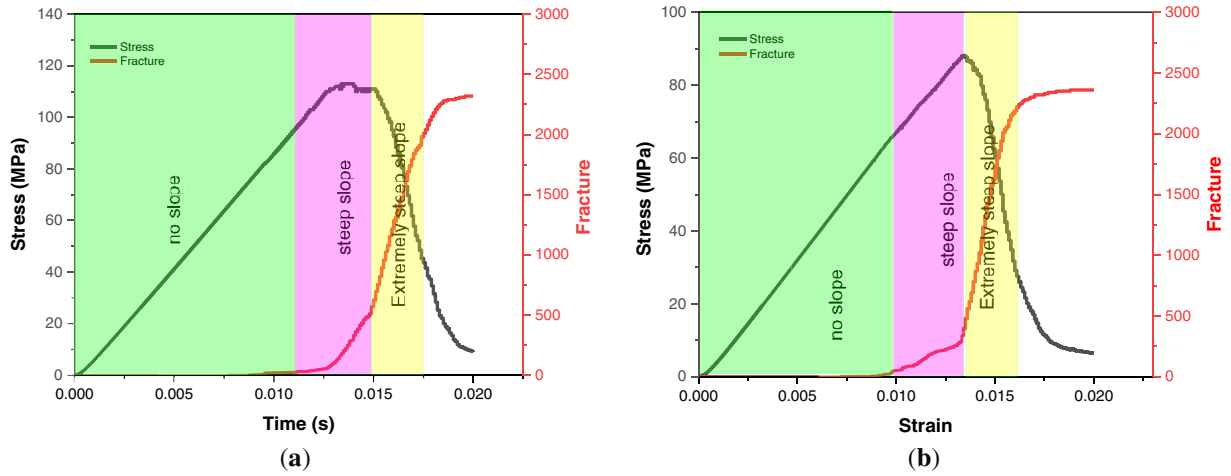

Figure 13: Trend of fractures in dry and saturated sandstone samples under simulation; (**a**) dry; (**b**) saturated.

4.6 Fracture Development with AIRT

4.6.1 AIRT Characteristics in the Fracture Process

Initially, the strains of experimental and simulated sandstone samples were normalized before reaching a peak value for the development of a connection between the observed number of fractures and the change in IR. This normalization process allowed a direct link to be made between the two variables. A visual representation of the normalization approach is presented in Fig. 14. Following normalization of the strain data for both the experimental and simulated samples, the figure shows the relationship between the AIRT and the expansion of internal fractures. The proportion of AIRT is presented along the x-axis, whereas the number of fractures is indicated along the y-axis. Fig. 14 demonstrates that there is an evident negative correlation between AIRT and fracture count: in Stage I and Stage II, the AIRT varies considerably, whereas the fracture density is relatively low, indicating the prevalence of elastic deformation and microcrack early formation. The change rate of AIRT in Stage III, at stress 0.72 (dry) and 0.75 (saturated), is significantly slowed down. This transition is characterized by a reduction in AIRT slope from the steep gradient of the plastic phase (16°–30°) to a moderate gradient (<16°), signaling the onset of the precursor phase. This decrease in slope is accompanied by a simultaneous increase in cumulative AE energy and simulated fracture counts, which are all a multi-modal precursor signal to rock failure. Stage IV, the propagation of the fractures is at its peak, and therefore, the AIRT change is insignificant, in line with the redistribution of stored elastic energy along the prevailing fracture plane instead of further heating of the bulk specimen.

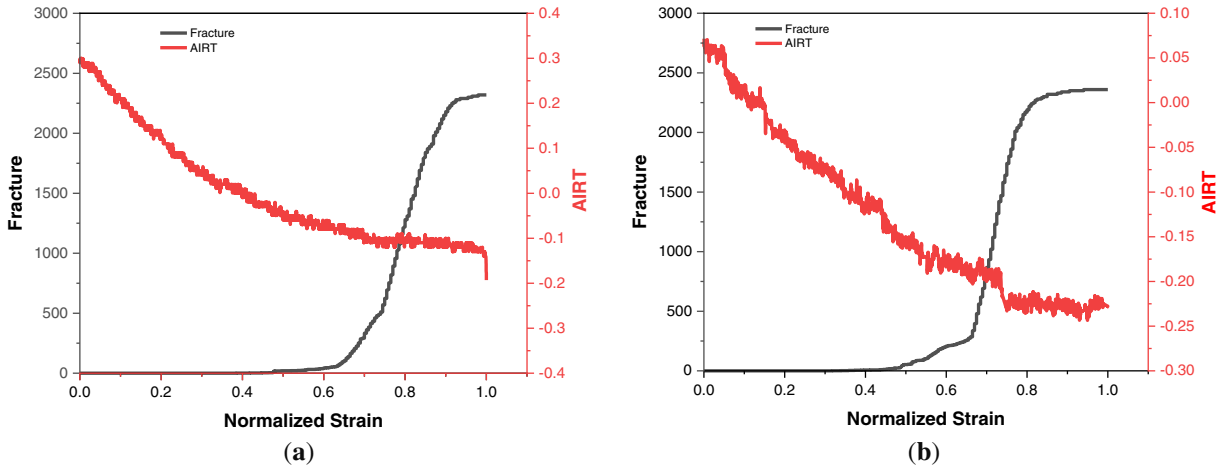

Figure 14: Relationship of AIRT with simulated sample fracture; (**a**) dry sample; (**b**) saturated sample.

4.6.2 Fracture Equation Using Stress–Strain

In this work, a novel attempt is made to estimate fractures from stress and strain values. Previously, Hou et al. [58] established a correlation between fractures and infrared (IR) data via simulated rock samples. However, since simulated rock samples do not perfectly replicate real samples, the previously derived relationship may predict fracture values accurately for simulated scenarios but could yield less accurate results for experimental data. To address this limitation, this study introduces a new approach to estimate fractures as described below in Eq. (3).

$$Fracture \begin{cases} 0.01 \times e^{0.1\times\varepsilon^2+0.09\times\left(\varepsilon^2+\sigma\right)+0.05\times\varepsilon+0.05\times\sigma+0.1\times R}, \varepsilon > 0.017844 \\ e^{0.01\times\varepsilon^2+0.09\times\left(\varepsilon^2+\sigma\right)+0.05\times\varepsilon+0.05\times\sigma+0.1\times R}, \varepsilon < 0.017844 \end{cases} \tag{3}$$

PFC simulation data were obtained to determine the value of fractures under controlled loading conditions across a systematically varied range of stress (i.e., σ) and strain (i.e., ε) conditions. The most physically appropriate structure was found to be a piecewise exponential functional form, which is consistent with the long-standing use of exponential functions in the literature to model damage accumulation and fracture propagation of brittle rock materials in compressive loading. The strain threshold number of 0.017844 in Eq. (3) does not mark a random number, but the physically significant inflection point in the simulated fracture strain curves, which represents the change from the crack initiation stage to the crack propagation stage. The PFC simulation data were then used to perform nonlinear least-squares regression, which produced the coefficients of the exponential expressions, with a high coefficient of determination (R2 = 0.99). The equation was then calibrated to experimental fracture curves by applying an iterative refinement, where the coefficients were varied to allow the equation to predict and observed values as close as possible without imposing any constraints on the physically motivated functional form or the strain threshold. This calibration is similar to usual constitutive model adjustment in rock mechanics, and is not unconstrained data fitting.

In Eq. (3), a variability constant R is added to widen its application to dry and saturated rock. The binary definition of R = 0 dry rock and R = 1 saturated rock is a conscious and pragmatic simplification based on the experimental nature of this experiment, which was designed to compare the two discrete end-member saturation states most frequently documented in rock mechanics experimental work and most indicative of reality in the engineering world. The data available could not be rigorously calibrated to a continuous saturation parameter since the intermediate saturation levels were not tested in the experimental program. It is known that the level of saturation of water can be continuous in both the natural and engineering contexts. But in the modern version of Eq. (3), R is considered as a binary constant (R = 0 in case of dry; R = 1 in case of saturated), which only represents the two end-member cases that are experimentally tested. It is proposed that the extension of Eq. (3) to use a continuous saturation index (e.g., degree of saturation S_w: [0, 1]) is thus a direction to future research, awaiting the provision of experimental fracture data over an intermediate range of saturation.

All experimental samples are analyzed using Eq. (3) regardless of the PFC simulation, and this is used to obtain the fracture curve of rocks during the entire loading process as shown in Fig. 15. Where ε is the strain, σ is the stress, and R is the constant of variability, which is 0 in the case of dry rock and 1 in the case of saturated rock. This conditional expression calculates fracture values in accordance with the current strain regime: at a strain greater than 0.017844, the first exponential expression of the crack propagation stage is used; at a strain less than 0.017844, the second exponential expression of the crack initiation stage is used. Fig. 16 also depicts how the fractures develop in the samples of sandstones in the presence of dry and saturated conditions, reaffirming the suitability of the proposed technique in each water content scenario. The AIRT values for dry sandstone and saturated samples are shown in Fig. 15. This finding reveals that during slow fracture development, the rate of AIRT is greater, whereas during high fracture development, the AIRT rate is lower. This phenomenon is due to various factors related to fracture mechanics and heat transfer. When fractures grow rapidly, as observed under high-stress or dynamic conditions, most of the energy released is in the form of mechanical energy rather than heat. Consequently, less energy is available to heat the surrounding material, resulting in a slower increase in temperature detected by infrared radiation. Conversely, during slow fracture growth, characterized by lower stress and longer durations, more mechanical energy is converted into heat, leading to a greater temperature rise. Additionally, factors such as the material's thermal properties, experimental setup, and cooling methods also influence temperature differences.

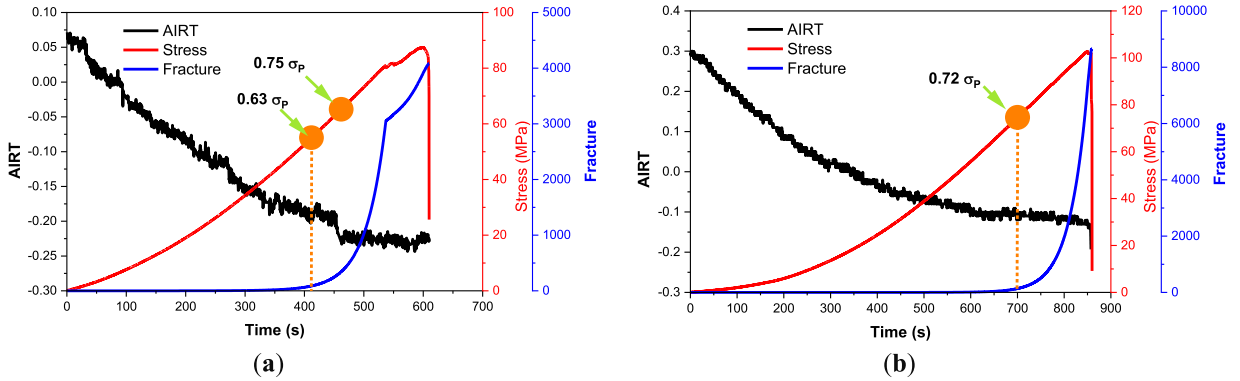

Figure 15: Relationships between AIRT, stress, and fracture in dry and saturated sandstone; (**a**) dry and (**b**) saturated samples.

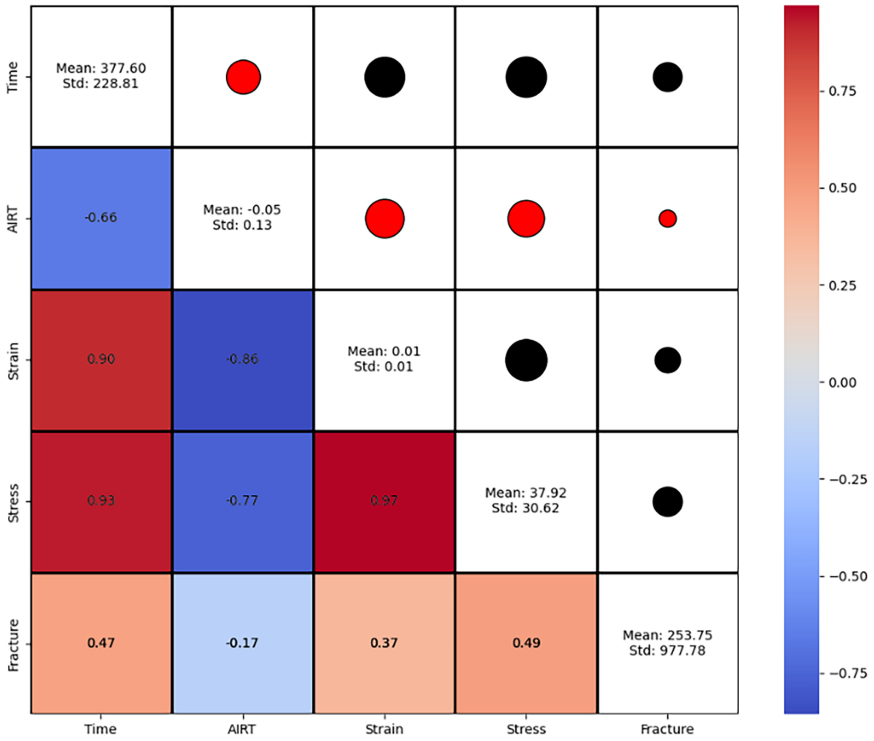

Figure 16: Correlation matrix of the dependent and independent variables.

The AIRT trends observed across various sandstone samples are consistent. Notably, the rate of AIRT within crack closures is initially high, indicating elastic deformation. However, as the stress–strain curve progresses from elastic deformation to plastic deformation, the rate of AIRT decreases gradually. The AIRT values follow an almost linear trajectory, with a noticeable drop at the failure point. Thus, AIRT rates are quicker at the initial stages of the stress-strain curve and slower at the latter part. The linear and higher rate of AIRT was observed during these two phases, which can be attributed to several mechanisms. The closure of microcracks under stress is the first case of this role. In other words, when the external forces act on the rock, the microcracks in the rock's matrix close gradually, which results in a rapid decrease in the air contained inside the rock. This effectively leads to a linear drop off in AIRT, as air trapped isotopically in the rock is either pushed out or squeezed down. In addition, the accelerated AIRT during elastic deformation, characterized by reversible strain in the rock without lasting damage, highlights the ability of the rock to both capture and release air without incurring considerable structural damage. The rock porosity and permeability also affect the AIRT rate. Rocks with increased porosity and connected pore flow networks generally show higher AIRT rates due to increased air release during deformation. Overall, the tangible and amplified AIRT rates associated with these linear trends are beneficial for understanding the mechanical structure and internal composition of rock, and how it responds to external stress and deformation. The shear failure of sandstone after elastic deformation had an initial stress level of stress equal to 0.75 at dry rock as well as 0.63 for wet rock (Fig. 17) and was significantly increasing after elastic deformation, which was defined as an early precursor to rock failure.

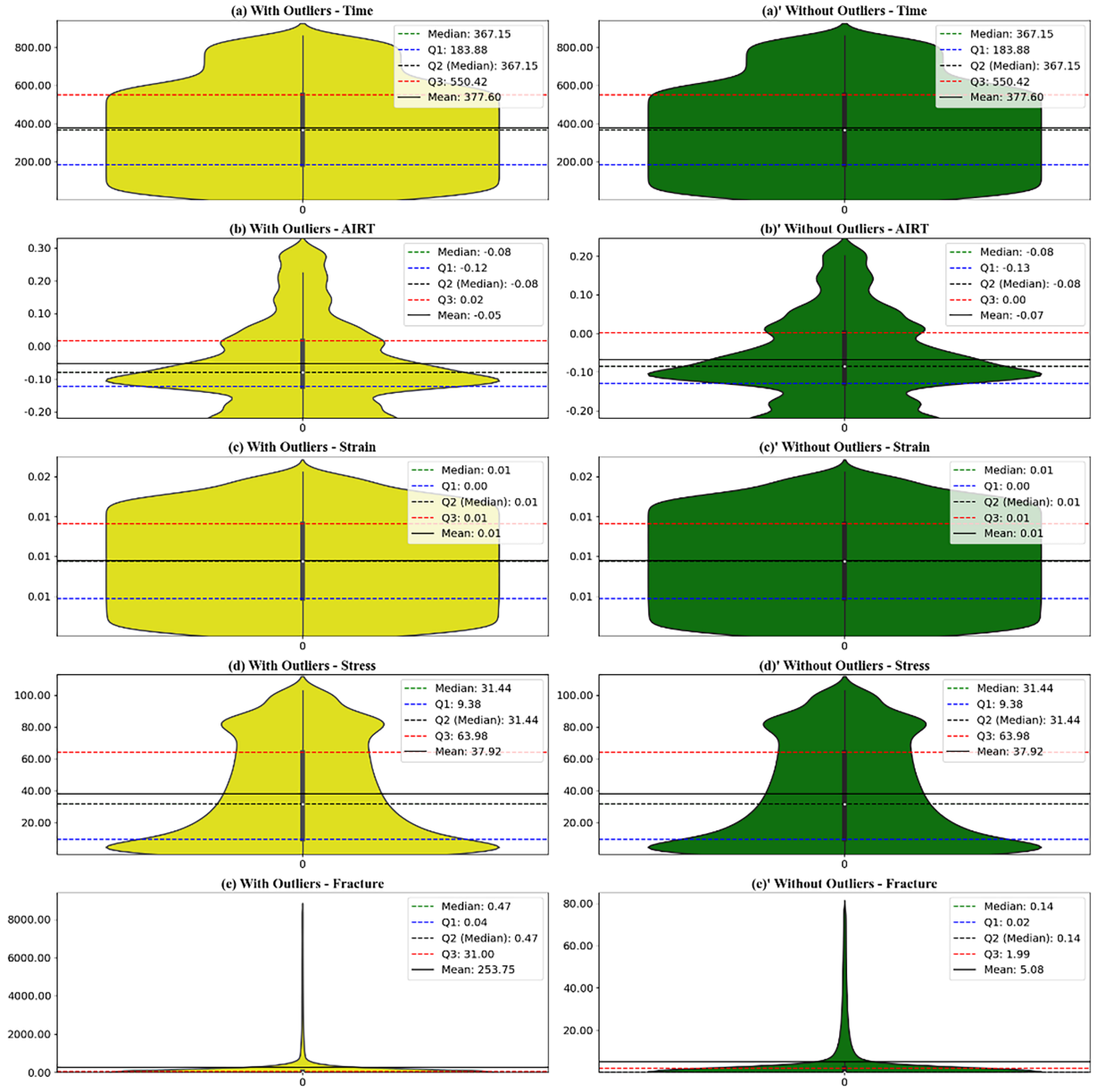

Figure 17: Violin plots for research data distribution.

4.7 ML Prediction Models

4.7.1 Data Description and Splitting

K-fold cross-validation (k = 5) is used in this study to split the dataset into training and testing sets for the XGB, AdaBoost, LGBM, and CatBoost algorithms. In contrast, the GEP model development 80:20 training-to-testing split was much simpler. In this method, the entire database is divided into k folds of equal size and used as a testing set once, and the $k - 1$ folds are used for training. It is repeated k times, allowing each data point to be used for both training and testing purposes. This technique allows us to effectively train machine learning algorithms and accurately assess their performance on unseen data, thus evaluating the models' ability to generalize and reduce the risk of overfitting of the algorithms to the training set. Table 5 presents a statistical overview of the dataset used in this investigation.

Table 5: Model input and output parameters statistical details.

Indexes	Time (s)	AIRT (°C)	Strain	Stress (MPa)	Fracture
Count	14,664	14,664	14,664	14,664	14,664
Mean	377.602	−0.052	0.009	37.923	253.746
Std	228.814	0.127	0.005	30.619	977.781
Min	0.6	−0.244	0	0.024	0
25%	183.875	−0.124	0.005	9.384	0.037
50%	367.15	−0.08	0.009	31.44	0.467
75%	550.425	0.016	0.014	63.978	31
Max	857.5	0.3	0.02	102.796	8550

4.7.2 Data Distribution and Correlation

Before implementing the prediction model, the dataset was thoroughly cleaned to eliminate outliers (the points significantly deviating from other datasets can distort the analysis by affecting measures of central tendency and variability). Detecting and eliminating outliers is essential for ensuring the precision and trustworthiness of data analysis. A statistical summary including features such as standard deviation, minimum, mean, and maximum values was computed for each variable both before and after outlier removal, as depicted in Fig. 17. The examination revealed diverse impacts of outlier removal on different variables. While time and stress displayed consistent central tendencies and variability after outlier removal, AIRT exhibited a slight decrease in the mean and standard deviation, indicating a trend toward lower values and decreased variability. Strain experienced minimal alterations in summary statistics, suggesting that outlier removal had negligible effects. In contrast, for fractures, the average and variance dropped drastically once we removed undesirable outliers, again indicating how much outliers can affect the fracture mean value. The results emphasize the need to detect and exclude extremes for proper data analysis. Knowledge of these effects is fundamental to interpreting the data correctly and making informed decisions based on statistical analysis.

The correlation matrix for the dataset in this study is depicted in Fig. 16. The matrix illustrates the correlation values of inputs and outputs, where each cell in the matrix represents a correlation coefficient value between the two features. A value close to 1 is a strong positive correlation, while a value closer to −1 is a strong negative correlation. The table below the matrix gives a summarized numerical representation of those correlations. Identifying multicollinearity in between the columns is a critical step in model preparation—it is imperative to get both a matrix representation and a tabular representation for analysis.

4.7.3 AI Model Development

This study explores the use of AI-based regression models (GEP, GBR, XGB, LGBM, Adaboost, and the CatBoost regressor) to estimate sandstone strain, stress, and fracture under variable water conditions. The subsequent sections provide the details of the implementation of these algorithms and their validation using error estimation.

GEP Model Development

The GEP method was implemented in this study by the use of GeneXpro Tools version 5.0, a tool for genetic programming. A data-splitting option was used to separate the dataset into an 80:20 ratio before the algorithm training phase. As different datasets and ML algorithms need different hyperparameters, several GEP hyperparameters need to be specified before the training process can be initiated. These hyperparameters have a dramatic impact on the performance of the model, and finding the optimal values

requires testing and experiments over a large set of possibilities. The main GEP hyperparameters are chromosome number, chromosome size, functions, and number of genes. The literature suggests that a higher number of chromosomes might impact the convergence of the model and its complexity. And hence, these parameters are fine-tuned using an iterative trial–and–error approach coupled with information from the pertinent literature [59]. In this study, GEP models for predicting stress, strain, and fracture were developed by exploring a feasible range of values for these properties, respectively, while identifying the best performing configuration. The GEP parameter setting used for the models is given in Table 6. In the table, as an example, we can see that there are 4 genes associated with prediction for stress and 5 with prediction for strain. Therefore, four subexpression trees were generated for the prediction of stress, five for strain, and four for fracture. As the linking mechanism, we combined these subexpression trees through the addition function in order to generate the final output equations. Essentially, higher values on these parameters can help the model be more accurate, but they also increase the algorithm in complexity. Thus, appropriate features, namely the number of chromosomes and head size, were selected to avoid overfitting or underfitting computation time. These final GEP equations, used for predicting stress, strain, and fracture, are based on simple arithmetic functions, as well as the square root function, to strike a balance between computational efficiency and the simplicity and utility of the resultant equations.

Table 6: Algorithm fitting parameters used in the study.

Parameters	Significance	Settings		
		Stress Prediction	**Strain Prediction**	**Fracture Prediction**
Constants per Gene	Quantifies the number of constants in each gene	10	10	10
No. of Genes	Shows the number of subexpression trees	4	5	4
Linking Function	Used to link the subexpressions	Addition	Addition	Addition
No. of Chromosomes	Total number of chromosomes made by GEP	50	50	50
Functions	Arithmetic functions that make the equation	+, −, ×, ÷, sqrt, x^2, x^3, x^4, x^5, *cube root*	+, −, ×, ÷, sqrt, x^2, x^3, x^4, x^5, *cube root*	+, −, ×, ÷, sqrt, x^2, x^3, x^4, x^5, *cube root*

Table 6 shows that the final equation for strain, stress, and fracture contains the following elementary arithmetic functions: (+, −, ×, ÷). The selection attempts to minimize the complexity of the resulting equation. Furthermore, addition is selected as the equation's linking function, which means that the subexpressions resulting from the decoding of each gene are added to produce the finished result. Eqs. (4)–(6) provide the strain, stress, and fracture equations derived from the GEP models.

$$Stress = X_1 + X_2 + X_3 + X_4 \tag{4}$$

where

$$X_1 = \frac{(178.213 * \mathrm{AIRT}) + \mathrm{T}}{7.4313}$$

$$X_2 = \left\{\sqrt{\sqrt[3]{T}} \times \left(\frac{AIRT \times T}{-5.833}\right)\right\} - 7.190$$

$$X_3 = \left\{2AIRT - \sqrt[3]{T}\right\} + 22.95$$

$$X_2 = \sqrt{\left\{\sqrt[3]{T} + 24.256\right\} + \{(3.79T) + (6.40 + AIRT)\}}$$

$$Strain = X_5 \div X_6 \div X_7 \div X_8 \div X_9$$

where

$$X_5 = T - \sqrt{T + AIRT}$$

$$X_6 = 6.871 + 2\,(AIRT) \tag{5}$$

$$X_7 = 28.619$$

$$X_8 = -12.33 - \sqrt{T + 7.67}$$

$$X_9 = -6.51 - AIRT - \left\{\left(-9.92 - \sqrt{T}\right) \times (AIRT)\right\}$$

$$Fracture = X_{10} \times X_{11} \times X_{12} \times X_{13} \times X_{14}$$

where

$$X_{10} = \sqrt{\sigma} \times (\varepsilon)^2 - \left[\left\{1.84 - (-8.48 - \varepsilon) - \sqrt[3]{\varepsilon}\right\} - \left\{\sqrt[3]{\varepsilon} - \sqrt{\sigma}\right\}\right]$$

$$X_{11} = \frac{-7.27}{\sigma}$$

$$X_{12} = \frac{\left[\left\{\left(\frac{1.815 - \sigma}{\sigma}\right) - \sigma\right\} + \{A + B\}\right] \times [\sigma - 6.01]}{\varepsilon}$$

$$X_{13} = \varepsilon$$

$$X_{14} = \left[\left\{\left(\left\{\left(6.55 - \sqrt{\varepsilon}\right) \times -8.763\right\} + \sigma\right) + N\right\} - 2.061\right] - \sqrt{\sigma}$$

where (6)

$$A = \left(\sqrt[3]{-0.869 + \left(5.98 \times \sqrt[3]{\sigma + 1.320}\right)}\right) - \left(\varepsilon - \sqrt[2]{S}\right)$$

$$B = \frac{\sigma}{\frac{(3.50\varepsilon + \sigma) \times \frac{2.195}{\sigma}}{\varepsilon} \times \left(\sqrt[3]{\varepsilon} - \varepsilon\right)}$$

$$N = \begin{bmatrix} \\ \dfrac{0.561 - \sigma + \varepsilon}{\dfrac{0.68\sigma + \sigma + \varepsilon}{\dfrac{0.879\varepsilon}{\sqrt{\sqrt{\sigma}} + 6.554}}} \end{bmatrix} + \left[\sqrt[3]{\{\varepsilon - \sigma^2\} - \left\{68.162 + \frac{9.85}{\sigma}\right\}}\right]$$

Ensemble Algorithms Model Development

Although hyperparameter tuning for GEP was performed manually, involving extensive testing with a range of values, hyperparameter tuning for XGB, LGBM, CatBoost, BR, and AdaBoost was more efficient.

Specifically, a widely used method known as grid search, which is available in the Python scikit-learn library through the GridSearchCV function, was employed for this purpose. Importantly, not only hyperparameter optimization but also data loading, data splitting, model fitting on the training set, and testing on unseen data were performed via the Python language within the Anaconda software environment. The grid search method was used as a preferred choice for parameter tuning because of its efficiency and extensive use in optimizing machine learning model parameters [46]. The grid search approach involves evaluating every possible combination of hyperparameter values to select the most feasible set of parameter values for the model. In this approach, a parameter is changed over a range of values to find the optimal setting, holding all other parameters constant. This continues for every single parameter until the best set of values is found for the respective model. In this study, the same method is applied to find the best hyperparameters for XGB, AdaBoost, CatBoost, LGBM, and BR algorithms. Table 7 presents the grid search results of these models, reporting the parameter values that return the highest accuracy for each algorithm.

Table 7: Optimized Hyperparameter settings for XGB, AdaBoost, GBR, and CatBoost algorithms.

Algorithm	Hyperparameter	Optimal Value
XGB	Number of trees Maximum depth of each tree Shrinkage rate applied to each tree	n_estimators: 100 max_depth: 6
AdaBoost	Number of trees Maximum depth of each tree Shrinkage rate applied to each tree	learning_rate: 0.3 n_estimators: 50 max_depth: 8
GBR	Shrinkage rate applied to each tree Maximum depth of each tree Number of trees	learning_rate: 1.0 learning_rate: 0.1 max_depth: 3
CatBoost	Number of iterations (trees) Shrinkage rate applied to each tree Maximum depth of each tree	n_estimators: 100 iterations: 1000 learning_rate: 0.03

Notice some common hyperparameters, such as n_estimators, max_depth for XGB, AdaBoost, GBR, CatBoost, and BR algorithms, which are keys for model training. As previously mentioned, n_estimators indicates the number of base learners the algorithm will use to make predictions, while max_depth specifies a tree's maximum depth. The other really important hyperparameter common across these models is the learning rate. Essentially, the learning rate is a shrinkage that is applied to the predictions of each individual new tree. In boosting algorithms, trees are added incrementally to minimize the residuals of previous trees; a large learning rate can contribute to a fast fitting [44], leading to potential overfitting. To allow the model to adjust more slowly to mitigate overfitting, a lower learning rate is recommended. Consequently, the learning rates employed in this study were cautiously chosen based on advice from the literature and practical assessment. For CatBoost, a learning rate of 0.03 was chosen; for XGB, a learning rate of 0.3 was used; and for GBR, a learning rate of 0.1 was selected. These choices are reflected in Table 7, which details the optimal hyperparameters for each algorithm.

4.8 ML Model Performance

4.8.1 GEP

Table 8 shows the performance metrics for GEP-based predictions of strain, stress, and fracture in both phases. The results show that all the predicted parameters have good performance for GEP, with $R^2 > 0.80$. Specifically, the R^2 for fracture prediction on the testing set was 0.985, demonstrating that the GEP-based equation is capable of predicting fracture values with up to 98.5% accuracy for previously unseen data. The curve fitting graphic in Fig. 18 also shows the predictive nature of the GEP equation. The plot also indicates GEP model correctly predicts strain, stress, and fracture. Small differences between the real and predicted values indicate the robustness of the GEP algorithm in each moment, showing the efficiency of this technique in estimating rock properties.

Table 8: Summary of the error evaluation of GEP.

	Stress		Strain		Fracture	
	Training	**Testing**	**Training**	**Testing**	**Training**	**Testing**
MAE	7.65	7.69	4.73	4.73	46.824	46.521
RMSE	9.29	9.40	6.02	6.01	130.239	12.767
R^2	0.910	0.912	0.948	0.950	0.982	0.985

4.8.2 GBR

In Table 9, the active prediction results of the GBR model for different evaluation metrics with very low prediction errors are demonstrated. Starting with the training dataset, it can be seen that MAE = 0.0127, RMSE = 0.0134, while for the testing data, MAE = 0.0005 and RMSE = 0.0005; thus, the accuracy of the testing dataset is high. These metrics illustrate the GBR model's strong performance in predicting the target value accurately, rendering it the best-performing model in this evaluation. Sandstone property prediction under different water conditions is given in Fig. 18.

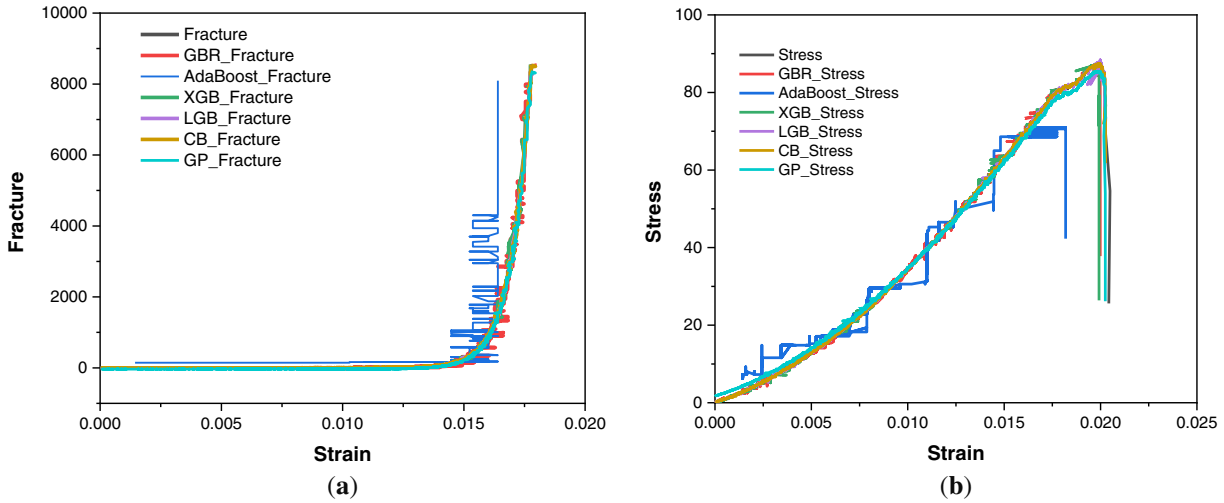

Figure 18: (Continued)

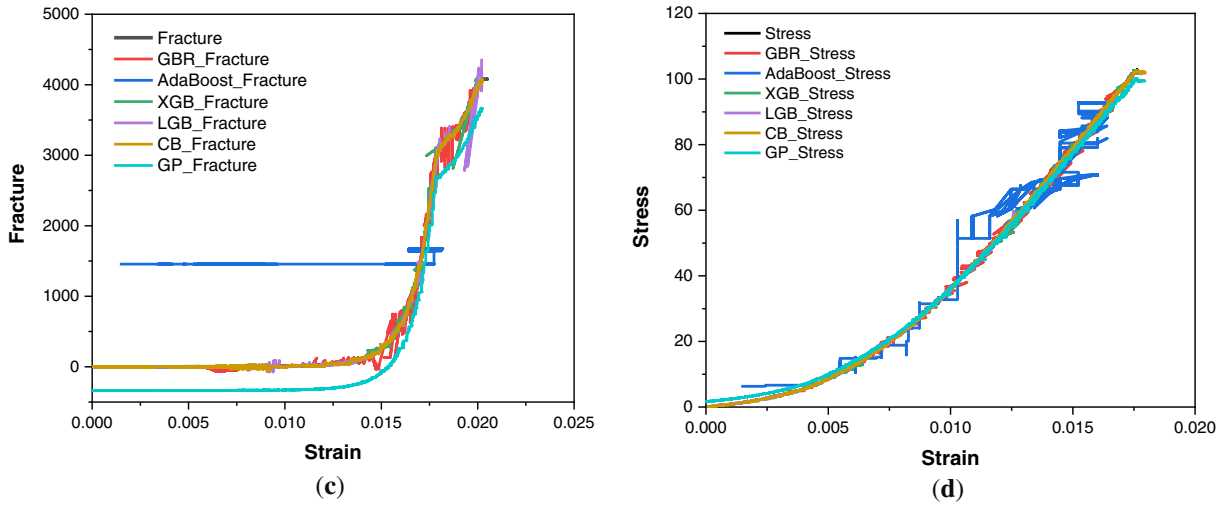

Figure 18: Sandstone property prediction under different water conditions: (**a**) stress–strain prediction under dry conditions; (**b**) fracture prediction under dry conditions; (**c**) stress–strain prediction under saturated conditions; (**d**) fracture prediction under saturated conditions.

Table 9: Model performance of the ensemble learning algorithms.

Models	Training MAE	Testing MAE	Training RMSE	Testing RMSE	Training R^2	Testing R^2
GBR	0.012679	0.013353	0.000473	0.000501	0.999526	0.9995
AdaBoost	0.11869	0.119579	0.02022	0.020474	0.979649	0.9800
XGB	0.00452	0.005545	0.000398	0.000597	0.999597	0.9994
LGBM	0.005966	0.006618	0.000515	0.000627	0.999479	0.9993
CatBoost	0.006042	0.006727	0.000454	0.000568	0.999542	0.9994

4.8.3 XGB

XGB stands out by providing better accuracy metrics than its counterparts. For instance, MAE of 0.0045 and 0.0055 RMSE are excellent performances as shown in Table 9. Its performance is ongoing, strong, and stable on the trying-out dataset, producing an MAE of 0.0004 and an RMSE of 0.0006. These results further reinforce XGB's excellent suitability for prediction tasks.

4.8.4 AdaBoost

Although the AdaBoost Regressor did a good job at predicting the required values, the error rates are slightly higher compared to GBR. The performance metric of 0.1187 for MAE and 0.1196 for RMSE for the training phase proves that although the errors are much bigger than those of the GBR, the model has a pretty high accuracy rate. On the test dataset, the AdaBoost regressor is also consistent, as its MAE is 0.0202 and RMSE is 0.0205, which gives validation of its effectiveness in predicting the target variable with fairly consistent performance.

4.8.5 LGBM

LGBM performed exceptionally well, with error rates comparable to XGB. On the training dataset, it exhibited an MAE of 0.0060 and an RMSE of 0.0066. Moreover, it showed an excellent predictive power on

unseen data, achieving very low Mean Absolute Error (MAE = 0.0005) and Root Mean Square Error (RMSE = 0.0006) scores on the testing dataset. The LGBM model shows accuracy based on these metrics, which indicate its efficiency and effectiveness in prediction.

4.8.6 CatBoost

CatBoost also predicts exceptionally well in both training and testing phases. The model achieves an MAE of 0.0060 and RMSE of 0.0067 on the training dataset, and this performance is similar to the results achieved by the LGBM and XGB models. It is stable because its MAE is 0.0005 and its RMSE is 0.0006 as a result of the testing data set, which emphasizes the uniformity of its performance on each data set.

5 Discussion

Naturally, rocks can disintegrate more when water dissolves the clay minerals in the rocks, transforming the bonding structure itself, making the bonding structure even further weak. Water-induced rock fracture has been elaborately discussed in previous studies [5]. Especially, this process is important for clay minerals. For the sandstone used in this experiment, XRD analysis showed that it contained 7.9% clay minerals. This study also suggests that water can significantly affect the degradation behavior of the internal structure of sandstone [60,61]. It consists primarily of clay minerals, calcite, dolomite, quartz, and clay minerals that have hydrophilic characteristics. Water molecules can wedge themselves into the molecular structure of these minerals, which can lead to further dissolution. To date, existing research has not adequately studied water-rock interactions under diverse water conditions, along with the role of mineral constituents on the mechanical behavior and structural resilience of rock masses. The study should focus on the role of dissolution of clay minerals and other hydrophilic materials in the mechanical properties and stability of rock under various physical conditions. As water critically governs the fracturing process of sandstone, the infrared monitoring of saturated sandstone was therefore investigated. The experiments demonstrate that wet rock reflects more IR off the surface of the sandstone. This conclusion is in line with existing academic research [62], despite the very limited attempts that have been made to better understand this phenomenon. Thermodynamically speaking, the stored heat in the entire mass of saturated sandstone is more than that of dry sandstone, hence the IR on the saturated sandstone surface would distribute a wider range of temperature. As a result, the AIRT appears to be lower in wet than in dry samples. The authors will conduct additional studies to study the specific heat capacity of the same rock with different amounts of water in subsequent studies. Moreover, they want to explore how this affects the IR properties of the sandstone surface.

The temporal and geographical feature distributions of the IR temperature on the sandstone bearing surface are investigated in this research via the use of data from image processing methods. It is necessary to use statistical techniques to explain the internal fractures that are present inside the rock samples, since an investigation revealed that the temperature of the infrared radiation is not distributed evenly. In the past, the process of characterizing rock fractures was accomplished by examining fluctuations in IR indicators during the failure. This work focused on increased infrared accuracy through PFC grain flow modeling and numerical modeling of sandstone. The motive behind this work was to investigate the evolution of microcracks in sandstone samples with varying water exposure and to define a correlation that allows us to predict the variations of the IR and the features of the fractures of sandstone. In addition, the study compared multiple AI algorithms to find the best technique for predicting fractures. The next research direction will be the establishment of a fracture and artificial intelligence (AI) model, which could be generalized to various loading situations, considering the size of the sample and the types of rock. The purpose of this specific study is to enhance the efficiency of fracture prediction in rocks by employing infrared radiative temperature information acquired at multiple time points from the exterior of sandstone. This study

provides a quantitative and reliable approach to characterizing the temporal distribution and spatial features of fractures.

In addition, to investigate rocks that contain water, a solid understanding of the patterns of water distribution within rocks through the use of particle flow modeling is essential. Previous research has employed nuclear magnetic resonance imaging (NMR) to explore water dispersion in coal samples [5,28]. Similarly, this research utilized the discrete element method for modelling the effect of water content on coal. However, because coal and sandstone have different internal compositions and properties, this study worked under the assumption that water is dispersed uniformly throughout the sandstone. To increase the accuracy of these simulations, further research will be undertaken for the investigation of the water dispersion characteristics of sandstone. This study aimed to develop an accurate prediction model for fractures that may serve as early indicators of failure.

5.1 Field Implementation

Fig. 19 presents a flowchart providing real-time field guidance for engineers. The protocol begins at the Continuous Field Monitoring node (teal), or the situation in which engineers have continuous monitoring of the infrared and acoustic emission signal, both online, even without known absolute levels of stress. Starting here, the protocol bifurcates into two parallel lines of evaluation. Rule 1: IR Trend (blue) is a constant reminder of the slope of the AIRT, and Rule 2: AE Validation (purple) is a constant reminder of cumulative AE energy. These two rules act at the same time and independently of each other; neither of them is thought to be enough to cause a response. In the IR branch, as long as the slope of the AIRT is greater than 16°, then the specimen is considered to be in a steep plastic deformation phase, and no further action is justified than continued observation. But as the slope decreases to the moderate gradient regime (below 16°), the main infrared warning signal fires, which in the experiment correspond to the precursor stress threshold of 0.72 sigma (dry) and 0.75 sigma (saturated). In the same branch (AE), at the same time, a steep rise in cumulative AE energy is used to verify the IR warning, telling the difference between a real precursor event and an artefact of measurement or noise. These two branches then meet at one junction point, which is a requirement of the protocol that both conditions should be met at the same time before any action can be undertaken. This linkage gives way to the Operational Threshold Reached node (coral), where the 0.72–0.75 sigma critical stress window is set, and active response is deemed to have been reached. Lastly, the flowchart ends at the Implement Immediate Safety Measures node (red), which indicates that the catastrophic rock failure is about to occur and that preventive measures should be taken immediately.

5.2 Explainable Machine Learning (XML)

As the use of ML techniques to predict several important rock and soil parameters increases, there is a growing need to make the ML-based models transparent and easy to use in practical settings (Sagi and Rokach 2020). In this context, XML techniques play a crucial role in understanding complex black-box ML models. In the realm of civil engineering and rock mechanics, SHAP and ICE analysis are the most commonly used XML techniques. Thus, these two methods are used in this study to explore how input variables affect the rock fracture prediction under uniaxial loading. SHAP analysis identifies and ranks the most important inputs for predicting the outcome, while ICE analysis helps to visualize how the output is affected by input variables.

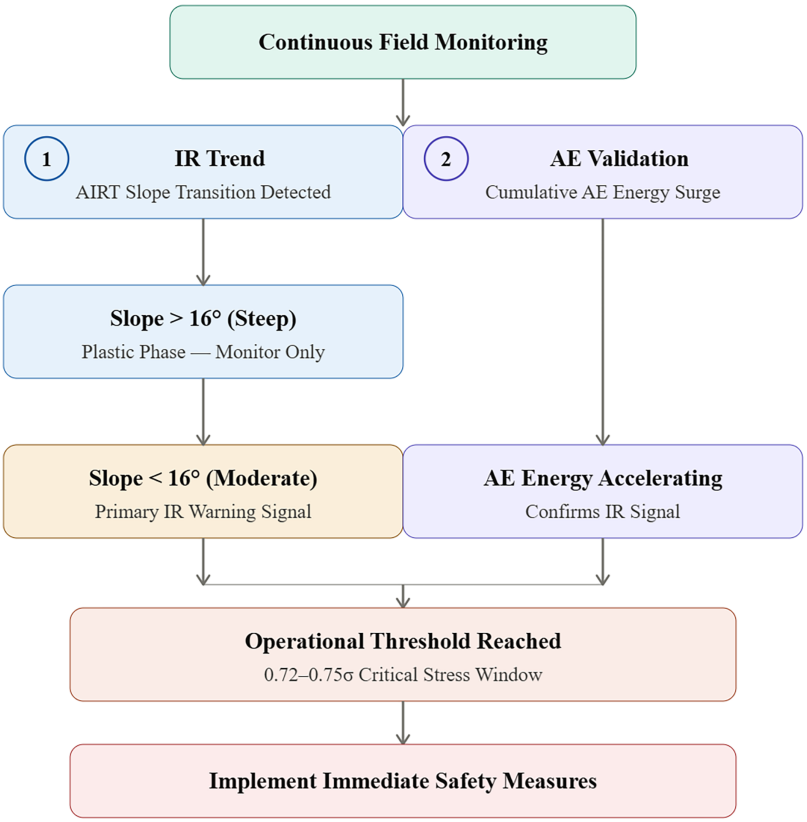

Figure 19: Illustrates the protocol of real-time rock failure monitoring using IR and AE.

SHapley Additive exPlanations (SHAP) Analysis

SHapley analysis, created by Lundberg and Lee [62], helps explain ML models based on game theory, where features are seen as players, and the model's outcomes are treated as the rewards. To interpret tree-based models like XGB, a tool called TreeSHAP is commonly used. In this study, TreeSHAP is implemented using Python to interpret the predictions made by the XGB model for predicting rock fracture under uniaxial loading. SHAP ranks the input variables according to their contribution in predicting the output by means of its global interpretation analysis which can be visualized by means of a SHAP summary plot (also known as beeswarm plot) in Fig. 20. The contributors are arranged from top to bottom in order of decreasing significance. Each input variable has its own row, with dots representing its SHAP values for different data instances. The dots are color-coded in such a way that blue represents lower values of the inputs, while red shows higher values. This plot not only ranks the variables but also gives the idea of how the output is affected by input values, i.e., if red dots are present on the negative side, it indicates that increasing input values decrease its SHAP value (consequently lowering its significance in predicting the outcome), and *vice versa*. Fig. 21 shows that the time of loading is the most dominant factor in predicting the rock fracture, having the widest range of SHAP values in the positive direction. It shows that increasing loading time increases its SHAP value and consequently increases rock fracture. The lower values for loading time are, however, clustered closer to the zero line on the negative side of the SHAP plot. It indicates that at lower loading times, there will be fewer chances of rock fracture. Over time, stress on the rock is the most crucial factor, and it exhibits similar behaviour to the time of loading. However, the strain developed in rocks under loading shows a different behaviour compared to the other two inputs, as the red dots (indicating higher strain values)

are stretched on both positive and negative sides, while the blue dots of lower strain values are consistently clustered along the zero line. It shows that when there is less strain developed in the rock, the tendency to develop a fracture is also low, as explained previously by [26,31–33]. The least contributing factor, according to the SHAP plot, is IR, and it can be seen from Fig. 20 that the SHAP values for IR are lying practically on the zero-line, indicating its lesser significance in predicting rock fracture under uniaxial loading.

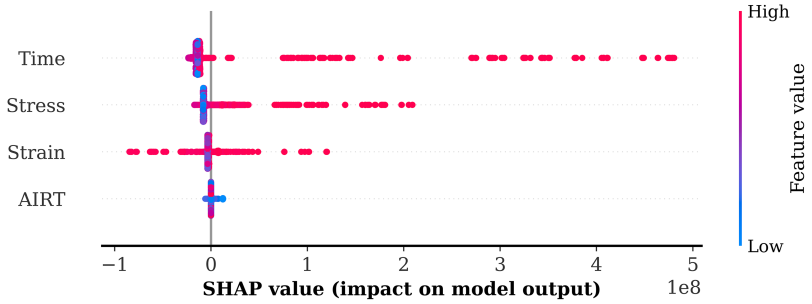

Figure 20: Input contribution ranking to predict rock fracture.

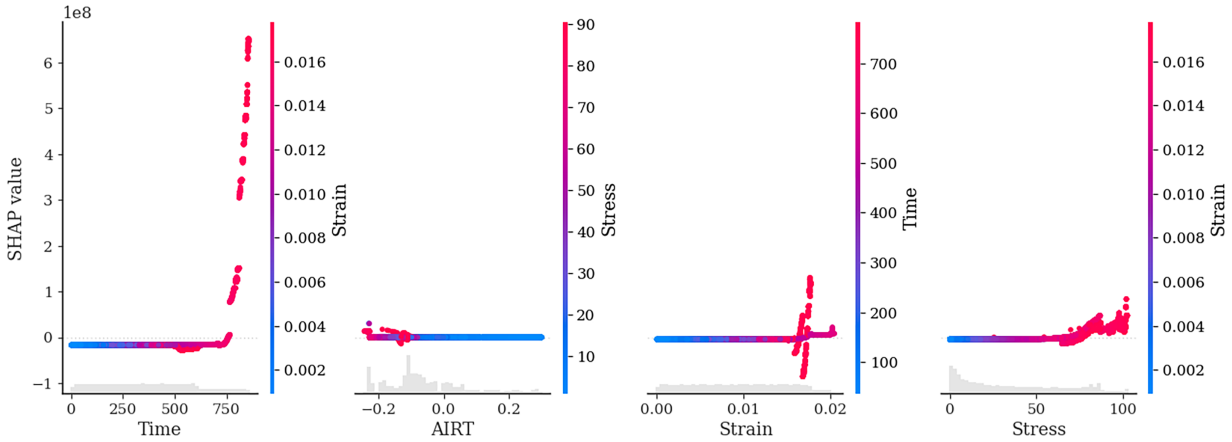

Figure 21: SHAP feature interaction plots between input variables.

After ranking the input variables based on their importance in the prediction of the outcome, Fig. 20 unravels the useful relationships between input variables and their corresponding SHAP values through SHAP partial dependence plots or more commonly known as feature interaction plots. A condensed representation of these plots for the variables considered in this study is shown in Fig. 21, where the values of the variables are displayed along the bottom axis and the right-hand side, while the SHAP values for the variable on the x-axis are shown on the left-hand side. This arrangement allows for observation of how the SHAP value of one variable changes with variations in another. Overall, the SHAP values associated with time exhibit greater variability compared to other variables, indicating its dominant role in predicting the outcome. In addition, the SHAP value of time (and thus its importance in the prediction of the output) increases with an increase in strain values, which unveils the multivariate relationship between time and strain in determining rock fracture. The SHAP values of other variables show some variability, but it is not as pronounced as time, which seamlessly aligns with the ranking of variables given by the SHAP summary plot in Fig. 20.

Additionally, each partial dependence plot includes a background histogram showing the frequency distribution of the variable on the x-axis. Therefore, these plots not only illustrate interactions between variables through SHAP values but also provide insight into the distribution of the input data.

5.3 ICE Analysis

SHAP analysis helps to identify the most influential variables in the prediction of output. However, it is also important to investigate the impact of changing each input on the ML algorithm's predictions. This can be done by ICE analysis, in which one input is varied across its range while all other parameters are held constant at their average values. The results of change in the output values are plotted against the original input values to get an ICE plot for a variable as shown in Fig. 22. The impact of change in IR on fracture of rocks under uniaxial loading can be seen from Fig. 22a. As IR increases from −0.3 to almost −0.1, the fracture tendency decreases rapidly and after that it becomes constant for the whole range of IR. It must be noted that the change in IR over its whole range only brings about a change in output from 4.25×10^8 to almost 2.5×10^8. This is a comparatively smaller range, and it is due to the fact that IR was previously listed as the least contributing variable to the output by means of SHAP analysis. The rock fracture experiences an increase with an increase in loading time, as evident from Fig. 22b. It can be seen that increasing loading time from 0 to 800 s increases the fracture from almost 0 to more than 2.5×10^8. However, the predominant effect is seen after a loading time of 600 s. It can be attributed to the fact that the time of loading has been identified as the most crucial factor for determining target output by SHAP analysis. The stress and strain in the rock exhibit somewhat similar behaviour in relation to the fracture under uniaxial loading. The effect of both inputs is minimal in the beginning, but it increases rapidly after strain and stress surpass 0.015 and 60, respectively (Fig. 22c,d). This happens because initially, there is less stress acting on the rock and consequently less strain. However, when the rock is exposed to loading for longer time periods, both stress and strain increase, resulting in the formation of fractures.

5.4 Data-Driven Decision Making and GUI Development

It has been explained in previous sections that the ML models are reliable for predicting the subject output. Also, the XML techniques depicted that the predictions made by the most accurate algorithm, i.e., XGB, align with the observations obtained from laboratory experiments and previous literature. It is, however, imperative to make the decision-making process of the algorithm more transparent to make sure that engineers and different stakeholders can effectively utilize the findings of this study. This can be done by means of the SHAP decision plot, as shown in Fig. 23, which helps to visualize the prediction-making process of the XGB algorithm for all experimental instances considered in this study. This plot shows how the input variables affect the model's final prediction. The effect of input variable values on the predicted outcome can be investigated through the colour of different lines, where blue lines indicate lower output values while red lines represent higher values. The lines originate from the bottom and intersect the graph's top, indicating the predicted value for that particular instance. From the decision plot, it is evident that higher values of inputs like stress and loading time push the output towards higher values compared to the rest of the variables. Also, these two inputs are the most influential variables according to the SHAP summary plot. This indicates the effectiveness of SHAP analysis in ranking the variables correctly according to the decision-making rules of the algorithm.

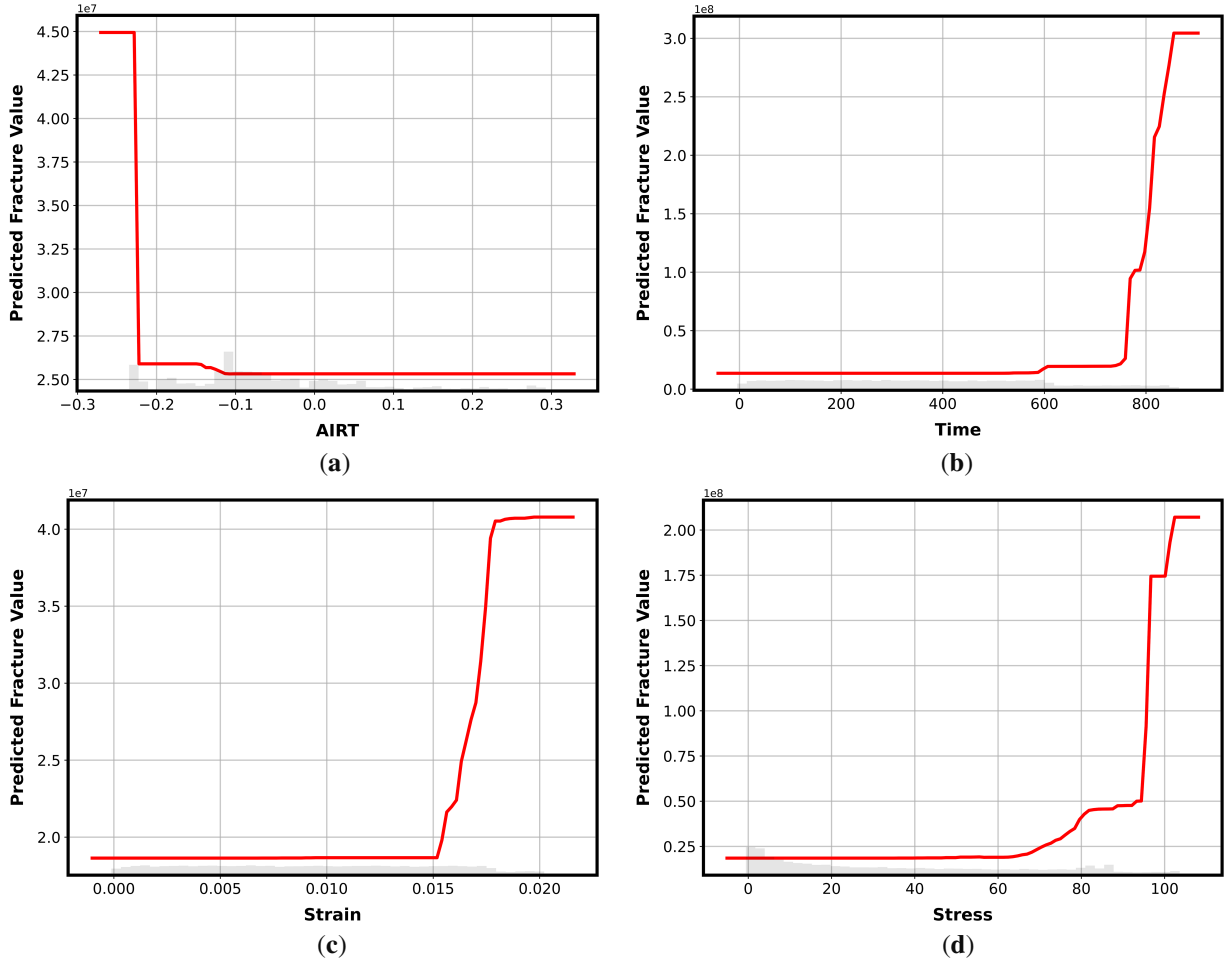

Figure 22: Effect of input variables on the predicted outcome by means of ICE analysis; (**a**) IR; (**b**) time; (**c**) strain; (**d**) stress.

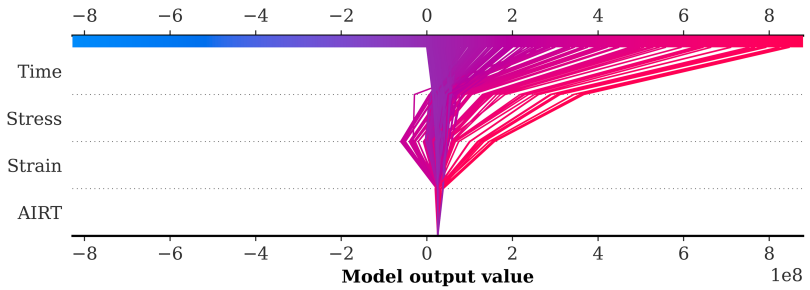

Figure 23: SHAP decision plot to depict the decision-making process of the algorithm.

It has been proven by the implementation of XML techniques that the XGB algorithm is making predictions according to the findings of previous studies. Moreover, its decision-making process has been made more transparent by means of a SHAP decision plot to overcome its black-box nature. To enhance the usability and accessibility of the developed XGB model for predicting rock fracture, a graphical user interface (GUI) has been developed to provide a platform for determining the output without requiring

explicit programming knowledge. The users can manually insert the input values to obtain an immediate prediction, as shown in Fig. 24. The GUI can be accessed at: (https://github.com/WaleedBinInqiad/Rock-Fracture-Prediction). In addition to providing accurate predictions, this tool integrates the SHAP force plot to provide a visual interpretation of how each input parameter influences the predicted output. A SHAP force plot decomposes the baseline value $f(x)$ into contributions from each feature such that the blue arrows indicate a negative impact on the outcome, while red arrows show a positive impact. The example used in Fig. 24 presents an instance where the XGB model predicted the value of rock fracture as 1075.43. From the accompanying SHAP plot, it can be seen that time is the most influential variable pushing the output towards a lower value. The stress and strain are also reducing the predicted outcome, highlighted by their blue color. Notice that the significance of variables in guiding the final output value agrees with the pattern obtained by the SHAP summary plot and ICE analysis in the previous sections. This plot helps to dissect individual predictions and helps to investigate which variables contribute to increasing or decreasing the outcome for a particular instance. Thus, this interpretability feature enhances the transparency of the XGB model by allowing users to understand the relative contribution of individual predictors, thereby improving confidence in model outputs and supporting informed engineering decisions. Overall, by combining efficient prediction capabilities with XML analysis, this tool demonstrates significant potential for adoption in geotechnical engineering in supporting decision-making processes for the safe and reliable design of underground excavated structures.

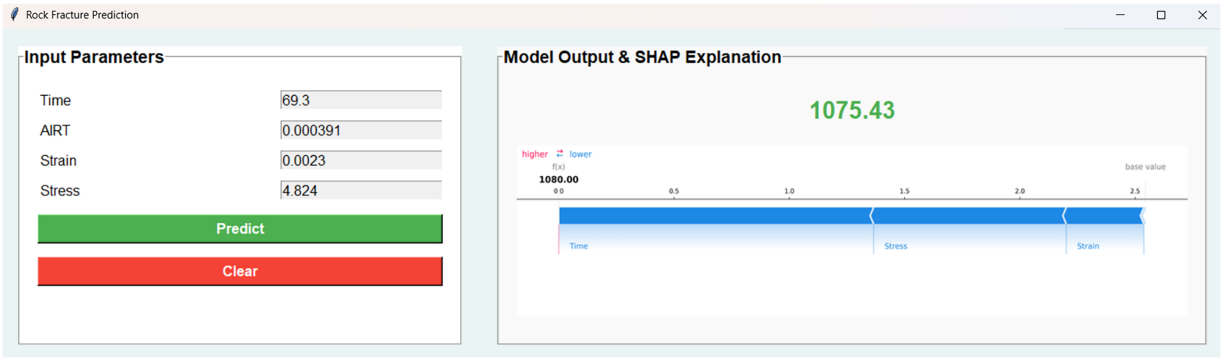

Figure 24: Graphical user interface developed for efficient output prediction.

6 Conclusions

The following conclusions were drawn from the current research:

- The mineral composition determined by XRD revealed that clay minerals accounted for approximately 6.5% of the sample. Over time, the presence of clay in saturated samples leads to progressive dissolution, causing the expansion of the pore structure of sandstone and the formation of interconnected pores, which consequently promotes the development of cracks and ultimately affects the structural strength of sandstone. Analysis of the AE data and PFC grain flow simulations reveals that saturated sandstone experiences more cracks than dry sandstone does during the process of loading.
- The rate of change of the AIRT exhibits excessively steep slopes, approximately >60° during the crack closure stage, extremely steep slopes (approximately 31°–60°) during the elastic deformation stage, and steep slopes (approximately 16°–30°) during the plastic stages, with a moderate or near-horizontal slope preceding rock failure. Similarly, the cumulative acoustic energy shows steep slopes within the AIRT moderate slope range, with dense energy counts in this region. The steepness of slopes is greater in

saturated rock than in dry rock. Additionally, a moderate slope can serve as an early warning precursor of rock failure.

- PFC grain flow simulations reveal that rock fractures occur during the loading process. As these fractures develop, there is a surge in fractures and cumulative energy. Simultaneously, the slope of the AIRT transitions from steep to moderate. This convergence of fractures, cumulative energy, and the AIRT slope serves as an early warning precursor. Specifically, this precursor manifests at stress levels of 0.72 and 0.75 for dry and saturated sandstone, respectively.
- Different AI techniques were used to estimate sandstone strain, stress, and fracture under varying water conditions. The accuracy comparison revealed that XGB outperformed all the other algorithms, achieving R^2 values equal to 0.999 for both datasets, followed by GBR, CatBoost, LGBM, GEP, and AdaBoost. In contrast, GEP provided a novel empirical equation for the prediction of stress, strain, and rock fracture, whereas all other employed AI algorithms did not demonstrate this capability.
- Explainable machine learning techniques like SHAP and ICE were implemented to make the black-box XGB algorithm explainable and transparent. Also, the SHAP decision plot was used to depict the decision-making process of the algorithm.
- An interactive GUI has been developed to efficiently predict the rock fracture under uniaxial loading. The knowledge gained from this research will improve the understanding of sandstone behavior under various conditions and facilitate the development of more accurate models to aid fracture forecasting and early failure detection.

This research has a few limitations, which limit the generalization of the research results. First, only two discrete moisture conditions (dry at 0% and saturated at 3.34%–3.89%) were investigated and the experiment focused only on sandstone materials with uniform geometry (50 mm × 50 mm × 50 mm cubes) that were tested under uniaxial compression, thus ignoring intermediate moisture conditions and other loading regimes, such as triaxial, cyclic, or dynamic conditions used in practice. Second, the sample size was quite small (12 samples, 6 dry, 6 saturated), which could potentially undermine the statistical strength and predictive capabilities of the models that were created. Third, the PFC numerical models adopted a uniform distribution of water in the sandstone matrix, a simplification that fails to reflect the heterogeneous pore structure and moisture gradient found in natural occurrences of rocks. Lastly, the models were only trained on laboratory-scale data, and their extrapolation to field-scale processes or other lithologies has not been tested, which requires further studies to determine how well the models can be transferred to different geological settings.

Acknowledgement: The authors would like to express their gratitude to China University of Mining and Technology, China for enabling this work.

Funding Statement: This work was supported by the Xinjiang Uygur Autonomous Region Key R&D Programme Projects (2023B03009).

Author Contributions: Study conception and design: Naseer Muhammad Khan; data collection: Naseer Muhammad Khan, Sajjad Hussain; analysis and interpretation of results: Naseer Muhammad Khan, Sajjad Hussain, Waleed Inqiad, Majid Khan; supervision: Liqiang Ma, Danial Jahed Armaghani; draft manuscript preparation: Naseer Muhammad Khan, Sajjad Hussain, Waleed Inqiad, Majid Khan, Tariq Feroze. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available on request from the corresponding authors.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

GEP	Gene Expression Programming
ML	Machine Learning
GBR	Gradient Boosting Regressor
PFC	Particle Flow Computation
LGBM	Light Gradient Boosting Machine
IRV	Infrared Radiation Variance
GP	Genetic Programming
GBDT	Gradient Boosted Decision Tree
GOSS	Gradient-based One-Sided Sampling
XGB	Extreme Gradient Boosting
AI	Artificial Intelligence
AE	Acoustic Emission
IR	Infrared Radiation
AIRT	Average Infrared Radiation Temperature
VDIIT	Variance of Differential Infrared Image Temperature
ET	Expression Tree
EFB	Exclusive Feature Bundling

References

1. Luo Y. Influence of water on mechanical behavior of surrounding rock in hard-rock tunnels: an experimental simulation. Eng Geol. 2020;277(B7):105816. doi:10.1016/j.enggeo.2020.105816.
2. Wu W, Gong F, Jiang Q. Influence of water on rockburst of surrounding rock in deep circular tunnels under triaxial internal unloading conditions. Tunn Undergr Space Technol. 2023;138(B7):105165. doi:10.1016/j.tust.2023.105165.
3. Zhao J, Li D, Jiang J, Luo P. Uniaxial compressive strength prediction for rock material in deep mine using boosting-based machine learning methods and optimization algorithms. Comput Model Eng Sci. 2024;140(1):275–304. doi:10.32604/cmes.2024.046960.
4. Armaghani DJ, Liu Z, Khabbaz H, Fattahi H, Li D, Afrazi M. Tree-based solution frameworks for predicting tunnel boring machine performance using rock mass and material properties. Comput Model Eng Sci. 2024;141(3):2421–51. doi:10.32604/cmes.2024.052210.
5. Khan NM, Ma L, Emad MZ, Feroze T, Gao Q, Alarifi SS, et al. Predicting sandstone brittleness under varying water conditions using infrared radiation and computational techniques. Water. 2024;16(1):143. doi:10.3390/w16010143.
6. Gong S, Zhou L, Wang W, Wang C. Investigation of dynamic fracture behavior and energy dissipation of water-bearing coal under impact load. Eng Fract Mech. 2022;275(6):108793. doi:10.1016/j.engfracmech.2022.108793.
7. Zhou Z, Cai X, Ma D, Du X, Chen L, Wang H, et al. Water saturation effects on dynamic fracture behavior of sandstone. Int J Rock Mech Min Sci. 2019;114(9):46–61. doi:10.1016/j.ijrmms.2018.12.014.
8. Hao RQ, Li JT, Cao P, Liu B, Liao J. Test of subcritical crack growth and fracture toughness under water-rock interaction in three types of rocks. J Cent South Univ. 2015;22(2):662–8. doi:10.1007/s11771-015-2568-9.
9. Ma C, Li H, Niu Y. Experimental study on damage failure mechanical characteristics and crack evolution of water-bearing surrounding rock. Environ Earth Sci. 2018;77(1):23. doi:10.1007/s12665-017-7209-1.
10. Zhang H, Lu K, Zhang W, Zhou Y, Yang G, Li J, et al. A comprehensive study of damage characteristics and acoustic emission response mechanism of sandstone with different water contents. Eng Fract Mech. 2023;288(3):109392. doi:10.1016/j.engfracmech.2023.109392.
11. Du K, Li X, Tao M, Wang S. Experimental study on acoustic emission (AE) characteristics and crack classification during rock fracture in several basic lab tests. Int J Rock Mech Min Sci. 2020;133(6):104411. doi:10.1016/j.ijrmms.2020.104411.

12. Wu L, Liu S, Wu Y, Wang C. Precursors for rock fracturing and failure—Part I: IRR image abnormalities. Int J Rock Mech Min Sci. 2006;43(3):473–82. doi:10.1016/j.ijrmms.2005.09.002.
13. Meng Q, Zhang M, Han L, Pu H, Nie T. Effects of acoustic emission and energy evolution of rock specimens under the uniaxial cyclic loading and unloading compression. Rock Mech Rock Eng. 2016;49(10):3873–86. doi:10.1007/s00603-016-1077-y.
14. Wang C, He B, Hou X, Li J, Liu L. Stress-energy mechanism for rock failure evolution based on damage mechanics in hard rock. Rock Mech Rock Eng. 2020;53(3):1021–37. doi:10.1007/s00603-019-01953-y.
15. Du F, Ma J, Guo X, Wang T, Dong X, Li J, et al. Rockburst mechanism and the law of energy accumulation and release in mining roadway: a case study. Int J Coal Sci Technol. 2022;9(1):67. doi:10.1007/s40789-022-00521-0.
16. He S, Qin M, Qiu L, Song D, Zhang X. Early warning of coal dynamic disaster by precursor of AE and EMR "quiet period". Int J Coal Sci Technol. 2022;9(1):46. doi:10.1007/s40789-022-00514-z.
17. Cai M, Morioka H, Kaiser PK, Tasaka Y, Kurose H, Minami M, et al. Back-analysis of rock mass strength parameters using AE monitoring data. Int J Rock Mech Min Sci. 2007;44(4):538–49. doi:10.1016/j.ijrmms.2006.09.012.
18. Du K, Li X, Wang S, Tao M, Li G, Wang S. Compression-shear failure properties and acoustic emission (AE) characteristics of rocks in variable angle shear and direct shear tests. Measurement. 2021;183(4):109814. doi:10.1016/j.measurement.2021.109814.
19. Yang SQ, Jing HW, Wang SY. Experimental investigation on the strength, deformability, failure behavior and acoustic emission locations of red sandstone under triaxial compression. Rock Mech Rock Eng. 2012;45(4):583–606. doi:10.1007/s00603-011-0208-8.
20. Moradian ZA, Ballivy G, Rivard P, Gravel C, Rousseau B. Evaluating damage during shear tests of rock joints using acoustic emissions. Int J Rock Mech Min Sci. 2010;47(4):590–8. doi:10.1016/j.ijrmms.2010.01.004.
21. Wu L, Wang J. Infrared radiation features of coal and rocks under loading. Int J Rock Mech Min Sci. 1998;35(7):969–76. doi:10.1016/S0148-9062(98)00007-2.
22. Ma L, Zhang Y, Cao K, Wang Z. An experimental study on infrared radiation characteristics of sandstone samples under uniaxial loading. Rock Mech Rock Eng. 2019;52(9):3493–500. doi:10.1007/s00603-018-1688-6.
23. Liu S, Wei J, Huang J, Wu L, Zhang Y, Tian B. Quantitative analysis methods of infrared radiation temperature field variation in rock loading process. Chin J Rock Mech Eng. 2015;34(S1):2968–76. doi:10.13722/j.cnki.jrme.2014.0656.
24. Shen R, Li H, Wang E, Chen T, Li T, Tian H, et al. Infrared radiation characteristics and fracture precursor information extraction of loaded sandstone samples with varying moisture contents. Int J Rock Mech Min Sci. 2020;130(5):104344. doi:10.1016/j.ijrmms.2020.104344.
25. Cao K, Ma L, Wu Y, Khan NM, Spearing AJS, Hussain S, et al. Cyclic fatigue characteristics of rock failure using infrared radiation as precursor to violent failure: experimental insights from loading and unloading response. Fatigue Fract Eng Mater Struct. 2021;44(2):584–94. doi:10.1111/ffe.13362.
26. Lin Y, Zheng Z, Zhou K, Gu Z, Pan Z. Research on the characteristics of infrared radiation and energy evolution law of red sandstone with different porosity during uniaxial compression. Sustainability. 2022;14(13):7698. doi:10.3390/su14137698.
27. Cai X, Zhou Z, Tan L, Zang H, Song Z. Water saturation effects on thermal infrared radiation features of rock materials during deformation and fracturing. Rock Mech Rock Eng. 2020;53(11):4839–56. doi:10.1007/s00603-020-02185-1.
28. Yuan Z, Ren F, Liu D. Experimental investigation on the infrared precursors of rockburst in sandstone with different bedding dip angles. Infrared Phys Technol. 2023;128(3):104518. doi:10.1016/j.infrared.2022.104518.
29. Tang Y, Liu J, Hao T, Li F, Zhao L. Failure mechanics and infrared radiation characteristics of soft coal at various moisture contents. Geomat Nat Hazards Risk. 2021;12(1):1371–85. doi:10.1080/19475705.2021.1920479.
30. Cao K, Dong F, Liu W, Khan NM, Cui R, Li X, et al. Infrared radiation denoising model of sub-region-Gaussian kernel function in the process of sandstone loading and fracture. Infrared Phys Technol. 2023;129(1):104583. doi:10.1016/j.infrared.2023.104583.

31. Cao K, Ma L, Wu Y, Khan NM, Yang J. Using the characteristics of infrared radiation during the process of strain energy evolution in saturated rock as a precursor for violent failure. Infrared Phys Technol. 2020;109(1):103406. doi:10.1016/j.infrared.2020.103406.
32. Cao K, Ma L, Wu Y, Spearing AJSS, Khan NM, Hussain S, et al. Statistical damage model for dry and saturated rock under uniaxial loading based on infrared radiation for possible stress prediction. Eng Fract Mech. 2022;260(1):108134. doi:10.1016/j.engfracmech.2021.108134.
33. Chen L, Yin S, Wang E, Qi H, Hou Y, Kang X. Coupled infrared radiation temperature and acoustic monitoring of damage characteristics on saturated red sandstone under uniaxial compression loading. Case Stud Therm Eng. 2025;73(3):106472. doi:10.1016/j.csite.2025.106472.
34. Poli R, Koza J. Genetic programming. In: Search methodologies: introductory tutorials in optimization and decision support techniques. Berlin/Heidelberg, Germany: Springer; 2014. p. 143–85. doi:10.1007/978-1-4614-6940-7_6.
35. Shahverdiloo MR, Zare S. A genetic programming model for estimating the rock mass deformation modulus based on analytical parameters. Bull Eng Geol Environ. 2025;84(7):388. doi:10.1007/s10064-025-04387-9.
36. Zhou J, Yazdani Bejarbaneh B, Jahed Armaghani D, Tahir MM. Forecasting of TBM advance rate in hard rock condition based on artificial neural network and genetic programming techniques. Bull Eng Geol Environ. 2020;79(4):2069–84. doi:10.1007/s10064-019-01626-8.
37. Özdemir 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. 2022;47(1):629–39. doi:10.1007/s13369-021-05761-x.
38. Oltean M, Grosan C. A comparison of several linear genetic programming techniques. Complex Syst. 2003;14(4):285–313. doi:10.25088/complexsystems.14.4.285.
39. Sonebi M, Cevik A. Genetic programming based formulation for fresh and hardened properties of self-compacting concrete containing pulverised fuel ash. Constr Build Mater. 2009;23(7):2614–22. doi:10.1016/j.conbuildmat.2009.02.012.
40. Kamdem EG, Tiam FFK, Ngueyep LLM, Wounabaissa O, Nnomo HRL, Kanmogne A. Prediction rotary drilling penetration rate in lateritic soils using machine learning models. Discov Civ Eng. 2025;2(1):53. doi:10.1007/s44290-025-00217-9.
41. Schapire RE. The strength of weak learnability. Mach Learn. 1990;5(2):197–227. doi:10.1007/BF00116037.
42. Shahani NM, Qin X, Xin W, Li J, Aizitiliwumaier T, Ma X, et al. Hybrid PSO with tree-based models for predicting uniaxial compressive strength and elastic modulus of rock samples. Front Earth Sci. 2024;12:1337823. doi:10.3389/feart.2024.1337823.
43. Ahmad Khan A, Chaudhari O, Chandra R. A review of ensemble learning and data augmentation models for class imbalanced problems: combination, implementation and evaluation. Expert Syst Appl. 2024;244(2):122778. doi:10.1016/j.eswa.2023.122778.
44. Jan MS, Hussain S, e Zahra R, Emad MZ, Khan NM, Rehman ZU, et al. Appraisal of different artificial intelligence techniques for the prediction of marble strength. Sustainability. 2023;15(11):8835. doi:10.3390/su15118835.
45. Shi K, Shi R, Fu T, Lu Z, Zhang J. A novel identification approach using RFECV–optuna–XGBoost for assessing surrounding rock grade of tunnel boring machine based on tunneling parameters. Appl Sci. 2024;14(6):2347. doi:10.3390/app14062347.
46. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci. 1997;55(1):119–39. doi:10.1006/jcss.1997.1504.
47. Ferreira AJ, Figueiredo MAT. Boosting algorithms: a review of methods, theory, and applications. In: Ensemble machine learning: methods and applications. Berlin/Heidelberg, Germany: Springer; 2012. p. 35–85. doi:10.1007/978-1-4419-9326-7_2.
48. Ceryan N. A review of soft computing methods application in rock mechanic engineering. In: Handbook of research on advanced computational techniques for simulation-based engineering. Hershey, PA, USA: IGI Global; 2016. p. 1–70.

49. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. Adv Neural Inf Process Syst. 2018;31:1–11. doi:10.48550/arxiv.1706.09516.
50. Liu Z, Jahed Armaghani D, Fakharian P, Li D, Vladimirovich Ulrikh D, Nikolaevna Orekhova N, et al. Rock strength estimation using several tree-based ML techniques. Comput Model Eng Sci. 2022;133(3):799–824. doi:10.32604/cmes.2022.021165.
51. Lin H, Zhang Q, Zhang L, Duan K, Xue T, Fan Q. The influence of water content on the time-dependent mechanical behavior of argillaceous siltstone. Rock Mech Rock Eng. 2022;55(7):3939–57. doi:10.1007/s00603-022-02826-7.
52. Noël C, Baud P, Violay M. Effect of water on sandstone's fracture toughness and frictional parameters: brittle strength constraints. Int J Rock Mech Min Sci. 2021;147(B7):104916. doi:10.1016/j.ijrmms.2021.104916.
53. Yao W, Li C, Zhan H, Zhou JQ, Criss RE, Xiong S, et al. Multiscale study of physical and mechanical properties of sandstone in three gorges reservoir region subjected to cyclic wetting–drying of Yangtze River water. Rock Mech Rock Eng. 2020;53(5):2215–31. doi:10.1007/s00603-019-02037-7.
54. Yu L, Yao Q, Chong Z, Li Y, Xu Q, Liu Z. Experimental study on the moisture migration and triaxial mechanical damage mechanisms of water-bearing coal samples. Int J Rock Mech Min Sci. 2022;160(2):105263. doi:10.1016/j.ijrmms.2022.105263.
55. Zhou XP, Jiang DC, Zhao Z. Digital evaluation of micro-pore water effects on mechanical and damage characteristics of sandstone subjected to uniaxial, cyclic loading–unloading compression by 3D reconstruction technique. Rock Mech Rock Eng. 2022;55(1):147–67. doi:10.1007/s00603-021-02662-1.
56. Abdulagaov IM, Abdulagatova ZZ, Kallaev SN, Omarov ZM. Heat-capacity measurements of sandstone at high temperatures. Geomech Geophys Geo Energy Geo Resour. 2019;5(1):65–85. doi:10.1007/s40948-018-0099-1.
57. Hou L, Cao K, Muhammad Khan N, Jahed Armaghani D, Alarifi S, Hussain S, et al. Precursory analysis of water-bearing rock fracture based on the proportion of dissipated energy. Sustainability. 2023;15(3):1769. doi:10.3390/su15031769.
58. Hou L, Ma L, Cao K, Muhammad Khan N, Feng X, Zhang Z, et al. Analysis of fracture characteristics of saturated sandstone based on infrared radiation variance. Phys Chem Earth Parts A/B/C. 2024;133(11):103517. doi:10.1016/j.pce.2023.103517.
59. Ali Shahmansouri A, Akbarzadeh Bengar H, Ghanbari S. Compressive strength prediction of eco-efficient GGBS-based geopolymer concrete using GEP method. J Build Eng. 2020;31(3):101326. doi:10.1016/j.jobe.2020.101326.
60. Shan T, Li Z, Zhang X, Niu Y, Tian H, Zhang Q, et al. Infrared radiation and acoustic emission of damage evolution and failure precursory for water-bearing coal. Rock Mech Rock Eng. 2022;55(12):7657–74. doi:10.1007/s00603-022-03042-z.
61. Li H, Kang T, Zhang R, Zhang X, Liang X, Zhu W, et al. Numerical study of the effects of loading parameters on high-energy gas fracture propagation in layered rocks with peridynamics. Eng Fract Mech. 2024;310:110516. doi:10.1016/j.engfracmech.2024.110516.
62. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:1–10.